Wireless Communications and Mobile Computing

Detectable, Traceable, and Manageable Blockchain Technologies

Lead Guest Editor: Yin Zhang
Guest Editors: Jeungeun Song and Iztok Humar

Detectable, Traceable, and Manageable Blockchain Technologies

Wireless Communications and Mobile Computing

Detectable, Traceable, and Manageable Blockchain Technologies

Lead Guest Editor: Yin Zhang
Guest Editors: Jeungeun Song and Iztok Humar

Chief Editor

Associate Editors

Academic Editors

Contents

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 2381365, 13 pages
https://doi.org/10.1155/2022/2381365

Research Article

A Blockchain-Based Privacy-Preserving Publish-Subscribe Model in IoT Multidomain Data Sharing

Zhendong Liu,[1] Liang Meng,[1] Qingyuan Zhao,[1] Fei Li,[1] Manrui Song,[1] Dongxu Dai,[1] Xiujuan Yang,[1] Song Guan,[1] Yue Wang,[1] and Hongliang Tian[2]

[1]*State Grid Benxi Electric Power Supply Company, Benxi 117000, China*
[2]*Jilin Northeast Electric Power University Science and Technology Development Co., Ltd., Jilin 132000, China*

Correspondence should be addressed to Hongliang Tian; hltian@foxmail.com

Received 24 June 2022; Accepted 18 September 2022; Published 14 October 2022

Academic Editor: Yin Zhang

With the dramatically increasing deployment of intelligent devices, the Internet of Things (IoT) has attracted more attention and developed rapidly. It effectively collects and shares data from the surrounding environment to achieve better IoT services. For data sharing, the publish-subscribe (PS) paradigm provides a loosely coupled and scalable communication model. However, due to the loosely coupled nature, it is vulnerable to many attacks, resulting in some security threats to the IoT system, but it cannot provide the basic security mechanisms such as authentication and confidentiality to ensure the data security. Thus, in order to protect the system security and users' privacy, this paper presents a secure blockchain-based privacy-preserving access control scheme for the PS system, which adopt the fully homomorphic encryption (FHE) to ensure the confidentiality of the publishing events and leverage the ledger to store the large volume of data events and access crossdomain information. Finally, we analyze the correctness and security of our scheme; moreover, we deploy our proposed prototype system on two computers and evaluate its performance. The experimental results show that our PS system can efficiently achieve the equilibrium between the system cost and the security requirement.

1. Introduction

With the rapid development of Internet of Things (IoT) in recent years, IoT devices deployed in application scenarios such as smart grid, smart city and smart home have increased sharply [1–3]. It was estimated that there will be over 24.9 billion IoT devices connected to the Internet by 2025 [4]. These interconnected mass terminal devices store and forward data to better realize system functions. As an attractive communication paradigm, publish-subscribe (PS) system can be used to build distributed data sharing across the Internet by separating the sender from the receiver. However, due to the loose coupling between publishers and subscribers, it is a challenge to provide security mechanisms such as authentication and confidentiality among each domain of the IoT [5]. Thus, we need to find out a method to ensure the data is only delivered to eligible subscribers who are interested and protect the confidentiality of the published events and the privacy of sensitive information in the process [6, 7].

Access control technology can protect the confidentiality, integrity, and availability of PS service and user data in the traditional IoT PS system. However, the traditional access control schemes cannot be used directly to provide fine-grained and scalable requirements for publish-subscribe systems [8]. The original publish-subscribe model relies on a trusted third-party broker such as MQTT [9], LooCI [10], and NesC [11], where data from all devices flows to subscribers through a central broker. Such a centralized architecture makes the PS model have the following disadvantages:

(i) The centralized architecture is vulnerable to a single point of failure. Since the broker is a centralized server, which coordinates the communication between the publishers and subscribers, if the server fails or is attacked by a malicious adversary, it may cause a large amount of sensitive information be compromised, thus threatening the privacy of the users and even making the whole system down

(ii) The semitrusted broker may be immoral, and it may lead to unauthorized access, abuse, and tampering with data

(iii) Since centralized servers rely on computationally greedy encryption algorithms, this is not suitable for computing resources-constrained IoT devices

Therefore, a novel decentralized PS model needs to be designed to address these issues. Due to the advantages of decentralization, anonymity and nontampering of records of blockchain [12, 13], it can provide reliable subscription record storage, subscription content forwarding, and subscription information verification for the PS system. The application of blockchain in the PS system has the following benefits:

(i) Decentralization: the published encrypted data and the subscription records are stored in blocks in the distributed ledger, and the consistency of network records is maintained through the consensus mechanism. Due to the decentralized nature of blockchain, it can increase the fault tolerance and antiaggression of the system, thus avoiding the impact of a single point of failure

(ii) Anonymity: all subscription contents are stored in the blockchain in an encrypted way, and the subscriber can access the data through its public key address. However, malicious users can only link to the public key address through hash pointer but do not know the real identity of the users

(iii) Nontampering: the subscription information is added to the blockchain after consensus verification, and then it will be recorded by all nodes together and related to each other through cryptography; so, tampering the data is very difficult and expensive

In order to solve the mentioned challenges in the PS system, this paper designs a novel blockchain-based PS model and proposes an access control mechanism based on the fully homomorphic encryption (FHE) algorithm [14] to protect the privacy of data sharing among multiple domains in the IoT. The proposed model mainly includes four entities: publishers, subscribers, broker based on private blockchain, and consortium blockchain, where publisher is responsible for publishing specific encrypted data, and subscriber receives related content by subscribing to the interested topics. Each broker based on private blockchain is composed of multiple distributed and decentralized gateway devices, and it only serves a subset of IoT devices to match user needs, delivers subscription content, and stores the subscription records, whereas the consortium blockchain connects private blockchain to facilitate crossdomain data sharing.

It is noteworthy that with the dramatically increasing of mobile services and applications, the broker needs to be equipped with more computing and storage capacity, but IoT devices are usually resource constrained, and they cannot bear the resource consumption caused by complex verification calculation of blockchain; so, we mitigate this problem by using edge computing. Edge computing utilizes nearby edge servers to bring real-time computations and communications [13, 15, 16]. As one way to process data at the network edge, it greatly expands the capacity and feasibility of terminal devices. In our model, we make full use of the private blockchain that has been formed through the gateway in [17], and then use the edge servers to create the consortium blockchain and perform FHE. By this way, it can provide publishers and subscribers with effective privacy protection. Our contributions are as follows:

(i) We propose a blockchain-based PS model for data sharing among multiple domains of IoT. This model eliminates the disadvantages of traditional PS model based on centralized broker and can make full use of consortium blockchain to carry out cross-domain subscription services in the large-scale IoT

(ii) We combine edge computing to provide computing power for data validation and all cryptographic computations and make it possible to deploy blockchain in the resource-constrained IoT. In addition, the cryptographic accumulator is used to quickly verify whether the subscription information on the one private blockchain is valid or not, which reduces the cost and latency of cross-domain data sharing

(iii) We use FHE with IND-CPA security to realize the attribute-based access control mechanism, so that the edge servers can perform arbitrary calculation of ciphertext without decryption, in this way, while ensuring the confidentiality and privacy of the subscription information and realizing the fine-grained access control of user data

The rest of this paper is organized as follows. Section 2 introduces some related work and briefly analyzes the pros and cons of various solutions. Section 3 reviews the preliminaries used in this paper. In Section 4, we present a blockchain-based privacy-preserving PS model. Section 5 analyses the performance and security of our scheme by deploying it on two computers. Finally, we summarize the paper with a further research discussion.

2. Preliminaries

In this section, we review some of the relevant theoretical basis of this study and briefly introduce and analyze the related background technologies, which mainly include the concepts of publish-subscribe system, attribute-based authorization, blockchain, fully homomorphic encryption, and edge computing.

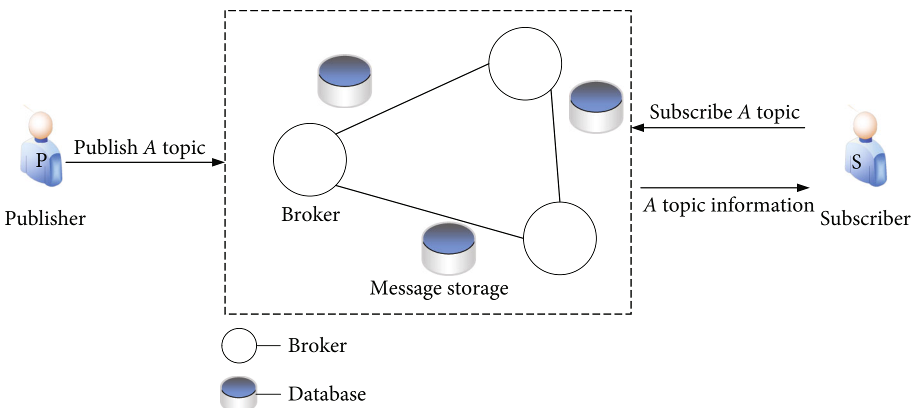


Figure 1: Publish-subscribe system architecture.

2.1. Publish-Subscribe System. Publish-subscribe system can be seen as a way of data-centric message distribution [18]. During the distribution of a message, the publisher can publish the message without specifying the identity of the user, and the subscriber also does not need to know the identity of the data owner to use message. In such a middleware solution, a message is represented as an event that can be detected in the application. As is shown in Figure 1, the PS model relies on three elements: publisher, subscriber, and the broker.

In the model, a publisher is an actor who generates any content and publishes it to the specified topic; subscriber is a user of events who subscribes the interested topics, and subscriber gets the published event when a publisher creates a publication for its subscription request. The broker is responsible for receiving the published events and notifying subscribers of the interested topics.

2.2. Attribute-Based Authorization [19]. An attribute A is defined as $A = (\mathrm{st}, \mathrm{value})$, meaning that the attribute st have value. A user has one attribute A that can be represented by conjunctive formula $A_1 \Lambda A_2 \Lambda \cdots \Lambda A_t$. For a given system event topic tp, authorization policy restricts access to event data with a tp topic by using a user's specific attribute value.

Definition 1. The expression for an authorization policy is $\Lambda_{\mathrm{tp}} = (A_{11} \Lambda A_{12} \Lambda \cdots \Lambda A_{1t})\mathrm{V} \cdots \mathrm{V}(A_{s1} \Lambda A_{s2} \Lambda \cdots \Lambda A_{st})$, which means that when a subscriber has at least a set of attributes from attribute concatenation $A_{11} \Lambda A_{12} \Lambda \cdots \Lambda A_{1t}$ to $A_{s1} \Lambda A_{s2} \Lambda \cdots \Lambda A_{st}$, the subscriber can access the data with topic tp.

For a subscriber whose attribute expression is $\omega = (A_{11}{}' \Lambda A_{12}{}' \Lambda \cdots \Lambda A_{1t}{}')\mathrm{V} \cdots \mathrm{V}(A_{h1}{}' \Lambda A_{h2}{}' \Lambda \cdots \Lambda A_{ht}{}')$, he/she has h group connection attributes. As long as one of the h group conjunctive attributes appears in Λ_{tp}, then ω is defined to satisfy Λ_{st}.

2.3. Blockchain and Edge Computing. Since Nakamoto [12] published the Bitcoin white paper in 2008, the blockchain, as the underlying technology of Bitcoin, has quickly attracted a lot of attention due to its characteristics such as decentralization, no tampering, public verification and anonymity. The blockchain works as a distributed database that records all transactions that have occurred in the peer-to-peer (P2P) network. As is shown in Figure 2, the blockchain is a series of blocks connected one by one by hash. Blocks are added to the longest main blockchain by consistency protocol among most nodes in the network. Each block contains two parts: block header and block body, where all transactions involved in the block body, and the block header consists of the link pointers of the previous block header, a Merkle root of all transactions and a timestamp. Hyperledger Fabric [13, 20, 21] is a consortium blockchain based on distributed ledger. Unlike public or private blockchain, it executes the verification of transactions by a set of preselected nodes in the consortium blockchain, and the nodes can change dynamically; so, the consortium blockchain is more suitable for the scenario that supports node scalability.

Due to the limited computing capacity and available energy consumption of IoT terminal device, it has become the key bottleneck restricting the application of blockchain in IoT, but edge computing can help mitigate this problem. Edge computing transfers data processing from the remote cloud center to the edge of the network, and the computation and data storage can be dispersed to the edge of the Internet near the endpoint of things, sensors, and users. It brings real-time computation and communication by leveraging nearby edge servers.

2.4. Fully Homomorphic Encryption [14]. Let q be prime, $\mathbb{Z}_q$ be the integer field of modulo q, and n be an integer. For the given plaintext $v \in \mathbb{Z}_q$ and the key K generated by the parameters q and n, there are encryption function $\mathrm{Enc}(K, v) = (c_1, c_2, \cdots, c_n)$ and decryption function $\mathrm{Dec}(K, (c_1, c_2, \cdots, c_n)) = v$, where ciphertext $(c_1, c_2, \cdots, c_n)$ is an n-dimensional vector. Public key PK generated by key K can be used to encrypt v, and then

$$\begin{aligned} \mathrm{Enc}(\mathrm{PK}, v) &= (c_1, c_2, \cdots, c_n), \\ \mathrm{Dec}(K, (c_1, c_2, \cdots, c_n)) &= v. \end{aligned} \tag{1}$$

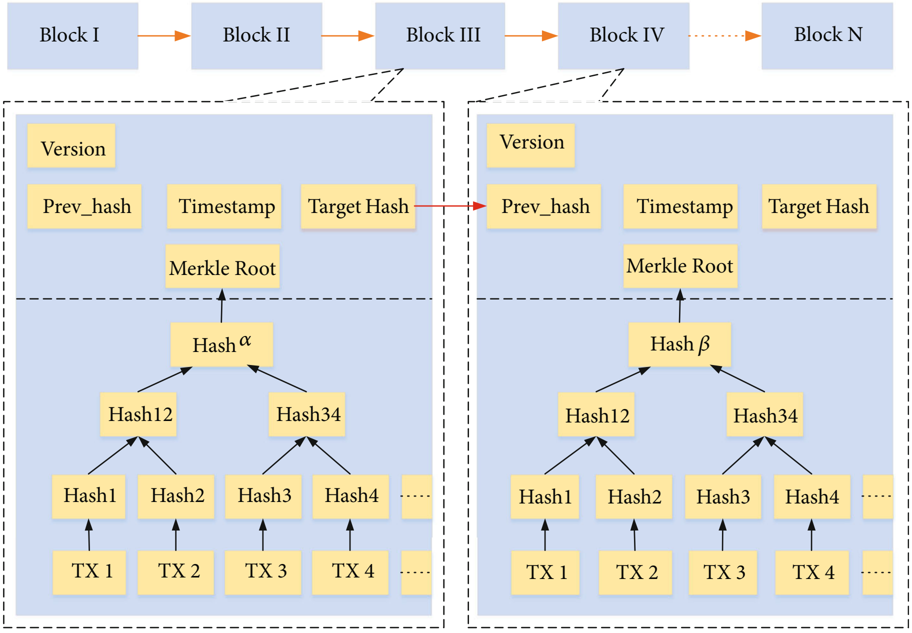


Figure 2: Blockchain structure.

Let $C = (c_1, c_2, \cdots c_n)$ and $C' = (c_1{}', c_2{}', \cdots, c_n{}')$. When $\mathrm{Dec}(K, C) = v$ and $\mathrm{Dec}(K, C') = v'$ exist in the decryption function, the FHE algorithm satisfies the following additional homomorphism properties:

$$\begin{aligned} \mathrm{Dec}\left(K, C \oplus C'\right) &= v + v' (\mathrm{mod}\ q), \\ \mathrm{Dec}(K, d \square C) &= d * v (\mathrm{mod}\ q), \end{aligned} \tag{2}$$

where ⊕ is vector addition, and □ is scalar multiplication of vectors.

The homomorphic operation of multiplication also requires the public evaluation key $\mathrm{PEK}_{ij} (1 \le i \le n, 1 \le j \le n)$, which is generated by K. For $v * v'$ obtained from ciphertext C and C', it can be expressed as

$$\left(\left(c_1 * c_1{}'\right) \square \mathrm{PEK}_{11}\right) \oplus \cdots \oplus \left(\left(c_i * c_j{}'\right) \square \mathrm{PEK}_{ij}\right) \oplus \cdots \oplus \left(\left(c_n * c_n{}'\right) \square \mathrm{PEK}_{nn}\right). \tag{3}$$

For a given publisher's secret key sk_p and subscriber's public key pk_s, the ciphertext encrypted with sk_p can be converted to the ciphertext encrypted with subscriber's secret key sk_s. The key exchange process is as follows:

Let $\mathrm{KeySwitch}(\mathrm{pk}_s, \mathrm{sk}_p)$ be the generating function of exchange key KS, and then $\mathrm{KS} = \{\mathrm{KS}_1, \mathrm{KS}_2, \cdots, \mathrm{KS}_n\}$, where any KS_i is an n-dimensional vector. Suppose there is $\mathrm{Decrypt}(\mathrm{sk}_p, (c_1, c_2, \cdots, c_n)) = v$, then the reencryption of ciphertext C with exchange key KS can be expressed as $\mathrm{ReEnc}(\mathrm{KS}, C) = (c_1 \square \mathrm{KS}_1) \oplus (c_2 \square \mathrm{KS}_2) \oplus \cdots \oplus (c_n \square \mathrm{KS}_n),$, let $C' = \mathrm{ReEnc}(\mathrm{KS}, C)$, and then $\mathrm{Dec}(sk_s, C') = v$.

3. Related Work

In recent years, most of the research on PS system has focused on effective event routing, event filtering, and composite event detection, and little has been done to address privacy issues. Here, we briefly summarize some relevant work in recent years and find that it can be divided into two categories: (1) PS system based on traditional broker server and (2) PS system based on P2P (peer-to-peer) network. This section mainly analyzes the current research status of privacy-preserving PS system.

3.1. Based on Traditional Broker Servers.

Duan et al. [22] proposed a comprehensive access control framework CACF to guarantee the data confidentiality and service privacy of the publish-subscribe model in different domains. It uses fully homomorphic encryption to encrypt data and bidirectional privacy-preserving policy to match access policies and subscription policy. We can see from the performance analysis result that the CACF scheme can provide confidentiality and privacy-preserving with acceptable latency, but the centralized message-oriented Java Message Service (JMS) broker can cause a single point of failure.

AKPS [23] is a privacy-preserving attribute-keyword-based data publish-subscribe scheme. This scheme uses attribute-based encryption with decryption outsourcing to encrypt the published data. While realizing the publisher's own control of data access, it transfers the main decryption overhead from subscribers to the cloud server. And subscribers who search by keyword can choose to receive the data according to their own interests. However, the publisher has only one identity; that is, it cannot receive the information as a subscriber.

In [24], Wang et al. proposed a privacy protection scheme for a content-based publish/subscribe system with

Table 1: The comparison with other schemes.

Scheme	Confidentiality	Decentralized	Privacy	Fine-grained access	Against collusion attack	Against spoofing attacks
Duan et al. [22]	✓	—	✓	✓	✓	✓
Yang et al. [23]	✓	—	✓	✓	—	—
Wang et al. [24]	✓	—	✓	—	✓	—
Diro et al. and Diro et al. [25, 26]	✓	—	✓	—	—	—
Borcea et al. [27]	✓	—	✓	—	—	—
Zhao et al. [28]	✓	✓	✓	—	✓	✓
Lv et al. [29]	✓	✓	✓	—	—	✓
Tariq et al. [30]	✓	✓	—	—	—	—
Our scheme	✓	✓	✓	✓	✓	✓

differential privacy in a fog computing environment. It used the U-Apriori algorithm to extract the collection of the first K frequent items from uncertain data sets and then applied the exponential and Laplace mechanism to ensure differential privacy. Brokers mine the first K item sets to eventually match the appropriate publishers and subscribers. This method reduces the cost of user computation and storage, but the complex attribute matching method increases the delay of matching time and increases with the number of users.

In order to provide basic security mechanisms for fog computing-based publish-subscribe system in IoT, Diro et al. [25] proposed a secure lightweight publish-subscribe protocol based on elliptic curve cryptography (ECC). It reduces the overhead of computations, storage, and communications in traditional security protocols such as SSL/TSL. In [26], Diro et al. proposed a resource efficient end-to-end security scheme by offloading computations and storage of security parameters to fog nodes in the vicinity. In addition, a symmetric-key payload encryption has been used to minimize the overhead of message communication in the resource-contested IoT environment.

Borcea et al. [27] introduced PICADOR, a topic-based publish-subscribe system designed using proxy reencryption. This system provides end-to-end encrypted information distribution service, and it ensures the information confidentiality between publishers and subscribers without sharing encryption and decryption keys. The system not only reduces the communication cost but also reduces the vulnerability of internal attack. However, reencryption also brings a heavy computing burden to proxy server.

3.2. Based on P2P Network. Zhao et al. [28] built a fair and secure publish-subscribe system (SPS) based on blockchain. In SPS, in order to realize fair data exchange, publishers publish a topic on the blockchain, and subscribers subscribe the interested topic by deposit. At the same time, the publisher and subscriber use hybrid encryption to ensure data confidentiality and take advantage of the pseudoanonymity of bitcoin system to ensure the identity privacy of both parties. However, because this scheme cannot provide fine-grained access control, it cannot provide users with more accurate and efficient services according to their own features.

In [29], Lv et al. propose a privacy-preserving publish/subscribe model by using the blockchain technique, which ensures the system confidentiality by employing public key encryption with equality test (PKEwET), and they solved the single point of failure and the anonymity of the participants by using the Ethereum.

Tariq et al. [30] proposed a new approach to provide authentication and confidentiality in broker-less content-based publish/subscribe system. Credentials are assigned to publishers and subscribers by adapting the pairing-based cryptography mechanisms. Because the private keys and ciphertext assigned to publishers and subscribers are marked with credentials, a particular subscriber can decrypt an event only if the credentials associated with the event match the private key. However, Tariq et al. do not consider the anonymity of subscriber.

In [31], the authors contributed Trinity, a novel distributed publish-subscribe broker with blockchain-based immutability. It distributes the published data to all brokers in the network and stores the distributed data in an immutable ledger by using the blockchain technology. In this way, it can guarantee persistence, ordering, and immutability across trust boundaries, but the Trinity framework increases the end-to-end delay while consuming bandwidth and computation resources.

Gao et al. [32] proposed a new trustworthy secure ciphertext-policy and attribute hiding access control scheme based on blockchain, named TrustAccess, to achieve trustworthy access. To address the privacy issues of access policy and user attribute in the TrustAccess, an optimized hidden policy CP-ABE named OHP-CP-ABE to ensure policy privacy while satisfying the large universe access requirement. In addition, the authors use the multiplicative homomorphic ElGamal cryptosystem to ensure the attribute privacy during authorization validation.

4. BPAC System Model

In this section, we mainly explain how the proposed blockchain-based IoT publish-subscribe system works. For convenience, some notations will appear in our BPAC scheme as shown in Table 1.

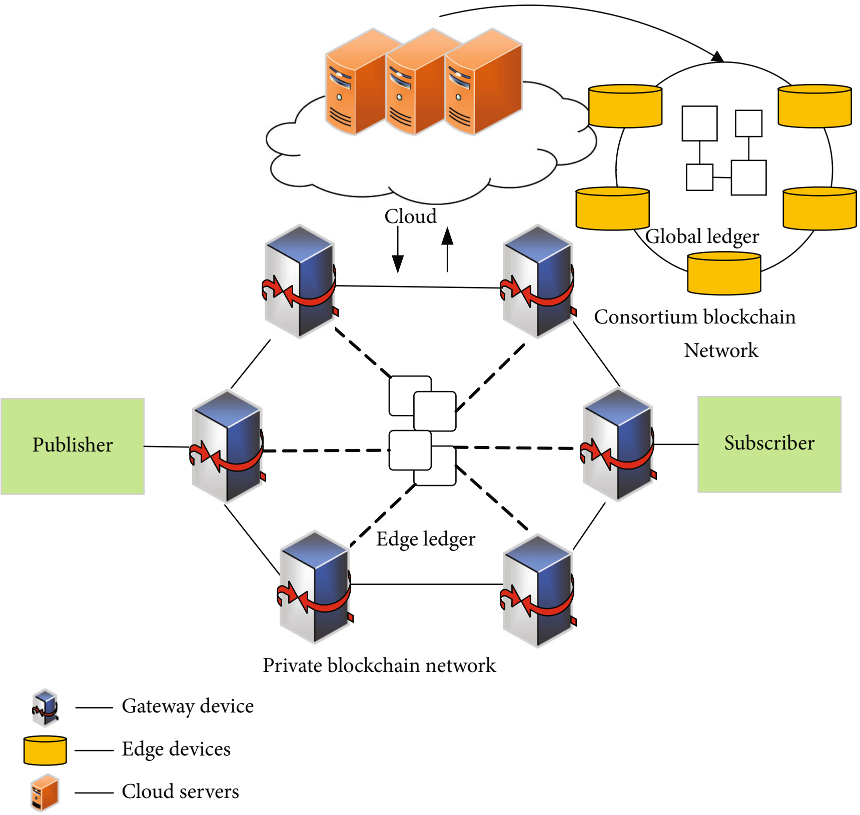


Figure 3: Security access control system model.

4.1. Security Model. In our work, we assume the certificate authority (CA) that creates the public/private keys for the publisher or subscriber and assigns public parameters to the system is honest; that is, the CA follows the rules to perform computations. And the publisher who can correctly and truly publish the encrypted data is legal. All published events are stored in the global ledger maintained by the edge devices, and all data validation and publish-subscribe services processing are performed by the edge devices to reduce the workload of an IoT device. It is worth emphasizing that the storage and protection of the published events are only performed by blockchain, without intervention of any other entity. Therefore, the security of our scheme is mainly guaranteed by blockchain. In our scheme, publishers and subscribers within the domain directly interact with each other through private blockchain, and the crossdomain users connect private blockchain through consortium blockchain for temporary crossdomain information interaction. In the actual collaborative IoT services, there may have a many-to-many relationship among multiple publishers and subscribers. Here, we just take one publisher and one subscriber to discuss the access control procedure in our framework. The system model is shown in Figure 3.

4.2. Blockchain-Based Security Publish-Subscribe System. We propose a secure PS scheme which is based on FHE [14]. Assume that a publisher P contains a key pair $(\mathrm{PK}_p, \mathrm{SK}_p)$, and a subscriber S contains a key pair $(\mathrm{PK}_s, \mathrm{SK}_s)$. The specific dynamic data flow is shown in Figure 4. The access control procedure mainly contains the following phases: Setup, Publish, Subscribe, Match, and Receive.

4.2.1. Setup. The setup algorithm takes the security parameter λ, a number of levels L, and $b \in \{0, 1\}$ as input parameters to generate the system parameter $\mathrm{Params} = (q, d, n, N, \chi)$. This algorithm is run by CA, and only CA knows the value of Params, where let $\mu = \mu(\lambda, L, b)$, whose modulus is prime q, and $d = d(\lambda, \mu, b)$, $n = n(\lambda, \mu, b)$, $N = N(\lambda, \mu, b)$, and $\chi = \chi(\lambda, \mu, b)$. Finally, the key pair PK and SK are generated as follows:

$$\begin{aligned} \mathrm{SecretKeyGen}(\mathrm{params}) &\longrightarrow \mathrm{SK}, \\ \mathrm{PublicKeyGen}(\mathrm{params}) &\longrightarrow \mathrm{PK}, \end{aligned} \tag{4}$$

where the key pair of publisher and subscriber is, respectively, $(\mathrm{PK}_p, \mathrm{SK}_p)$ and $(\mathrm{PK}_S, \mathrm{SK}_S)$.

4.2.2. Publish. The publisher randomly selects random number $r_{\mathrm{pp}}, r_{\mathrm{up}}, r_{\mathrm{ac}}$ and hash function h in advance, where r_{pp} is greater than the number of topics in the publishing event e_{tp}, then generates $h_{\mathrm{up}} = h(\mathrm{A}_{i1} \| \mathrm{A}_{i2} \| \cdots \| \mathrm{A}_{\mathrm{im}} \| r_{\mathrm{up}})$, and encrypts event e_{tp} with topic tp and policy $\Lambda_{\mathrm{tp}} = (\mathrm{A}_{11} \Lambda \mathrm{A}_{12} \Lambda \cdots \Lambda \mathrm{A}_{1t}) \mathrm{V} (\mathrm{A}_{s1} \Lambda \mathrm{A}_{s2} \Lambda \cdots \mathrm{A}_{\mathrm{st}})$ as C_{tp} through edge servers. For each set of attribute conjunction formula $\mathrm{A}_{i1} \Lambda \mathrm{A}_{i2} \Lambda \cdots \Lambda \mathrm{A}_{im} (1 \le i \le n)$, the publisher generates F_s through the attribute filter function $F(\mathrm{A}_{i1} \Lambda \cdots \Lambda \mathrm{A}_{im})$, uses the edge servers

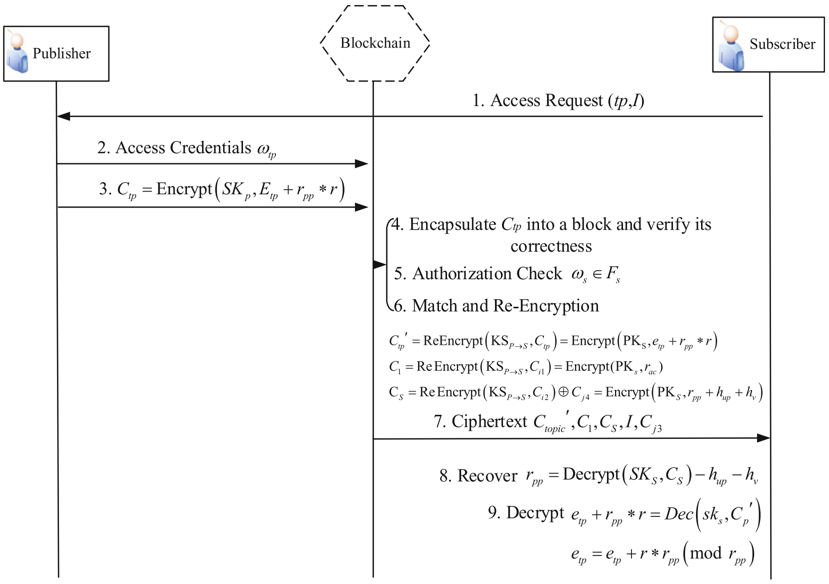

Publisher
Blockchain
Subscriber
1. Access Request (tp,I)
2. Access Credentials ω_{tp}
3. $C_{tp} = \text{Encrypt}\left(SK_p, E_{tp} + r_{pp} * r\right)$
4. Encapsulate C_{tp} into a block and verify its correctness
5. Authorization Check $\omega_s \in F_s$
6. Match and Re-Encryption
$C_{tp}{}' = \text{ReEncrypt}\left(\text{KS}_{P\rightarrow S}, C_{tp}\right) = \text{Encrypt}\left(\text{PK}_S, e_{tp} + r_{pp} * r\right)$
$C_1 = \text{Re Encrypt}\left(\text{KS}_{P\rightarrow S}, C_{i1}\right) = \text{Encrypt}(\text{PK}_s, r_{ac})$
$\text{C}_S = \text{Re Encrypt}\left(\text{KS}_{P\rightarrow S}, C_{i2}\right) \oplus C_{j4} = \text{Encrypt}\left(\text{PK}_S, r_{pp} + h_{up} + h_v\right)$
7. Ciphertext $C_{topic}{}', C_1, C_S, I, C_{j3}$
8. Recover $r_{pp} = \text{Decrypt}\left(SK_S, C_S\right) - h_{up} - h_v$
9. Decrypt $e_{tp} + r_{pp} * r = Dec\left(sk_s, C_p{}'\right)$
$e_{tp} = e_{tp} + r * r_{pp}\left(\text{mod } r_{pp}\right)$

FIGURE 4: Interactive time sequence in our scheme.

to convert it into access credentials:

$$\omega_{\text{topic}} = \begin{pmatrix} \text{KS}_{P\longrightarrow S}, \{(C_{11}, C_{12}, F_1), (C_{21}, C_{22}, F_2), \cdots, (C_{s1}, C_{s2}, F_s)\} \\ \{(C_{13}, C_{14}), (C_{23}, C_{24}), \cdots, (C_{h3}, C_{h4})\} \end{pmatrix}, \tag{5}$$

and finally publishes F_s and C_{tp} on a private blockchain. The encryption process for publishing events is as follows:

$$\begin{aligned} C_{i1} &= \text{Encrypt}\left(\text{SK}_P, r_{up}\right), \\ C_{i2} &= \text{Encrypt}\left(\text{SK}_P, h_{\text{up}} + r_{\text{pp}} - r_{\text{ac}}(h(A_{i1}) + h(A_{i2}) + \cdots + h(A_{im}))\right), \\ C_{j3} &= \text{Encrypt}(PK_S, r_S), \\ \text{C}_{j4} &= \text{Encrypt}\left(PK_S, h_v + r_{\text{ac}}\left(h\left(A_{j1}{}'\right) + h\left(A_{j2}{}'\right) + \cdots + h\left(A_{jm}{}'\right)\right)\right). \end{aligned} \tag{6}$$

When the private blockchain receives the encrypted event C_{tp}, the edge servers packaged it into a block and stored in the edge ledger after being authenticated by the whole network.

4.2.3. Subscribe. First, the subscriber S with property expression $\omega_s = (\text{A}_{11}{}'\Lambda\text{A}_{12}{}'\Lambda \cdots \Lambda\text{A}_{1t}{}')\text{V} \cdots \text{V}(\text{A}_{h1}{}'\Lambda\text{A}_{h2}{}'\Lambda \cdots \text{A}_{ht}{}')$ subscribes to an interested topic through edge ledger, and then subscriber encrypts its property index value j to $I = \text{Encrypt}(PK_s, j)$ and finally sends it to the private blockchain broker.

4.2.4. Match and Key Switching. When the publisher receives a subscription request from the subscriber, it first checks whether subscriber's attribute conjunction ω_s satisfies $\omega_s \in F_s$. If the condition is met, the subscriber is certified as a valid user, and his subscription request is allowed. Then, the publisher will reencrypt the ciphertext C_{tp}, C_{i1}, C_{i2} through edge servers to $C_{tp}{}', C_1, C_s$. The conversion process is as follows:

$$\begin{aligned} C_{\text{tp}}{}' &= \text{ReEncrypt}\left(\text{KS}_{P\longrightarrow S}, C_{\text{tp}}\right) = \text{Encrypt}\left(\text{PK}_\text{S}, e_{\text{tp}} + r_{\text{pp}} * r\right), \\ C_1 &= \text{Re Encrypt}(\text{KS}_{P\longrightarrow S}, C_{i1}) = \text{Encrypt}(\text{PK}_s, r_{ac}), \\ \text{C}_S &= \text{Re Encrypt}(\text{KS}_{P\longrightarrow S}, C_{i2}) \oplus C_{j4} = \text{Encrypt}\left(\text{PK}_\text{S}, r_{\text{pp}} + h_{\text{up}} + h_v\right). \end{aligned} \tag{7}$$

Finally, the publisher authorizes the subscriber S to access $C_{tp}{}', C_1, C_s, I$ and C_{j3} from the edge ledger.

If subscriber S fails to meet the requirement, the edge servers simply refuse the subscriber's access requests.

4.2.5. Receive. After subscriber S receives $C_{tp}{}', C_1, C_s, I$ and C_{j3}, it first decrypts I to obtain index j, thus obtaining the authorization attribute conjunction $\omega_j = A_{j1}{}'\Lambda A_{j2}{}'\Lambda \cdots \Lambda A_{jm}{}'$. Then it decrypts C_{j3} and C_1 to get the random values r_s and r_{ac}. Then, the subscriber uses hash function h to restore r_{pp}:

$$\begin{aligned} h_{\text{up}} &= h\left(\text{A}_{j1}{}' \| \text{A}_{j2}{}' \| \cdots \| \text{A}_{\text{jm}}{}' \| r_{\text{up}}\right), \\ h_v &= h\left(\text{A}_{j1}{}' \| \text{A}_{j2}{}' \| \cdots \| \text{A}_{\text{jm}}{}' \| r_S\right), \\ r_{\text{pp}} &= \text{Decrypt}(\text{SK}_S, C_S) - h_{\text{up}} - h_v. \end{aligned} \tag{8}$$

Finally, the subscriber decrypts the ciphertext $C_{\text{tp}}{}'$ and gets $e_{\text{tp}} + r_{\text{pp}} * r$, and the modular operation is then performed on r_{pp} to recover the event e_{tp}.

4.3. Efficient Crossdomain Access and Authentication. For the crossdomain PS system, there is no direct connection among edge ledgers, and no copies of other ledgers are

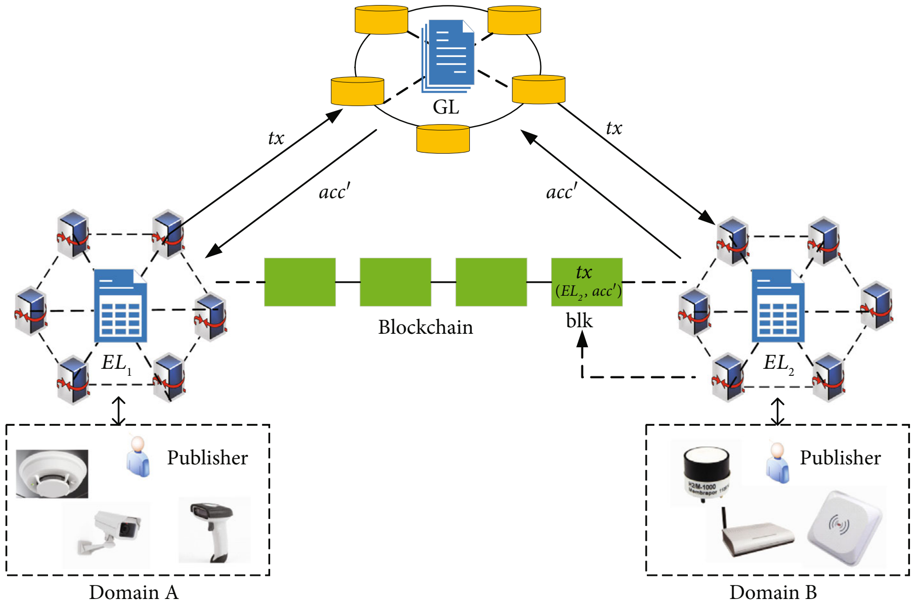


FIGURE 5: Crossdomain data verification.

kept. Therefore, after obtaining the authorization information, the subscriber needs to verify whether the authorization information block belonging to another edge ledger is valid.

Assume that EL_1 and EL_2 are two subscribers of edge ledger in different domains. EL_1 needs to access the publishing events in EL_2 through the global ledger GL and verifies its validity. The verification process after obtaining the authorization information block is shown in Figure 5.

(1) EL_2 processes the new authorization information block tx

(i) EL_1 initiates a verification request for information block tx to the global ledger GL. GL forwards it to EL_2 and EL_2 initializes the value acc of the accumulator after receiving the verification request

(ii) EL_2 packs tx into a new block blk and updates the accumulator value to acc'

(iii) All nodes el_{2j} in EL_2 run the consensus protocol to add blk and update accumulator value acc' to the blockchain

(2) EL_2 updates its status to GL

(i) EL_2 only updates the accumulator value to GL after a certain number of new blocks are created

(ii) GL checks whether EL_2 has achieved consensus on acc', if it passes the check, then the latest state of $(\mathrm{EL}_2, \mathrm{acc}')$ is included in the new block

(3) EL_1 checks the validity of tx

(i) EL_1 obtains the current accumulator value of EL_2 from GL

(ii) EL_1 requests EL_2 to provide evidence that block blk contains the authorization information block tx

(iii) EL_2 responses to EL_1's request and provides a proof that blk is included in the edge ledger EL_2

EL_1 verifies the evidence. After verification, it can utilize the information in tx.

5. Security and Performance Analysis

In this section, we first theoretically analyze the security of the proposed scheme and illustrate the correctness of our scheme, where our scheme only aims to resist collusion attack and spoofing attacks. Then, we implement the prototype system to evaluate its performance.

5.1. Security Analysis

5.1.1. Confidentiality. For our proposed publish-subscribe scheme, the security of data sharing is based on the security of blockchain and FHE algorithm. Among them, since the FHE is IND-CPA secure, that is to say, an adversary first gets a properly generated pk, then specifies message

$m_0, m_1 \in R_M$ (R_M is a message ring), and finally gets $\mathrm{Enc}_{pk}(m_b)$ for a random number b; it cannot guess the value of b with probability $>1/2+\varepsilon(\lambda)$, where ε is a negligible function in the security parameter λ. In other words, for a given ciphertext, an adversary is not able to know any useful information about the corresponding plaintext; that is, it is secure against chosen-plaintext attack. And we adopted the FHE algorithm to set up a credible PS system for IoT, which can separate data processing rights and data ownership, so as to prevent data privacy leakage while using edge servers computing power. In addition, blockchain lies on the hardness of preventing sibyl attacks and DDoS attacks. In the large-scale IoT environments, with more IoT devices connected to the blockchain network, the more gateway nodes in the network increases, and the more security will be improved; so, it is difficult for an attacker to launch a DDoS attacks in the blockchain network. This is because if you want to launch 51% attacks in the blockchain network, you need a lot of computing power to control the nodes that are distributed everywhere, since an adversary is not powerful enough to take over the majority of the nodes. Therefore, the scheme can guarantee the confidentiality of the message.

5.1.2. Resistance to Collusion Attack. For two collusive subscribers S_1 and S_2, they cannot successfully pass the inspection of the property filter function F in the edge servers, because neither of them has the authentication attribute authorized by the access control policy. Even if the edge servers are malicious and also participate in the collusion attack, consequently, make both pass the inspection and convert keys to generate $C_p{''}, C_1{'}, C_s{'}, I', C_{j3}{'}$ and $C_p{'''}$, $C_1{'}, C_s{''}, I'', C_{j3}{''}$. However, S_1 and S_2 will only get the following ciphertext:

$$C_S{'} = \mathrm{Encrypt}\left(\begin{array}{c} \mathrm{PK}_{S_1}, r_{\mathrm{pp}} + h_{\mathrm{up}}{'} + h_v{'} + \\ r_{ac} * \left(\begin{array}{c} h\left(A_{k1}{'}\right) + h\left(A_{k2}{'}\right) + \cdots + h\left(A_{\mathrm{km}}{'}\right) - \\ h(A_{i1} - A_{i2} - \cdots A_{\mathrm{im}}) \end{array}\right) \end{array}\right),$$

$$C_S{''} = \mathrm{Encrypt}\left(\begin{array}{c} \mathrm{PK}_{S_2}, r_{\mathrm{pp}} + h_{\mathrm{up}}{''} + h_v{''} + \\ r_{\mathrm{ac}} * \left(\begin{array}{c} h\left(A_{q1}{'}\right) + h\left(A_{q2}{'}\right) + \cdots + h\left(A_{\mathrm{qm}}{'}\right) - \\ h(A_{w1} - A_{w2} - \cdots - A_{\mathrm{wm}}) \end{array}\right) \end{array}\right). \tag{9}$$

But since S_1 and S_2 do not know the values of r_{ac}, A_k, A_ω, so S_1 and S_2 cannot recover r_{pp} and the event e_{tp}.

5.1.3. Resistance to Spoofing Attacks. In our scheme, an edge server is placed in the same local network as the IoT devices, aiming to help the IoT devices perform certain kinds of computations. If the edge server is fake, it may fake the access credentials to recover event e, but it does not have any private keys of the subscribers to decrypt ciphertexts. At the same time, if an edge device tries to forge encrypted data while performing cryptographic computations, it will be detected and excluded by other nodes in the consortium blockchain. In addition, the consortium blockchain composed of edge devices has a certain fault-tolerant. Even if there are false malicious nodes in the network, as long as the number does not exceed 1/3 of the total number of nodes, it can guarantee the normal and stable operation of the system. So, even if the edge devices are fake, as long as there are enough honest nodes in the network, our scheme is also available.

5.2. Correctness Analysis

Theorem 2. *For the access control policy $\Gamma_{topic} = (A_{11} \Lambda A_{12} \cdots \Lambda A_{1m}) V \cdots V (A_{n1} \Lambda A_{n2} \Lambda \cdots A_{nm})$ of an event e with a topic tp, and an attribute conjunction $\gamma = (A_{11}{'} \Lambda A_{12}{'} \Lambda \cdots \Lambda A_{nm}{'})$ of a subscriber S, when $1 \le j \le m$ and $1 \le j \le k$, and $A_{i1} = A_{j1}{'}, \cdots, A_{im} = A_{jm}{'}$, then S can access all events of topic tp.*

Proof. In our scheme, the edge servers generate $C_p{'}, C_1$, C_s, I and C_{j3} for subscriber S, and S finally gets event e by decrypting it. When $e_{tp} + r_{pp} * r = e_{tp} (\mathrm{mod}\ r_{pp})$, if $r_{pp} > e_{tp}$, then Theorem 2 is satisfied; so, our scheme satisfies correctness. □

We also compare our scheme with other related work from the aspects of confidentiality, data privacy, decentralization, fine-grained access, collusion resistance, and ant-spoofing attack in Table 1, and the specific comparison results are described in Table 1.

As is shown in Table 1, all solutions are realized data event confidentiality; however, the proposed PS systems adopt centralized architecture in literature [22–27], in which all data are published to the subscriber by central broker, such a centralized architecture is vulnerable to the effects of a single point of failure, and the broker who is not fully trusted may leak or tamper with data, thus causing some insecure factors and posing a threat to the stable operation of the system. On the other hand, the data owner should have the right to determine who can use the data it provides, while in [24–30], there did not reflect the control of publishers over the authorization granularity for different information and subscribers. And subscribing services can be dishonest in practice, and the subscribers may attempt to access unauthorized events by colluding with each other, but most of the other work did not consider this problem. On the contrary, our scheme can better solve the above problems.

5.3. Performance Analysis. In order to verify the availability and performance of our proposed BPAC mechanism, we deployed our prototype system on two computers: the publisher/subscriber and blockchain broker both ran on the configured with 8.0G of RAM, AMD 2.3GHz CPUs, and Windows10_64 operating system, which the private blockchain is built on Ethereum. Furthermore, we use the Hyperledger Fabric deployed on the IBM Cloud platform for the consortium blockchain. Here, we use system throughput and two types of time delay as the main performance

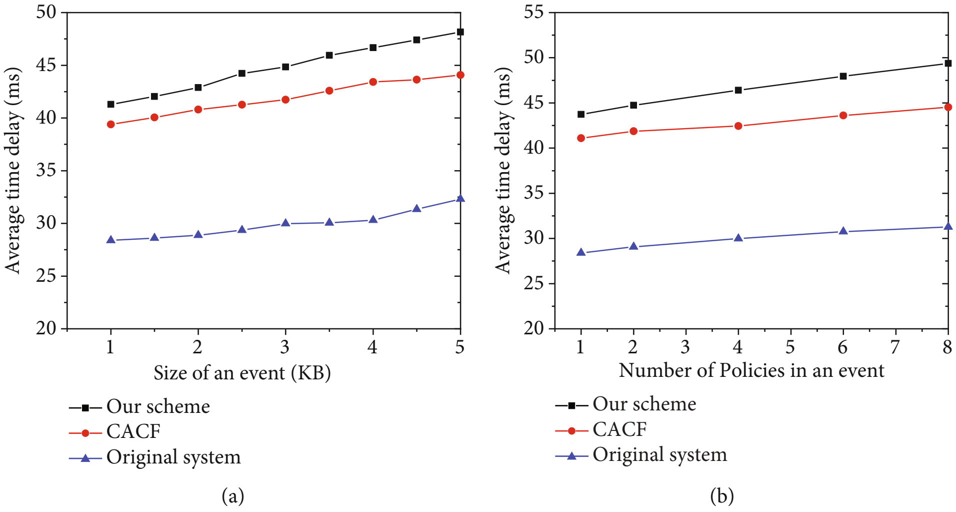


Figure 6: This is system delay and throughput with different event sizes: (a) latency with different sizes of one event (KB) and (b) throughput for different event sizes in KB.

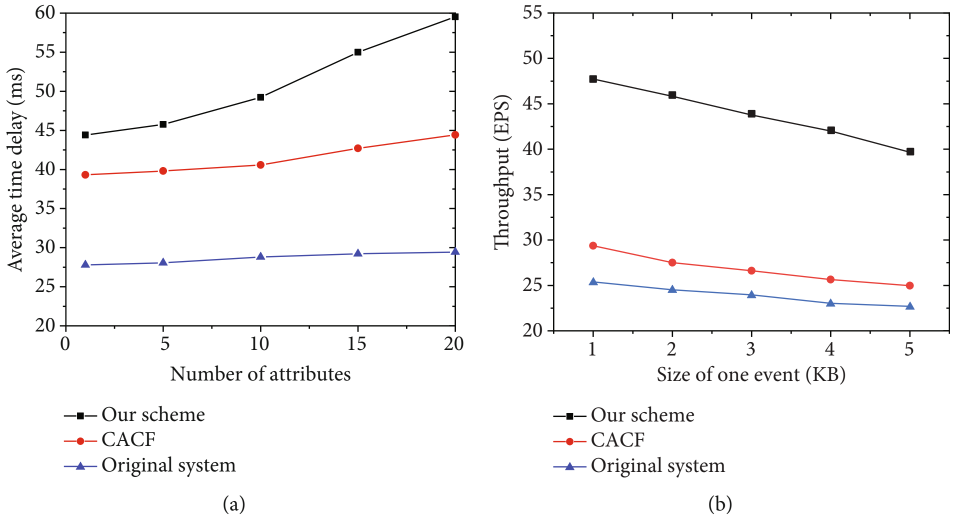


Figure 7: This is system delay with different numbers of attributes and policies: (a) latency with different numbers of attributes on one subscriber and (b) latency with different numbers of policies in one event.

evaluation criteria: (1) PS prototype system without using our proposed scheme and (2) using the proposed blockchain-based secure PS system. Among them, the time overhead of the prototype system is from the time the subscriber initiates the subscription request until the subscriber successfully obtains the publishing service or data. Our scheme would consist the additional time spent in running BPAC. This paper evaluates the proposed scheme in terms of the different event sizes of a publish event, the number of different policies, and the number of attributes of a subscriber, where the number of policies is 1, 2, 4, 6, and 8, and the number of attribute values is 1, 5, 10, 15, and 20. In addition, in order to better verify the efficiency of the proposed scheme, we compare our scheme with the CACF [22] scheme under the same test environment, which is a comprehensive access control framework using FHE scheme for publish/subscribe-based IoT services communication. The specific experimental results are shown as follows. It is worth noting that all data were obtained after running 100 times.

As is shown in Figure 6(a), with the publishing event sizes increases, the system delay gradually increases; that is, the size of the data event is one of the main factors that affect PS system latencies. Among them, the delay of the prototype system is significantly lower than our proposed scheme, and the CACF scheme is slightly higher than the prototype system but significantly lower than our scheme. This is due to the fact that the consensus validation process in our scenario consumes part of time and increase with the event complexity. Figure 6(b) shows the average sustainable throughput in processing the publishing events per second using different event sizes. Node that the throughput results are based on

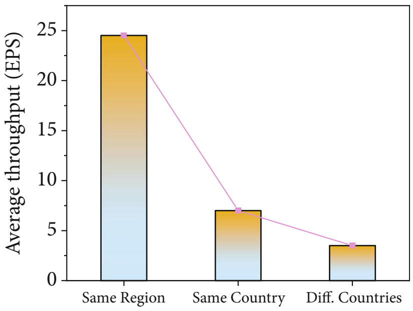


Figure 8: Throughput of query with nodes in different locations.

the average system latencies with or without our BPAC mechanism. As is shown in Figure 6(b), the system throughput decreases with the growth of data event sizes; that is to say, fewer the publishing events per second can be sent from the publisher to subscriber. In addition, we can know from the above two figures that the moderate amount of event data can complete PS service with low latency and acceptable throughput.

Figure 7 shows the impact on the system time overhead from both publisher and subscriber factors, where we mainly consider how the number of policies in one publishing event and attributes in one subscriber affect PS system latencies. In Figure 7(a), an increase in the number of subscriber attributes will result in an increase in the system time latency. This is because an increase in the number of attributes directly lead to more time in the attribute filtering and access control policy enforcement phases. Among them, the CACF scheme is still slightly lower than the scheme we proposed, which is because the FHE algorithm used in our scheme increases the time overhead. As shown in Figure 7(b), with the increase of access control policies, the time delay of the system gradually increases, and the delay of our scheme is about 43~50 ms. The time cost of the prototype system is significantly lower than ours, while CACF scheme is slightly higher than the prototype system but lower than our scheme. This is because our solution consumes part of the time and grows as the number of access control policies increases.

In Figure 8, in order to reflect the efficiency of crossdomain access operations, we test the throughput of our proposed PS system in different scenarios. All the experimental data is collected based on a minimum crossdomain access requirement that only involves one global ledger and two edge ledgers, and the average throughput in processing events per second is based on one KB event size. It is clear from Figure 8 that the physical location of the nodes also affects the performance of the PS system.

As can be seen from the results discussed above, although our proposed BPAC mechanism increases the system time delay compared with the CACF scheme, the absolute value of the delay increment is not large, and the application of blockchain in the PS system makes up for the lack of security and trust in the traditional scheme. We compromised the acceptable response time in exchange for higher reliability and solved the security problem in the PS system.

6. Conclusion and Future Work

In this paper, we propose an access control mechanism based on blockchain and FHE algorithm, which solves the security and privacy problems in the traditional centralized PS system. Our scheme protects the confidentiality of event data by encrypting the publishing data with the FHE algorithm. Meanwhile, it replaces the traditional central broker with the blockchain technology to realize decentralized distributed access control and realizes crossdomain information interaction by storing data in the global ledger. According to the theoretical analysis, it can guarantee the security and correctness of the system, and the experimental results show that our scheme is feasible and efficient to some extent.

However, our scheme also has certain deficiencies, such as our solution did not completely realize attribute revocation and update of access policies, and with the rapid growth of the IoT network scale, the attributes of one subscriber and access control policies for publishing events also become increasingly complex, as it may take more time in the matching stage, so as to further prolong system response time. In future research work, we will further solve the above problems. We plan to combine the two-strategy attribute-based authorization [33] and time-limited key management to realize more fine-grained access control and efficient key revocation and further adopt the Bloomer Filter [34] to optimize the matching process to achieve fast authentication.

Notations

λ:	Security parameter
L:	A number of levels
b:	Bit
q:	Prime
Params :	System parameter
$(\mathrm{PK}_p, \mathrm{SK}_p)$:	The key pair of publisher
$(\mathrm{PK}_S, \mathrm{SK}_S)$:	The key pair of subscriber
$r_{\mathrm{pp}}, r_{\mathrm{up}}, r_{\mathrm{ac}}$:	Random number
h:	Hash function
e_{tp}:	Publishing event
tp:	Topic
Λ_{tp}:	Access policy
C_{tp}:	The ciphertext of the publishing event
A_{im}:	Attribute collection
F:	Attribute filter function
ω_{topic}:	Access credentials
j:	Property index value
I:	The ciphertext of property index value
ω_s:	Attribute conjunction
$\mathrm{KS}_{P\longrightarrow S}$:	The exchanged key
$\mathrm{EL}_i/\mathrm{GL}$:	Edge ledger/global ledger.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Disclosure

A preprint has previously been published [35].

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the State Grid Liaoning Electric Power Supply Co., Ltd (No. 2022YF-22).

References

[1] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, "A survey on Internet of Things: architecture, enabling technologies, security and privacy, and applications," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1125–1142, 2017.

[2] F. Javed, M. K. Afzal, M. Sharif, and B. S. Kim, "Internet of Things (IoT) operating systems support, networking technologies, applications, and challenges: a comparative review," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 2062–2100, 2018.

[3] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: a survey on enabling technologies, protocols, and applications," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015.

[4] Y. A. Younis, K. Kifayat, and M. Merabti, "An access control model for cloud computing," *Journal of Information Security and Applications*, vol. 19, no. 1, pp. 45–60, 2014.

[5] "Ericsson mobility report," 2020, https://www.ericsson.com/en/internet-of-things.

[6] L. Malina, G. Srivastava, P. Dzurenda, J. Hajny, and R. Fujdiak, "A secure publish/subscribe protocol for internet of things," in *Proceedings of the 14th international conference on availability, reliability and security*, pp. 1–10, Canterbury CA UK, 2019.

[7] C. Esposito and M. Ciampi, "On security in publish/subscribe services: a survey," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 966–997, 2014.

[8] A. V. Uzunov, "A survey of security solutions for distributed publish/subscribe systems," *Computers & Security*, vol. 61, pp. 94–129, 2016.

[9] A. Banks and R. Gupta, "MQTT version 3.1.1," 2014, OASIS Standard, http://docs.oasis-open.org/mqtt/mqtt/v3.1.1/os/mqtt-v3.1.1-os.html.

[10] D. Hughes, K. Thoelen, W. Horré et al., "LooCI: a loosely-coupled component infrastructure for networked embedded systems," in *Proceedings of the 7th International Conference on Advances in Mobile Computing and Multimedia*, pp. 195–203, Kuala Lumpur Malaysia, 2009.

[11] P. A. Levis, S. Madden, D. Gay et al., "The emergence of networking abstractions and techniques in TinyOS," *NSDI*, vol. 4, pp. 1–1, 2004.

[12] S. Nakamoto, "Bitcoin: a peer-to-peer electronic cash system," 2009, https://metzdowd.com.

[13] J. Golosova and A. Romanovs, "The advantages and disadvantages of the blockchain technology," in *IEEE 6th workshop on advances in information, electronic and electrical engineering (AIEEE)*, pp. 1–6, Vilnius, Lithuania, 2018.

[14] Z. Brakerski, C. Gentry, and V. Vaikuntanathan, "(leveled) fully homomorphic encryption without bootstrapping," *ACM Transactions on Computation Theory*, vol. 6, no. 3, pp. 1–36, 2014.

[15] J. Ren, Y. Pan, A. Goscinski, and R. A. Beyah, "Edge computing for the Internet of Things," *IEEE Network*, vol. 32, no. 1, pp. 6-7, 2018.

[16] A. V. Dastjerdi and R. Buyya, "Fog computing: helping the Internet of Things realize its potential," *Computer*, vol. 49, no. 8, pp. 112–116, 2016.

[17] H. Tian, X. Ge, J. Wang, C. Li, and H. Pan, "Research on distributed blockchain-based privacy-preserving and data security framework in IoT," *IET Communications*, vol. 14, no. 13, pp. 2038–2047, 2020.

[18] P. T. Eugster, P. A. Felber, R. Guerraoui, and A. M. Kermarrec, "The many faces of publish/subscribe," *ACM computing surveys (CSUR)*, vol. 35, no. 2, pp. 114–131, 2003.

[19] J. Li, W. Yao, J. Han, Y. Zhang, and J. Shen, "User collusion avoidance CP-ABE with efficient attribute revocation for cloud storage," *IEEE Systems Journal*, vol. 12, no. 2, pp. 1767–1777, 2018.

[20] "Hyperledger Fabric," 2020, https://www.hyperledger.org/projects/fabric.

[21] E. Androulaki, A. Barger, V. Bortnikov et al., "Hyperledger fabric: a distributed operating system for permissioned blockchains," in *Proceedings of the thirteenth EuroSys conference*, pp. 1–15, Porto Portugal, 2018.

[22] L. Duan, C. A. Sun, Y. Zhang, W. Ni, and J. Chen, "A comprehensive security framework for publish/subscribe-based IoT services communication," *IEEE Access*, vol. 7, pp. 25989–26001, 2019.

[23] K. Yang, K. Zhang, X. Jia, M. A. Hasan, and X. S. Shen, "Privacy-preserving attribute-keyword based data publish-subscribe service on cloud platforms," *Information Sciences*, vol. 387, pp. 116–131, 2017.

[24] Q. Wang, D. Chen, N. Zhang, Z. Ding, and Z. Qin, "PCP: a privacy-preserving content-based publish–subscribe scheme with differential privacy in fog computing," *IEEE Access*, vol. 5, pp. 17962–17974, 2017.

[25] A. A. Diro, N. Chilamkurti, and N. Kumar, "Lightweight cybersecurity schemes using elliptic curve cryptography in publish-subscribe fog computing," *Mobile Networks and Applications*, vol. 22, no. 5, pp. 848–858, 2017.

[26] A. Diro, H. Reda, N. Chilamkurti, A. Mahmood, N. Zaman, and Y. Nam, "Lightweight authenticated-encryption scheme for internet of things based on publish-subscribe communication," *IEEE Access*, vol. 8, pp. 60539–60551, 2020.

[27] C. Borcea, Y. Polyakov, K. Rohloff, and G. Ryan, "PICADOR: end-to-end encrypted publish-subscribe information distribution with proxy re-encryption," *Future Generation Computer Systems*, vol. 71, pp. 177–191, 2017.

[28] Y. Zhao, Y. Li, Q. Mu, B. Yang, and Y. Yu, "Secure pub-sub: blockchain-based fair payment with reputation for reliable cyber physical systems," *IEEE Access*, vol. 6, pp. 12295–12303, 2018.

[29] P. Lv, L. Wang, H. Zhu, W. Deng, and L. Gu, "An IoT-oriented privacy-preserving publish/subscribe model over blockchains," *IEEE Access*, vol. 7, pp. 41309–41314, 2019.

[30] M. A. Tariq, B. Koldehofe, and K. Rothermel, "Securing broker-less publish/subscribe systems using identity-based encryption," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 2, pp. 518–528, 2014.

[31] G. S. Ramachandran, K. L. Wright, L. Zheng et al., “Trinity: a byzantine fault-tolerant distributed publish-subscribe system with immutable blockchain-based persistence,” in *IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, pp. 227–235, Seoul, Korea (South), 2019.

[32] S. Gao, G. Piao, J. Zhu, X. Ma, and J. Ma, “TrustAccess: a trustworthy secure Ciphertext-policy and attribute hiding access control scheme based on Blockchain,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 6, pp. 5784–5798, 2020.

[33] L. Duan, Y. Zhang, S. Chen, S. Wang, B. Cheng, and J. Chen, “Realizing IoT service’s policy privacy over publish/subscribe-based middleware,” *Springerplus*, vol. 5, no. 1, 2016.

[34] R. Barazzutti, P. Felber, H. Mercier, E. Onica, and E. Riviere, “Efficient and confidentiality-preserving content-based publish/subscribe with prefiltering,” *IEEE Transactions on Dependable and Secure Computing*, vol. 14, no. 3, pp. 308–325, 2017.

[35] H. Tian, X. Ge, J. Wang, and C. Li, “Exploiting blockchain and secure access control scheme to enhance privacy-preserving of IoT publish-subscribe system,” *Research Square*, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 3813841, 14 pages
https://doi.org/10.1155/2022/3813841

Research Article

BIoMT Modular Infrastructure: The Recent Challenges, Issues, and Limitations in Blockchain Hyperledger-Enabled E-Healthcare Application

Zaffar Ahmed Shaikh,[1] **Abdullah Ayub Khan**,[1,2] **Lin Teng**,[3] **Asif Ali Wagan**,[2] **and Asif Ali Laghari**[2]

[1]*Department of Computer Science and Information Technology, Benazir Bhutto Shaheed University Lyari, Karachi, 75660 Sindh, Pakistan*
[2]*Department of Computer Science, Sindh Madressatul Islam University, Karachi, 74000 Sindh, Pakistan*
[3]*Department of Software Engineering, Software College, Shenyang Normal University, Shenyang, China*

Correspondence should be addressed to Abdullah Ayub Khan; abdullah.ayub@bbsul.edu.pk and Lin Teng; 1532554069@qq.com

Received 9 June 2022; Revised 18 August 2022; Accepted 30 August 2022; Published 21 September 2022

Academic Editor: Cong Pu

This paper presents a layered hierarchy that depicts the progressive relationship between data, information, knowledge, and wisdom. To begin with, data is gathered and organized into information. Information is gathered, filtered, refined, and put through an investigation process to create knowledge. Wisdom is attained after knowledge discovery through the process of filtration and aggregation through experience. The layered hierarchy in the domain of e-healthcare necessitates higher scheduling costs for data collection, processing wisdom, and management, which is also an insecure and untrustworthy process for progressive medical service. The medical industry faces a difficult problem in providing collected data integrity, information reliability, and knowledge trustworthiness for the service of progressive medical relationships in the face of an increasing number of day-to-day records. The blockchain consortium hyperledger (fabric) has been used in this paper to act as a bridge that bridges the gap between electronic data, information, knowledge, and wisdom (DIKW) movement and processes by enabling the process of the layered hierarchy of schedule information and management and providing security and transparency. For e-healthcare information management and privacy, the DIKW-ledger, such as patients' consultancy information, availing medical services, personal records, appointments, treatment details, and other health-related transactions, a consortium hyperledger fabric-enabled efficient architecture is proposed. This proposed architecture creates two networks: a public network for medical stakeholders to exchange and agree on specific medical activities before being preserved on distributed storage (read-only after record registration) and a private network for complete DIKW process scheduling and management. We designed and created smart contracts for this purpose, as well as use-case diagrams to describe the overall execution process. The proposed architectural solution provides more efficient information integrity, provenance, and storage procedures to immutably preserve the medical ledger in a permissioned hash-encrypted structure.

1. Introduction

Data is one of the fundamental elements. It could be a common denominator in which all the constraints are connected and are preserved in the centralized storage. This data is derived from information and positioned through a continuum that exactly moves towards knowledge and then wisdom [1]. Representation of data is knowledge-driven and provides distinct attributes, for example, patient name, tracking id, gender, age, cause, symptoms, and prescription. The meaning of data attributes is semantic and drives proper information and knowledge. In the next step, this contextualized information is implied. Then, it is interpreted by the interpreter from the perspective of the information receivers.

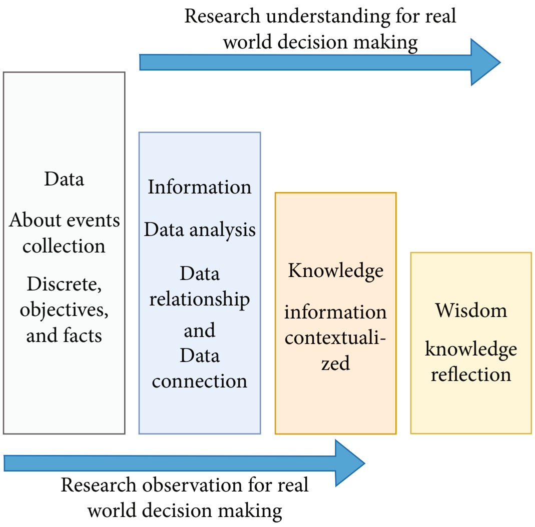

Research understanding for real world decision making
Data
About events collection
Discrete, objectives, and facts
Information
Data analysis
Data relationship and Data connection
Knowledge
information contextuali-zed
Wisdom
knowledge reflection
Research observation for real world decision making

Figure 1: Current process of DIKW analysis.

After this process, the high-level knowledge states that are pertinent to wisdom are shown in Figure 1. When knowledge is truly refined and sublimated, the receiver has the potential to minimize and maximize interaction with the medical environment [2, 3].

In many cases, there is an emerging ambiguity between the definitions of information and knowledge, especially in the medical DIKW domain. The distinctness between information and knowledge may be the interpretation of users; few of them call data "information," while others call it "knowledge" [4]. To reduce equivocation, several information systems use centralized storage to preserve metadata. This recorded metadata, on the other hand, aids in the process of interpreting and transforming data into well-formed information [5]. Giant enterprises often preserve the same data in different storage structures, which creates redundancy. For instance, the process of storing information requires more time and computational power to schedule, process, organize, and store the record in the file system (storage). Moreover, this complete scenario is insecure because of the procedure of records preserved in the centralized server-based storage structure.

For every information system, the individual entity must take metadata into account when attempting to interpret data [6]. Most of the time, these additional data entities must be considered together, such as records of patients by name recorded in three different data fields in terms of first name, last name, and middle name, before the information is driven from the data. However, from the perspective of healthcare organizations, the complexity involved in accessing sensitive patient information across distinct central server-based applications and organizational boundaries is most important, yet does affect the potential cost and generate errors [7, 8].

Recently, healthcare industries have utilized different mechanisms, procedural domains, criteria of facilities, and systems that are used for processing patient data [9]. Some patients may receive healthcare services at more than one physician's consultancy affiliated with the same hospital. For instance, sometimes, patients receive medical treatment by random occurrence and sometimes as part of the medical management process [10]. Further, this happens when there is no data exchange between the connected systems (nodes), and incomplete data occurs during data exchange from different disjoint nodes because the same patients end up with two different medical records within the same e-healthcare applications [11, 12]. However, these patients' identity redundancy creates another challenging aspect in terms of data management. Moreover, data management also requires record cleanup before organizing medical processed ledgers. Additionally, record cleanup involves medical duplication data detection, examination, removal, analysis of incomplete data, correction, tuning the format of records, and preservation.

There are numerous methods and techniques proposed for e-healthcare application integration, organization, and security practitioners. To ensure the validity, authentication, and reliability of e-healthcare systems and information processing and the overall management of records [9], it is imperatively significant to maintain the transparency and privacy of the complete process of medical data, information, knowledge, and wisdom over the network. To ascertain this, the identification of shreds of sensitive medical records is critical for the record-keeping purpose of an individual transaction that occurred while analyzing the medical data [10, 11]. The complexity of hiding innumerable kinds of medical information in a carrier channel (through signals) over wireless network-based connected edge devices to exchange information. The remote medical data acquisition mechanism itself is highly vulnerable while capturing digital medical data from distinct domains of the healthcare network [12]. These medical data node transactions and information exchange layered mechanisms are deemed insecure.

For further investigation, blockchain technology has been envisioned and utilized by several industrial production systems and supply chain management to achieve provenance, integrity, and tracking to enable record storage for further investigation [13]. Most medical analysts are planning to shift from centralized to decentralized care; for this purpose, blockchain distributed technology used as a decentralized secure infrastructure protects against network attacks [14]. Usually, it is intended for server-based central systems and structures. Blockchain also enables the robust performance of distributed nodes' defense ability during the process of sharing information from one to another [15]. Security is possible because of cryptographic hash-encryption (SHA-256) functions with the installation of intrusion detection and restricted solutions. Moreover, the technology can deploy a firewall with antidisclosure techniques to guarantee the medical ledger integrity, transparency, provenance, immutability, and trustworthiness of the stored records under examination.

However, this paper discusses a secure blockchain hyperledger fabric that enables a novel medical architecture for e-healthcare information management and privacy. For this purpose, we have designed a private network infrastructure for exchanging sensitive medical information among different nodes in a protected manner (encrypted ledger) over P2P network connectivity. This proposed medical-ledger provides overall data, information, knowledge, and wisdom of medical record provenance, track and trace, two-way protected communication, and assurance for performing all the medical-related operations. These working operations are (i) patients' registration, (ii) online medical services availing, (iii) medical alerts, (iv) physician consultancy and registration, (v) cost scheduling, (vi) payment criteria, (vii) wallet, etc. This scenario creates trust between the events while receiving the patient's medical data and preserving, examining, and interpreting the medical information. The main contributions of this paper are as follows:

(i) A blockchain consortium hyperledger network-enabled structure for medical information management and privacy is proposed

(ii) In this paper, we propose a secure process hierarchy of medical data, information, knowledge, and wisdom using blockchain-enabled serverless peer-to-peer (P2P) consortium (hybrid) network infrastructure

(iii) To automate transactions of e-healthcare, the pseudosmart contracts are designed and simulated to manage e-healthcare-related events and medical node transactions in a protected manner using the cryptographic hash-encryption (SHA-256) method. For this purpose, we create three distinct chaincodes, such as patients' device registration, new transactions and adding medical nodes, and updating the ledger

(iv) The proposed BIoMT modular medical-ledger architectural operation is simulated using an activity diagram in a permissioned private and permissionless public blockchain network

(v) Finally, we evaluate and examine the current e-healthcare applications and discuss the challenges and limitations of the proposed distributed applicational (DApp) architecture. The blockchain hyperledger fabric-enabled implementation's open issues and future directions are discussed

The remaining sections of this paper are organized as follows. In Section 2, we studied medical-related information management and privacy protection-based literature review and examine the radical impact on the previously proposed e-healthcare applications. The existing procedure of medical DIKW analysis and the communication between layered hierarchy is discussed in Section 3. In Section 4, we have presented a blockchain hyperledger fabric-enabled proposed architecture for medical DIKW management and privacy. The working operations of the proposed architecture are discussed in Section 5. Moreover, we have evaluated and analyzed the current implementation and involving challenges and limitations are addressed in the subsection (Section 5). Finally, we conclude this paper in Section 6.

2. Related Work

Blockchain technology is a decentralized database associated with the distribution chain of chronological order, the underlying system of cryptographic security. Essentially, a sequence of aligned data nodes is associated with the hash-encrypted (SHA-256) method. The individual node contains a chain of information about a specific blockchain batch of transactions over the peer-to-peer network [16]. The batch transaction is used to verify and validate the information and move to the next node (or generate a node). In the domain of medical information, each data unit can be encapsulated into a shell called a node, which creates the information-data chain-of-unit according to the defined consensus policy. However, various problems remain in the content of blockchain e-healthcare at present due to the different emphasis and development of DApp, a new way of blockchain distributed engineering in the medical environment [17].

Due to the popularity and wide use of cloud computing, medical information can be regarded as a collection. The application of distinct medical resources at different time frames in a decentralized domain reduces the cost of load [18]. The whole scenario ensures that each individual resource is independent and unique. Therefore, the distributed technology team medical resources and provides data, information, knowledge, and wisdom management with privacy. Blockchain-medical information is a new paradigm that ensures record preservation, transmission, and distributed connectivity, blockchain consensus policies, hash protection, and other information systems in the field of medical sciences [19].

Smart medical data examination and information recommendation and management based on blockchain, when combined with diverse medical data resources in different storage domains, provides organizations and data analysts with efficient identification, analysis, detection, and classification of a large number of medical records [16, 20]. The learning preferences and other activities related to medical data and wisdom are designed and recent e-healthcare assumptions are discussed in Table 1.

3. Current Process of DIKW and Clinical-Distributed Aware Technology for Usability

In this context, the concept DIKW is presented in the form of a scale that elaborates the process of moving towards increased understanding [27], as shown in Figure 2. For this reason, the process of DIKW usability in the medical domain is expressed through the applications of e-healthcare, where clinical associates with patients who are dissatisfied with the system utilization and evaluations of electronic health records. International medical standards categorize clinical trials into three main portions for the sake of usability: efficiency, effectiveness, and increased satisfaction level in terms of medical data processing, scheduling, and management [28, 29]. Include the patients' experience, historical ledger, and incorporated principles for designing clinical applications [29]. To protect patients' records' safety

Table 1: Management of medical data, information, knowledge, and wisdom, as well as blockchain security related literature reviews.

Research method	Research description	Research gaps	Similarity/difference with the proposed BIoMT
From trustworthy data to trustworthy IoT: a data collection methodology based on blockchain [21]	A data collection method based on blockchain-IoT is proposed for creating a trustworthy environment. A hyperledger fabric-enabled smart contract is designed and implemented to balance trust and privacy during the process of collection.	(i) Preprocessing issue (ii) Hybrid and complex systems' computational limitations (iii) Micro provider (iv) Data similarity-related challenges	(i) Hyperledger fabric (ii) Hash encryption (SHA-256) (iii) Permissioned, private network
MedSBA: a novel and secure scheme to share medical data based on blockchain technology and attribute-based encryption [22]	The authors of MedSBA presented a secure and efficient blockchain-enabled share medical records and attribute-based encryption system proposed to record and store medical data. This proposed system protects patients' privacy and allows fine-grain access control of medical services in the e-healthcare environment.	(i) PBFT consensus method (ii) Hybrid blockchain (iii) Communication and security related protocols issue (IoT/edge-enabled protocol) (iv) Patient's device registration limitation	(i) General data protection regulation (ii) Private blockchain (iii) OPNET software tool used (iv) BAN logic
A blockchain-based scheme for privacy-preserving and secure sharing of medical data [23]	The authors of this paper presented a blockchain-enabled privacy preservation scheme that enabled the secure exchange of sensitive medical information between participating stakeholders in a semitrusted cloud server. In addition, this proposed system achieves data availability between stakeholders where zero-authentication proof is employed.	(i) Hybrid communication channel (ii) Two-way authentication (iii) Cloud storage used	(i) Proxy-reencryption (ii) PBFT algorithm for transactions delivery and acknowledgement (iii) Zero-knowledge proof mechanism
Medical data sharing scheme based on attribute cryptosystem and blockchain technology [24]	A medical data sharing model based on attribute-hash-encrypted and blockchain is proposed. In this paper, the data is validated first and then preserved on an efficient storage medium (distributed). And so, reduce the possibility of irreversible modification. For this reason, the authors designed a many-to-many communication mechanism for sharing sensitive medical data between stakeholders.	(i) Data duplication issue (ii) ABS protocol (iii) Identity privacy for preservation (iv) Chosen cypher-text attack	(i) Attribute-based signature (ii) Attribute-based encryption (iii) Many-to-many communication channel
Design of a Secure Medical Data Sharing Scheme Based on Blockchain [25]	The authors of this paper proposed a blockchain-enabled authentication process for a network model of a medical cyber-physical system. This model is designed to ensure the data cannot be forged, tampered with, or untrackable.	(i) Sharing of medical big data (ii) Credibility problem (iii) Intractable challenges (iv) Bilinear mapping	(i) BAN logic (ii) Two-way authentication (iii) Permissionless network (iv) Consortium chain (v) Formal verification authentication
MEdge-chain: leveraging edge computing and blockchain for efficient medical data exchange [26]	The authors of this paper proposed a medical-ledger blockchain- (MEdge-chain-) enabled holistic framework for exploiting the integration to aggregate diverse healthcare entities. Such medical entities' scheduled processes of storage are, for example, swift first, then secure sharing, and lastly, preservation.	(i) Optimal blockchain configuration (ii) Data priority assignment challenge (iii) Limited resources (iv) Connectivity issue (v) Local healthcare service provider management (vi) Monitoring a large number of patients	(i) Delegated proof-of-stake consensus (ii) Edge-based remote monitoring (iii) Efficient data discovery (iv) Permissionless blockchain public network

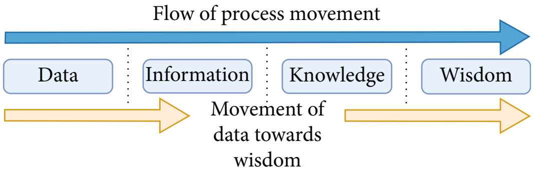


Figure 2: DIKW process movement.

and quality, the distributed system-based solution is proposed to enhance system usability, as dissatisfaction reduces and increases availability.

Still, most of the e-healthcare system relies on a centralized mechanism with a lack of security and availability. For this purpose, healthcare departments need to be concerned about the adoption of secure data, information, knowledge, and wisdom architecture as an integral component of medical information [28]. In this act, the data obtained from the system usability is evaluated by end-users (patients) and transformed into information. Whereas knowledge is the result of understanding the implications of information analysis. And so, the medical ledger of extracting wisdom is the distributed system of knowledge to enhance e-healthcare processes for clinical trials.

3.1. A Layered Hierarchy of Medical DIKW and Blockchain E-Healthcare Distributed Applications. The layered hierarchy has rigidly set building nodes in this pyramid-like structure, where the data comes first. It collects medical/clinical facts in an unorganized form, such as a number or character [30]. However, without patient context, medical data can mean little, such as a patient name, which cannot provide a complete, detailed understanding of the specific record. On the other hand, data is provided in the form of a tracking number, for example, through the patient ID "12011," which gets all the descriptions regarding the utilized medical services and their personal information as well. But the view in the context of data needs to transform the raw sequence of numbers into meaningful [31].

Information is the next node of the layer hierarchy. This is data that has been cleaned of errors and further processed in a way that makes it easier to evaluate, present, and analytically explain. To process information, the data processing mechanism involves distinct operations [32], for example, aggregation/accumulator, validation, and organizing in a way that explores the relationship between several disconnected points of data. If medical processed data/information is viewed as a description of collected objectives, facts, and discrete points, but also understood to apply it so to achieve the healthcare meta-information that helps in future investigations, it is turned into knowledge [33]. If medical processed data/information is viewed as a description of collected objectives, facts, and discrete points, but also understood to apply it so to achieve the healthcare meta-information that helps in future investigations, it is turned into knowledge [33]. This medical-related knowledge is often the edge that the healthcare sectors have over their research investigations. We uncover relationships that are not explicitly stated as information, as shown in Figure 3. However, in the domain of healthcare, wisdom is knowledge applied in action. Knowledge and wisdom are associated with what was achieved in the clinical investigation, whereas data and information are a look back in time.

To protect the layered hierarchy of healthcare DIKW, the proposed blockchain distributed ledger architecture facilitates the secure transfer of patient medical records, manages the healthcare-related supply chains, and helps e-healthcare systems/applications for privacy management. Giant organizations are adopting blockchain-healthcare technology, such as Akiri, Factom, MedicalChain, RoboMed, and Chronicled, for medical ledger integrity, transparency, provenance, immutability, and availability in a distributed nature [34, 35]. This technology keeps all the important medical data safe and secure at the moment (processing schedule dynamically). The blockchain decentralized mechanism manages all the patients' logs transparently and makes patient data available on a technology rife for security distributed applications. Substantially, blockchain does not only provide transparency; it manages medical ledgers privately, concealing the patient's identity with complex hash-based encryption and secure codes that tackle protection-related challenges and make sensitive medical records safe in the immutable storage [36, 37]. Moreover, the distributed nature of blockchain healthcare allows stakeholders to exchange the same information more quickly and efficiently.

4. Proposed Architecture for Information Management and Privacy

In this context, we proposed a distributed architecture of medical ledger management and privacy protection using a hyperledger fabric and blockchain-enabled smart contracts for secure information preservation in the immutable storage. This whole process is initiated after collection of medical data through the edge devices. One type of input that is allowed for handling real-time medical ledger management, and privacy from data to wisdom is as follows.

4.1. Fabric Endorsement and Service Orderer. The healthcare distributed ordering service (OS) starts with the application request of the medical node transactions, where the overall order in this ledger is endorsed by the connected nodes on the blockchain hyperledger fabric P2P network. The healthcare-related medical node transactions contain a unique signature (signed by each stakeholder before publishing transactions) and SHA-256 cryptographic encryption by the individual connection for committer/endorsement. This is all then submitted to the fabric orderer, where it is transmitted to the fabric commit for digital medical ledger security, shown in Figure 4. After that, the medical services are broadcast among participating stakeholders, from the orderer to the fabric committer on the blockchain network. For validation, the KAFKA verification predefined fabric mechanism for security purposed used; the defined consensus and hash reencryption are shown in Contract 1 and Figure 5.

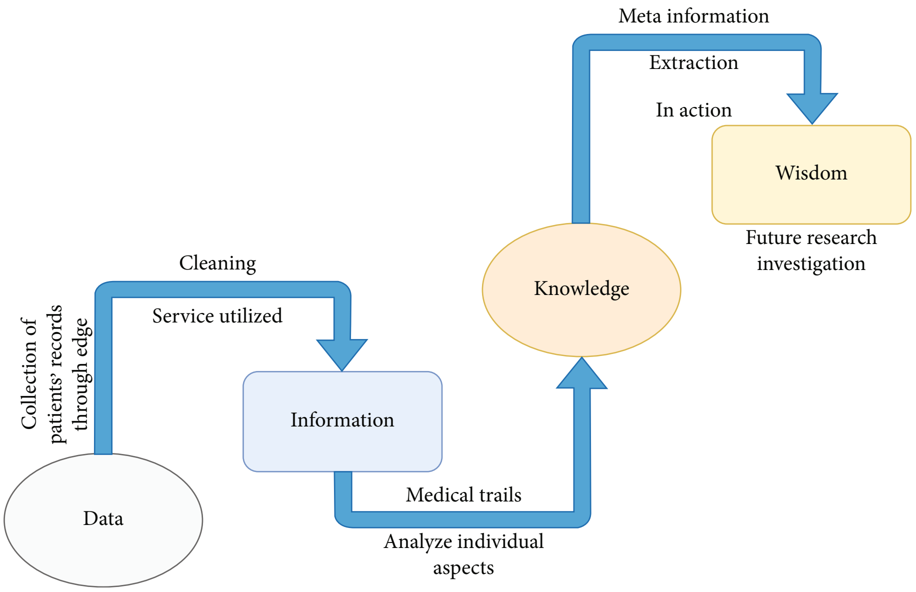


Figure 3: The layered hierarchy of DIKW in healthcare environment.

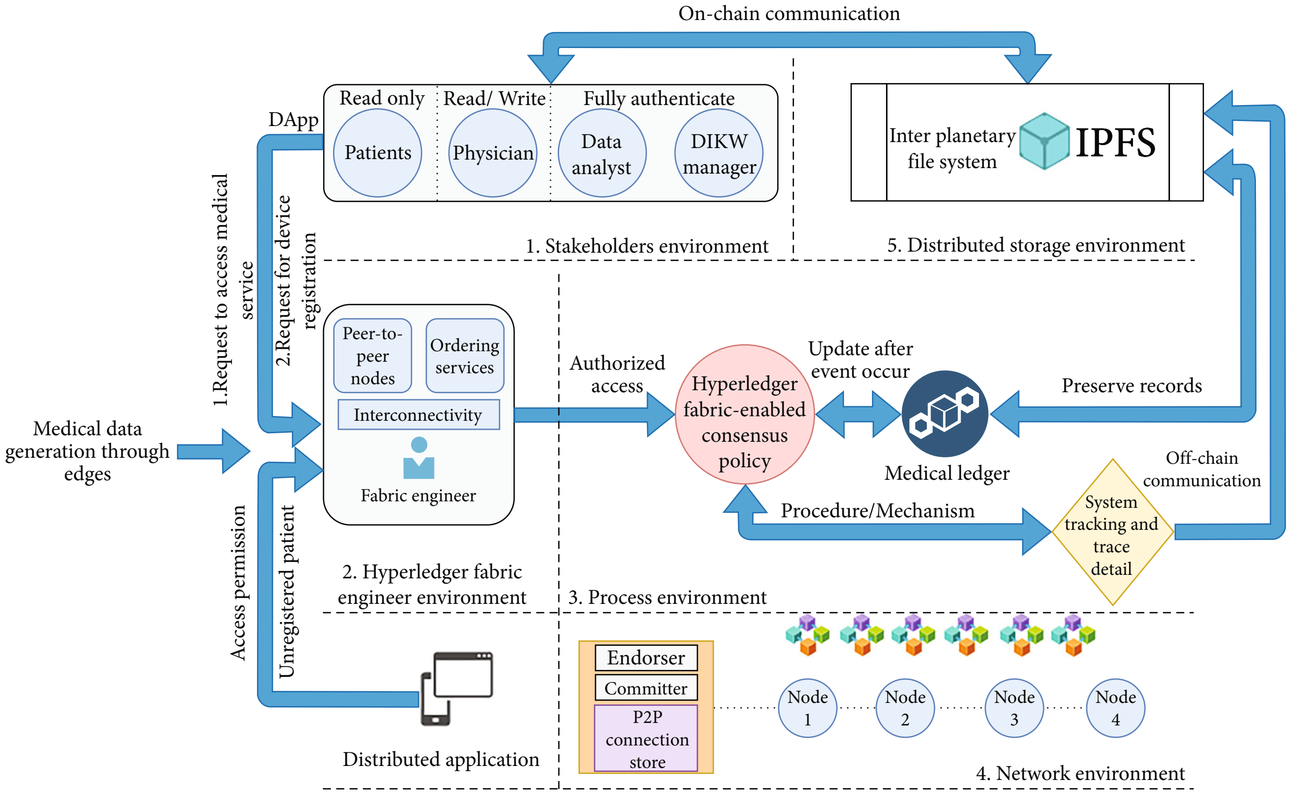


Figure 4: Proposed medical-ledger architecture for information management and privacy.

4.2. Certificate Authority of the Proposed Architecture. In the blockchain hyperledger fabric network, a certificate authority network is designed to analyze different untrusted connected stakeholders in the medical-ledger architecture, shown in Figure 4. If it has device registration and a root authenticate (participating identity), identify the stakeholders who are participating. The engineer provides certificate identity only to the participating stakeholders or requests participation after verification and validation. The healthcare system binds specific connected nodes and orders/requests. For certificate allocation, the blockchain fabric network engineer mimics all the transactions among stakeholders and is also responsible for managing overall renewal transactions, updates, and addresses. These private

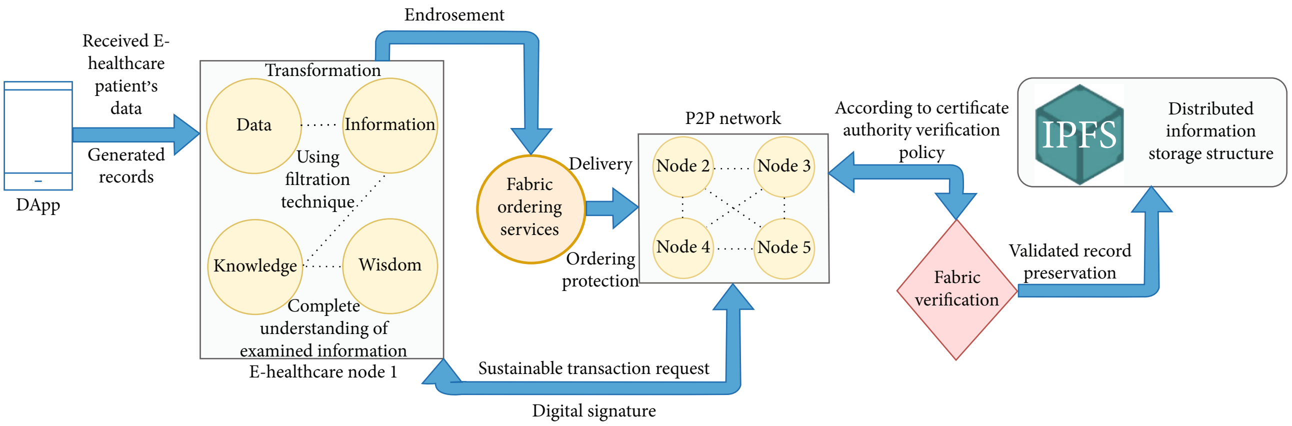

Endrosement
Received E-healthcare patient's data
Generated records
DApp
Transformation
Data
Information
Using filtration technique
Knowledge
Wisdom
Complete understanding of examined information
E-healthcare node 1
Fabric ordering services
Delivery
Ordering protection
P2P network
Node 2
Node 3
Node 4
Node 5
According to certificate authority verification policy
IPFS
Distributed information storage structure
Fabric verification
Validated record preservation
Sustainable transaction request
Digital signature

Figure 5: DIKW of e-healthcare node transaction and verification process.

on-chain and off-chain communications are singed digitally by the participating stakeholders and share private keys (for off-chain communication) to protect medical ledger management and preservation. These whole scenarios occur after the verification. It only uses the public key (on-chain communication) within the system.

4.3. Peer-to-Peer Permissioned Network (P2P). In the proposed architecture, the private network is designed to resist direct message delivery and pathway reception because of medical-ledger node transactions related to private security and integrity. The proposed medical ledger provides a strong communication channel, including transaction protection while exchange, secure stakeholders connection, and consensus workload, which is not directly accessible, so the channel can only be operated by the engineer and participating stakeholders with their registered devices. However, the execution of medical transactions is fully private and separate. This fabric network enables efficient medical transaction delivery and manages ledger maintenance as compared to other centralized systems. In this proposed network, the smart contract design protects individual transactions using a hash-based encryption mechanism and invokes specific types of node transactions execution on the secure private defined channel using blockchain protocols, as shown in Figure 5.

4.4. Distributed Nodes of Medical Transactions and Storage. The log and state execution are designed between multiple connected nodes in a private network channel. The system synchronizes automatically and runs two main objectives for managing information and security, such as committer to endorsement and endorsement to the committer. The medical node transaction request is submitted to the engineer according to the procedure of the predefined fabric endorser. As shown in Figure 4, this process is scheduled after the completion of node peering (connectivity). In this private network, blockchain distributed ledger management and storage are defined, where InterPlanetary File Storage (IPFS) is used for secure medical record preservation. IPFS is a third-party storage system that is utilized by just paying a small fee.

4.5. Smart Contracts. As shown in Contract 1 (and Figure 6), first patient device registration is required to turn on the blockchain ledger environment for the novel and secure smart contract-aware medical information management and privacy architecture. The blockchain hyperledger fabric-enabled engineer starts the system and implements the patient device registration contract (pDeviceReg()) and customized fabric stakeholders' privacy and exchange-related consensus policies designed for private device registration of individually connected stakeholders. It also records collected medical-related data such as service utilization and delivery of the healthcare application in accordance with the defined consensus, shown in Contract ((newNodeTransaction()) and (updateTransPreserve())). Furthermore, the pDeviceReg() function also stores additional records related to the medical-ledger, including device ID (dID()), patient device registration (pdReg()), patient ID (pID()), patient name (pName()), blockchain timestamp [execute], and all the activities are performed, as shown in Figure 5.

As a result, the blockchain hyperledger fabric-enabled engineer implemented and managed an automatically updated ledger for the (newNodeTransaction()) with hash-based events of node transaction protection (reEncryption()) for medical ledger privacy security, as shown in Figure 6. As well as storing medical/clinical services related to collected healthcare data, this adds new transactional details (after analysis of knowledge/wisdom) to the healthcare immutable storage for future research investigation. The function of newNodeTransaction() is created to update the medical-ledger with newly collected medical data and validate it against the daily scenario. In addition, this contract also records more details of the contract, such as patient service pService(), physician counseling (pCounseling()), new transaction (nTransaction()), updated ledger (uLedger()), blockchain timestamp [execute], and all the performed activities, shown in Contract 1.

The update ledger for immutable storage is designed and implemented to automatically update medical data whenever a new event of node transactions occurs. The updated contract function updateTransPreserve() records and evaluates the newly added details that have connected the previously-stored transactions related to medical nodes and

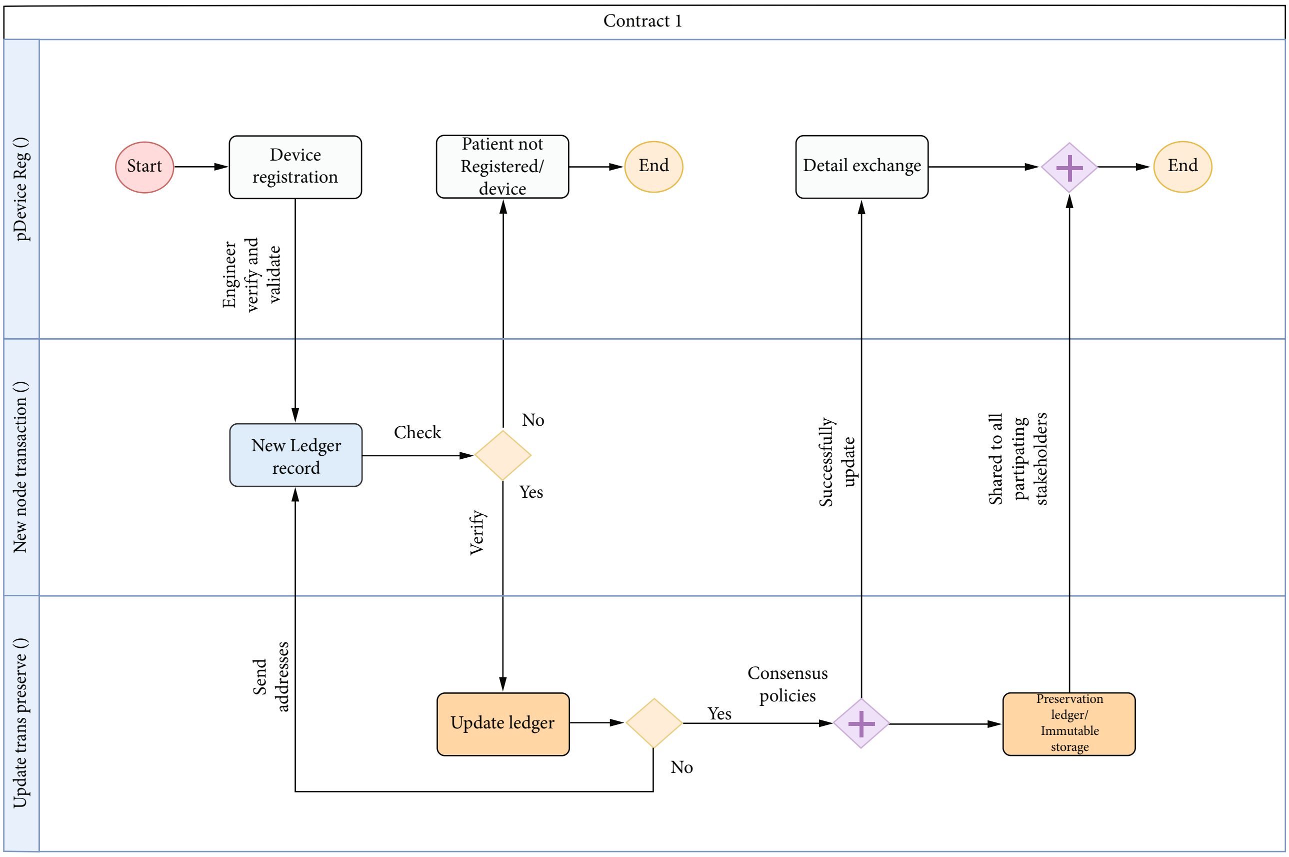


Figure 6: Operations of smart contract presented through flow of control diagram.

preservation descriptions in the distributed storage. The updateTransPreserve() contract also records the details of an updated ledger with reEncryption() to protect individual medical information, such as protecting each transaction (pETrans()), medical ledger management and privacy (mLMPrivacy()), generating hashes for individual records (gHash()), protecting storage (pStorage()), blockchain timestamp [execute], and all the activities in the immutable ledger, as shown in Appendix A (Table 2).

5. Comparison with Other State-of-the-Art Methods/Architectures/Models/Frameworks of Healthcare

The mobile-enabled healthcare application is proposed by Khan et al.; in this paper, the authors explore several related kinds of literature and present a viewpoint regarding artificial intelligence (AI) and big data analytics [38]. The purpose is to improve the process of the mobile-based medical system and patients' transactions for utilizing e-healthcare services. The collaborative strategy of AI and big data analysis is important with respect to the source of medical data, the process of filtration for information retrieval, and create knowledge towards wisdom. The applications of the collaborative approach provide insight to the patients and enable service plans and allow patients to manage services and schedules. Medical scheduler handles cost-efficient transactions between parties based on AI and metaheuristic-enabled techniques, such as consultant and patient. However, these systems are unsecure and unprotected in nature, such as server-based medical transactions and service deliveries, public networks, and most importantly the data tampering and forgery (unauthorize access and information integrity issues).

Alotaibi presented the current status of the Indian States and updated the health system regarding the use of point-to-care devices and efficient diagnosing and highlight healthcare system impact on this pandemic (COVID-19) [39]. As compared to conventional clinical devices, point-of-care devices are provided a solution in terms of acquiring particularly clinical information with less amount of cost and managing settings of medical resources and limitations. Although there is a lot of improvement in the healthcare diagnostic. Still, the use of point-of-care devices is in its nascent phase. However, there are various state-of-the-art solutions proposed in past few years, some of them which are related to the E-healthcare DIKW process (as mentioned in Table 3) are discussed as follows.

5.1. Open Research Issues, Challenges, and Limitations. In this domain, we discuss the proposed medical information management and ledger privacy protection-related healthcare distributed application limitations and challenges. Also, we mention and explain some critical aspects of medical analysis in the existing e-healthcare systems as follows.

Table 2: Contract 1: implementation of smart contracts for medical information management and privacy.

```
System constraints and initialization: blockchain hyperledger fabric-enabled healthcare medical information engineer system manage (pDeviceReg())
        Start e-healthcare distributed applications
        Schedule addresses and manage
        Preserve changes in the ledger
Data and constant: blockchain hyperledger fabric-enable engineer initiate process of edger/device registration and receive request via e-healthcare distributed applications
        Activities addresses schedule accordingly
        Int main():
        Type.File[a.txt];
        Device ID,
        (dID());
        Patient device registration,
        (pdReg());
        Patient ID,
        (pID());
        Patient name,
        (pName());
        Blockchain timestamp,
        [execute];
        Hyperledger fabric engineer maintain all the registration addresses,
        Records validation details,
        Update ledger,
        Counter (each time when new event occur);
        The engineer examine, analysis, verify, validate, and records all the details of patients' devices,
        Responsible and authorized set of nodes;
If        int main():
            Type.File[a.txt] = blockchain fabric engineer (true)
Then,
        If check device ID = true
        Then, change state of ledger
        And records additional details, device ID (dID()), patient device registration (pdReg()),
        Patient ID (pID()), patient name (pName()), blockchain timestamp [execute];
        Else
        Record, maintain ledger, error generation, change state,
        Traceback,
        Terminate;
Else
        Record, maintain ledger, error generation, change state,
        Traceback,
        Terminate;
Output: edger registration edgeregister()/pDeviceReg();
System constraints and initialization: blockchain hyperledger fabric-enabled healthcare medical information engineer system manage ((newNodeTransaction()) and (updateTransPreserve()))
        Start e-healthcare distributed applications
        Schedule addresses and manage
        Preserve changes in the ledger
Data and constant: blockchain hyperledger fabric-enable engineer initiate process of new transactions information and receive request via e-healthcare distributed applications
        Activities addresses schedule accordingly
        Int main ():
        Type.File[a.txt];
        Patient service,
        (pService());
        Physician counseling,
        (pCounseling());
        New transaction,
        (nTransaction());
        Update ledger,
        (uLedger());
        Blockchain timestamp,
        [execute];
        Hyperledger fabric engineer maintain all the new nodes transactions addresses,
        Records verification and validation details,
        Update ledger,
        Counter +1 (each time when new event occur);
        Engineer examines, analyses, verifies, validates, and records all the details of patients' devices,
        Responsible and authorized set of nodes;
If int main():
        Type.File[a.txt] = blockchain fabric engineer (true)
Then,
        If check device ID = true
        Then change state of ledger
        And records additional details, patient service pService(), physician counseling (pCounseling()), new transaction (nTransaction()), update ledger (uLedger()), blockchain timestamp [execute];
        Else
        Record, maintain ledger, error generation, change state,
        Traceback,
        Terminate;
Else
        Record, maintain ledger, error generation, change state,
        Traceback,
        Terminate;
Output: add new nodes transactions (newNodeTransaction()); and update nodes transactions
System constraints and initialization: blockchain hyperledger fabric-enabled healthcare medical information engineer system manage (hashRecord()))
        Start e-healthcare distributed applications
        Schedule addresses and manage
        Preserve changes in the ledger
Data: blockchain hyperledger fabric-enable engineer initiate process of hash-based re-encryption for e-healthcare distributed applications-related ledger security
        Activities addresses schedule accordingly
        Int main ():
        Type.File[a.txt];
        Protect each transaction,
        (pETrans());
        Medical ledger management and privacy,
        (mLMPrivacy());
        Generate hashes for individual record,
        (gHash());
        Protect storage,
        (pStorage());
        Blockchain timestamp,
        [execute];
If int main():
        Type.File[a.txt] = blockchain fabric engineer (true)
Then,
        If check device ID = true
        Then change state of ledger
        And records additional details, protect each transaction (pETrans()), medical ledger management and privacy (mLMPrivacy()), generate hashes for individual record (gHash()), protect storage (pStorage()), blockchain timestamp [execute];
        Else
        Record, maintain ledger, error generation, change state,
        Traceback,
        Terminate;
Else
        Record, maintain ledger, error generation, change state,
        Traceback,
        Terminate;
Output: ReEncrption (hashRecord());
```

5.1.1. Cross-Chain Interoperability Issue. Recently, most of the massive organizations that have large-scale data management requirements are moving to adopt blockchain hyperledger enabling technology for modular architectural solutions [44]. For this purpose, there is no specific platform specifically available and no effective protocols that achieve proper exclusivity. Interoperability issues are one of the challenging aspects. The cross-chain blockchain interoperable solution is required for designing a medical distributed enabling ecosystem. Provenance and transparency are required in the design, implementation, and deployment of blockchain hyperledger fabric-enabled medical information management and protection [45].

The restriction of node capacity in terms of size and gazette of scalability and protocols makes the e-healthcare application more reliable, such as node-time and increased security [46, 47]. Implementing a cross-chain solution that reduces transactional costs and improves communication with the fully connected network. The lack of direct connectivity between more than two distributed chains of healthcare makes interoperable communication within the ecosystem. This cross-chain technology facilitates blockchain hyperledger-enabled healthcare applications with distributed serverless transactions across different chains of nodes without involving vendor technology. Medical relays, atomic swaps, stateless server and scalability management, collaborative consensus mechanisms, and compliance for secure and protected chronological chains interconnective platforms must be considered by large healthcare sectors. Moreover, there are still some unsolved interoperable problems in the healthcare field, for example, transactional trust and the rate of bottlenecks in serverless node transactions.

5.1.2. Scope of Medical Record Privacy and Sustainability. In recent years, the healthcare system has drastically altered the environment of medical assessment and service delivery. As the system becomes digitalized, the scope of medical information and privacy has become a concerning issue because of its distinct vulnerability and the number of attacks that are increasing. For this purpose, a secure system is required for the protection of medical information to maintain patients' records, physician information, and consultant and hospital ledger preservation. Significantly, a network that stores large amounts of sensitive medical information exchanged between various medical engineers creates a forgery opportunity for information leakage and alteration [48, 49]. While the blockchain hyperledger fabric enabled distributed platform provides incentives for e-healthcare application adaptation, it also increases the associated patients' records security and privacy rights under the policy and regulation, designs the blockchain immutable structure, and deploys it for ledger preservation in a permissioned network [39, 47]. And so, it creates a new burden for engineers to tackle regulatory compliance-related issues, and private network enables intelligence verification and validation limitations and pertains to healthcare information sustainability management and challenges.

5.1.3. Sensitive Medical Information Protection and Scalability Limitations. Information concerning healthcare means sensitive personal records related to cognitive and physical health information, such as potential services of medical ledger details for physician consoling. Without getting any details, the physician cannot initiate treatment. These records need to be fully protected and cannot reveal any type of data about the patients. The types of medical information that fall under a critical category are information concerning cognitive and physical health. At the same time, handling large numbers of patients' records is a big task for cloud engineers, in which the data is continuously added to the cloud storage. To tackle such kinds of problems, we used blockchain hyperledger fabric technology that provides information integrity, transparency, provenance, immutability, and ledger scalability. Moreover, blockchain reencryption hash-based algorithms ensure the protection of medical information and also retain information confidentiality while sharing with the connected stakeholders in the private network [45, 46, 48]. However, in a permissioned private network, the fabric engineers are responsible for device registration, managing addresses of individual events and preserving all the details in the distributed storage. Dynamic management of engineer activities and smart scheduling is still an active problem in the hyperledger fabric.

5.1.4. Compliance and Policy Management Related Challenges. The various problems associated with the existing e-healthcare systems include errors in the digital medical services related transactions, such as record-keeping in centralized server-based storage and relying on cloud-based storage and related security scalability solutions [44, 49]. In addition, inappropriate and unreliable tools are used to protect medical record integrity and collect patients' transactions from portable, ubiquitous devices, and after analysis, they are submitted to the ledger storage through different network communication protocols, which is an insecure strategy.

Table 3: Comparison with other state-of-the-art proposed methods.

Other state-of-the-art methods	Research description	Research objectives and contributions	Comparison with the proposed BIoMT modular architecture
A blockchain-enabled healthcare system (HSBC) proposed for revocable attribute-based digital signature and access control [40, 41]	The highlight of this paper is (i) Proposed attribute-based signature scheme (ii) With attribute revocation (iii) For the purpose to protect the privacy of the registered patients (iv) Patient's identity in HS-BC	The main features of the proposed model are discussed as follows: (i) Security: blockchain (ii) Network: public network (iii) Ledger protection mechanism: blockchain-based predefined protection (iv) Hyperledger: no hyperledger (v) Consensus: predefined (vi) Node size: not defined (vii) Storage: cloud storage (viii) Response: not applicable (ix) Transactions executions delay: not applicable (x) User: patients	The proposed blockchain hyperledger fabric-enabled secure distributed e-healthcare architecture is designed for scheduling and managing DIKW medical processes in a protected manner. The main attributed and architectural features are defined as follows: (i) Security: blockchain-enabled privacy and security (ii) Network: consortium network structure (iii) Ledger protection mechanism: hash-based encryption (SHA-256) (iv) Hyperledger: fabric (v) Consensus: customized consensus policies (shown in contract 1) (vi) Node size: variable in between 2-4 MB (vii) Storage: IPFS (viii) Response: depend of traffic/direct (ix) Transactions executions delay: less delay (x) User: distributed e-healthcare registered patients
A secure and scalable control policy and access management for healthcare system using collaborative blockchain, IoT, and artificial intelligence techniques [41–43]	The paper discussed the collaborative nature and the impact of the current e-healthcare systems. The contribution of this paper are as follows: (i) An enhanced Bell–LaPadula is used to scalable digital ledger (ii) Dynamic access control policies developed by creating smart contracts using blockchain (iii) Provide dynamic access control and functionality (iv) Other state-of-the-art is used artificial neural network technique for classification of medical records for focusing on the personal healthcare records (v) IoT-blockchain-enabled real-time monitoring and medical diagnostic-based on four layers of data processes	The critical characteristics and attributes of the proposed model are discussed as follows: (i) Security: blockchain (ii) Network: public network (iii) Ledger protection mechanism: blockchain-based predefined protection (iv) Hyperledger: no hyperledger (v) Consensus: predefined (vi) Node size: not defined (vii) Storage: cloud storage (viii) Response: not applicable (ix) Transactions executions delay: not applicable (x) User: patients	

6. Conclusion and Future Work

This paper discusses the core concepts of data, information, knowledge, and wisdom and their management and privacy-related issues in the current e-healthcare systems using a centralized database. The layered hierarchical process of DIKW collaborates with the blockchain hyperledger fabric to secure the process of scheduling and management. One of the critical limitations is the protection of sensitive medical information through real-time distributed processing, management, and monitoring. The current scenario of information management and privacy of e-healthcare applications has gaps and challenges, including two-way authentication issues, the event of node transactions execution, adding new or updated transactions approval, and preservation in secure storage. It is proposed to use blockchain hyperledger fabric for unified dynamic medical information management and DIKW hierarchical processing. In this paper, we have also added three different folds, which highlight the main contributions of this paper, such as the blockchain hyperledger fabric-enabled information management process. So, smart contracts are for privacy (hash-encrypted SHA-256) and preservation and an efficient blockchain P2P communication (hybrid channel). A hyperledger fabric is provided with a modular infrastructure, which enables the process of capturing data from the private channel, managing all the nodes' transactions between stakeholders, and storing medical ledgers in the IPFS distributed data storage. A detailed design of information management by a blockchain hyperledger fabric engineer is substantial, including raw medical data capture, examination to drive information, analysis of individual aspects to form knowledge, refining knowledge to present it in the form of wisdom, and storage of the ledger. We also designed and deployed chaincode (smart contracts) to secure and protect the ledger and patients' device credentials. Three contracts are created for efficient P2P communication, such as device registration (pDeviceReg()), adding new node transactions (newNodeTransaction()), and updating transactions and preservation (updateTransPreserve()). The deployment of a blockchain-enabled secure distributed DIKW architecture is becoming a necessity for the healthcare sector. The private network deals with the efficient execution of transactions, privacy, and security-related challenges of medical nodes. Furthermore, the proposed architecture maintains overall events of node integrity, provenance, transparency, and immutability and provides effective performance in information management, monitoring, and privacy. The working of events of node transactions and preservation in an action of this architecture is presented through the use case diagram. While designing the proposed BIoMT modular architecture, we highlight and separate a number of open issues that need expert concern. These become our fundamental objectives where we are expected to work in the future and provide possible solutions to the emerging issues and challenges discussed in this paper.

Data Availability

Data sharing is not applicable to this research work as no new data (simulated) were created or analyzed in this paper.

Disclosure

The sponsors have not been involved in this research work design and implementation, data collection, analysis, or decision-making related to the publication or preparation of the paper.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

A.A.K. have written original draft and preparation. A.A.K, Z.A.S., L.T., A.A.W., and A.A.L. have reviewed, rewrote, performed part of the literature survey, and edited, investigated, designed the architecture, and explored software tools. All authors of this paper read and agreed to the published version (online) of this paper.

Acknowledgments

This paper is partially funded by Lin Teng (Software Collage, Shenyang Normal University).

References

[1] Y. Duan, E. Kajan, and Z. Maamar, *Crossing "Data, Information, Knowledge, and Wisdom" Models—Challenges, Solutions, and Recommendations*, Information (Mdpi), 2022.

[2] M. Alisie, "Blockchain and the evolution of information society," *Theories of Change: Change Leadership Tools, Models and Applications for Investing in Sustainable Development*, pp. 351–374, 2021.

[3] A. A. Khan, A. A. Laghari, A. A. Shaikh, M. A. Dootio, V. V. Estrela, and R. T. Lopes, "A blockchain security module for brain-computer interface (BCI) with multimedia life cycle framework (MLCF)," *Neuroscience Informatics*, vol. 2, no. 1, article 100030, 2022.

[4] Z. Sardar, "The smog of ignorance: knowledge and wisdom in postnormal times," *Futures*, vol. 120, article 102554, 2020.

[5] M. Hussain, F. A. Satti, S. I. Ali et al., "Intelligent knowledge consolidation: from data to wisdom," *Knowledge-Based Systems*, vol. 234, article 107578, 2021.

[6] P. Müürsepp, "Making sense of wisdom management," *International Journal for Applied Information Management*, vol. 1, no. 2, pp. 63–69, 2021.

[7] A. A. Khan, A. A. Shaikh, O. Cheikhrouhou et al., "IMG-forensics: multimedia-enabled information hiding investigation using convolutional neural network," *IET Image Processing*, vol. 16, no. 11, pp. 2854–2862, 2022.

[8] M. Jakubik and P. Müürsepp, "From knowledge to wisdom: will wisdom management replace knowledge management?," *European Journal of Management and Business Economics*, vol. 31, no. 3, pp. 367–389, 2022.

[9] A. A. Khan, Z. A. Shaikh, L. Baitenova et al., "QoS-ledger: smart contracts and metaheuristic for secure quality-of-service and cost-efficient scheduling of medical-data processing," *Electronics*, vol. 10, no. 24, p. 3083, 2021.

[10] K. Ali, Z. A. Shaikh, A. A. Khan, and A. A. Laghari, "Multiclass skin cancer classification using EfficientNets - a first step towards preventing skin cancer," *Neuroscience Informatics*, vol. 2, no. 4, article 100034, 2022.

[11] L. Tamine and L. Goeuriot, "Semantic information retrieval on medical texts," *ACM Computing Surveys*, vol. 54, no. 7, pp. 1–38, 2022.

[12] M. G. Rhodes, K. E. Fletcher, F. Blumenfeld-Kouchner, and E. A. Jacobs, "Spanish medical interpreters' management of challenges in end of life discussions," *Patient Education and Counseling*, vol. 104, no. 8, pp. 1978–1984, 2021.

[13] V. Chang, P. Baudier, H. Zhang, X. Qianwen, J. Zhang, and M. Arami, "How Blockchain can impact financial services–

the overview, challenges and recommendations from expert interviewees," *Technological Forecasting and Social Change*, vol. 158, article 120166, 2020.

[14] U. Bodkhe, S. Tanwar, K. Parekh et al., "Blockchain for industry 4.0: a comprehensive review," *IEEE Access*, vol. 8, pp. 79764–79800, 2020.

[15] D. Berdik, S. Otoum, N. Schmidt, D. Porter, and Y. Jararweh, "A survey on blockchain for information systems management and security," *Information Processing & Management*, vol. 58, no. 1, article 102397, 2021.

[16] S. Liu, Y. Dai, Z. Cai, X. Pan, and C. Li, "Construction of double-precision wisdom teaching framework based on blockchain technology in cloud platform," *IEEE Access*, vol. 9, pp. 11823–11834, 2021.

[17] J. Ducrée, "Research - a blockchain of knowledge?," *Blockchain: Research and Applications*, vol. 1, no. 1-2, article 100005, 2020.

[18] R.-G. J. Pablo, D.-P. Roberto, S.-U. Victor, G.-R. Isabel, C. Paul, and O.-R. Elizabeth, "Big data in the healthcare system: a synergy with artificial intelligence and blockchain technology," *Journal of Integrative Bioinformatics*, vol. 19, no. 1, 2022.

[19] W. Chen, "Exploration and practice of university students health education promotion model under big data information," in *EAI International Conference, BigIoT-EDU*, pp. 375–384, Cham, 2021.

[20] R. Huang, Y. Jiang, and X. Le, "Prevention and nursing research of PICC catheter-related complications in patients with digestive system malignant tumor based on smart medical block chain," *Journal of Healthcare Engineering*, vol. 2021, Article ID 5519722, 11 pages, 2021.

[21] C. A. Ardagna, R. Asal, E. Damiani, N. El Ioini, M. Elahi, and C. Pahl, "From trustworthy data to trustworthy IoT," *ACM Transactions on Cyber-Physical Systems*, vol. 5, no. 1, pp. 1–26, 2021.

[22] S. M. Pournaghi, M. Bayat, and Y. Farjami, "MedSBA: a novel and secure scheme to share medical data based on blockchain technology and attribute-based encryption," *Computing*, vol. 11, no. 11, pp. 4613–4641, 2020.

[23] H. Huang, F. Peng Zhu, X. S. Xiao, and Q. Huang, "A blockchain-based scheme for privacy-preserving and secure sharing of medical data," *Computers & Security*, vol. 99, article 102010, 2020.

[24] X. Yang, T. Li, X. Pei, L. Wen, and C. Wang, "Medical data sharing scheme based on attribute cryptosystem and blockchain technology," *IEEE Access*, vol. 8, pp. 45468–45476, 2020.

[25] X. Cheng, F. Chen, D. Xie, H. Sun, and C. Huang, "Design of a secure medical data sharing scheme based on blockchain," *Journal of Medical Systems*, vol. 44, no. 2, pp. 1–11, 2020.

[26] A. A. Abdellatif, L. Samara, A. Mohamed et al., "MEdge-chain: leveraging edge computing and blockchain for efficient medical data exchange," *IEEE Internet of Things Journal*, vol. 8, no. 21, pp. 15762–15775, 2021.

[27] I. Mahmood and H. Abdullah, "WisdomModel: convert data into wisdom," *Applied Computing and Informatics*, 2021.

[28] K. D. Cato, K. McGrow, and S. C. Rossetti, "Transforming clinical data into wisdom," *Nursing Management*, vol. 51, no. 11, pp. 24–30, 2020.

[29] S. Khan and M. Shaheen, "From data mining to wisdom mining," *Journal of Information Science*, 2021.

[30] D. Alvarez-Coello, D. Wilms, A. Bekan, and J. M. Gómez, "Towards a data-centric architecture in the automotive industry," *Procedia Computer Science*, vol. 181, pp. 658–663, 2021.

[31] D. Yin, X. Ming, and X. Zhang, "Understanding data-driven cyber-physical-social system (D-CPSS) using a 7C framework in social manufacturing context," *Sensors*, vol. 20, no. 18, p. 5319, 2020.

[32] Y. Yao, "Tri-level thinking: models of three-way decision," *International Journal of Machine Learning and Cybernetics*, vol. 11, no. 5, pp. 947–959, 2020.

[33] X. Peng and W. Bian, "How is data-driven precision teaching possible? From the perspective of cultivating teacher's data wisdom," *Journal of East China Normal University (Educational Sciences)*, vol. 39, no. 8, p. 45, 2021.

[34] A. Ayub Khan, A. A. Laghari, A. A. Shaikh, S. Bourouis, A. M. Mamlouk, and H. Alshazly, "Educational blockchain: a secure degree attestation and verification traceability architecture for higher education commission," *Applied Sciences*, vol. 11, no. 22, 2021.

[35] I. Abu-Elezz, A. Hassan, A. Nazeemudeen, M. Househ, and A. Abd-Alrazaq, "The benefits and threats of blockchain technology in healthcare: a scoping review," *International Journal of Medical Informatics*, vol. 142, p. 104246, 2020.

[36] A. A. Khan, Z. A. Shaikh, A. A. Laghari, S. Bourouis, A. A. Wagan, and G. A. Ali, "Blockchain-aware distributed dynamic monitoring: a smart contract for fog-based drone management in land surface changes," *Atmosphere*, vol. 12, no. 11, p. 1525, 2021.

[37] A. A. Khan, A. A. Laghari, T. R. Gadekallu et al., "A drone-based data management and optimization using metaheuristic algorithms and blockchain smart contracts in a secure fog environment," *Computers and Electrical Engineering*, vol. 102, article 108234, 2022.

[38] A. A. Khan, A. A. Laghari, D. S. Liu et al., "EPS-ledger: blockchain Hyperledger Sawtooth-enabled distributed power systems chain of operation and control node privacy and security," *Electronics*, vol. 10, no. 19, p. 2395, 2021.

[39] S. R. Alotaibi, "Applications of artificial intelligence and big data analytics in m-health: a healthcare system perspective," *Engineering*, vol. 2020, article 8894694, pp. 1–15, 2020.

[40] A. N. Konwar and V. Borse, "Current status of point-of-care diagnostic devices in the Indian healthcare system with an update on COVID-19 pandemic," *Sensors International*, vol. 1, article 100015, 2020.

[41] Q. Su, R. Zhang, R. Xue, and P. Li, "Revocable attribute-based signature for blockchain-based healthcare system," *IEEE Access*, vol. 8, pp. 127884–127896, 2020.

[42] A. A. Khan, A. A. Wagan, A. A. Laghari, A. R. Gilal, I. A. Aziz, and B. A. Talpur, "BIoMT: a state-of-the-art consortium serverless network architecture for healthcare system using blockchain smart contracts," *IEEE Access*, vol. 10, pp. 78887–78898, 2022.

[43] S.-K. Kim and J.-H. Huh, "Artificial neural network blockchain techniques for healthcare system: focusing on the personal health records," *Electronics*, vol. 9, no. 5, p. 763, 2020.

[44] T. Alam, "Blockchain-enabled mobile healthcare system architecture for the real-time monitoring of the COVID-19 patients," 2021.

[45] S. Namasudra, G. C. Deka, P. Johri, M. Hosseinpour, and A. H. Gandomi, "The revolution of blockchain: state-of-the-art and research challenges," *Archives of Computational Methods in Engineering*, vol. 28, no. 3, pp. 1497–1515, 2021.

[46] N. Upadhyay, “Demystifying blockchain: a critical analysis of challenges, applications and opportunities,” *International Journal of Information Management*, vol. 54, article 102120, 2020.

[47] S. Dhar, A. Khare, and R. Singh, “Advanced security model for multimedia data sharing in internet of things,” *Transactions on Emerging Telecommunications Technologies*, no. article e4621, 2022.

[48] G. Srivastava, J. Crichigno, and S. Dhar, “A light and secure healthcare blockchain for iot medical devices,” in *2019 IEEE Canadian conference of electrical and computer engineering (CCECE)*, pp. 1–5, Edmonton, AB, Canada, 2019.

[49] A. D. Dwivedi, “Brisk: dynamic encryption based cipher for long term security,” *Sensors*, vol. 21, no. 17, p. 5744, 2021.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 4389729, 13 pages
https://doi.org/10.1155/2022/4389729

Research Article

Authenticated Wireless Links between a Drone and Sensors Using a Blockchain: Case of Smart Farming

Kahlid S. Alqarni,[1] **Faris A. Almalki**,[2] **Ben Othman Soufiene**,[3] **Obaid Ali**,[4] **and Faisal Albalwy**[5,6]

[1]*Department of Information Technology, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia*
[2]*Department of Computer Engineering, College of Computers and Information Technology, Taif University, Taif 21944, Saudi Arabia*
[3]*PRINCE Laboratory Research, ISITcom, Hammam Sousse, University of Sousse, Tunisia*
[4]*Department of Computer Science & Information Technology, Ibb University, Ibb, Yemen*
[5]*Department of Computer Science, College of Computer Science and Engineering, Taibah University, Madinah, Saudi Arabia*
[6]*Division of Informatics, Imaging and Data Sciences, Stopford Building, University of Manchester, Oxford Road, Manchester M13 9PL, UK*

Correspondence should be addressed to Obaid Ali; obaid.alii2016@gmail.com

Received 23 June 2022; Revised 3 August 2022; Accepted 25 August 2022; Published 5 September 2022

Academic Editor: Yin Zhang

Agriculture is confronted with several significant difficulties, such as rising air temperatures and population growth, causing the implementation of smart farming operations as an optimum solution. This research aims to contribute to the growing knowledge of the potential role of blockchain technology in promoting the concept of smart farming by enhancing the efficiency of farming operations by boosting agricultural production, lowering environmental impact, and automating the work of farmers. It proposes a secure blockchain-based framework to establish trust among smart farming users. The framework utilizes asymmetric key exchange mechanism using an ECC authentication algorithm and SHA-256 hash function cryptography to secure communication between sensors and drones in the farm field. The SHA-256 hashing function ensures data integrity as attempts to tamper with data result in a different hash value, breaking the chain of blocks. To demonstrate the feasibility of the proposed framework, a proof-of-concept implementation was developed on the Ethereum blockchain, in which smart contracts were used to model the framework operations. The proof of concept's performance was examined using Hyperledger Caliper for latency, throughput, and transaction success rate. The findings clearly indicate that blockchain technology can provide an efficient and scalable mechanism to advance smart farming and address some of the barriers that inhibit smart farming, particularly regarding to data integrity and availability.

1. Introduction

Smart farming refers to the use of various technologies and gadgets, such as the Internet, cloud, and IoT devices. By 2050, the world's population is projected to reach 9.7 billion people, requiring greater agricultural production to feed those billion people [1]. Because of causes like as industrialization, commercial marketplaces, and residential structures being constructed on agricultural areas, the population is increasing, while agricultural land is decreasing. Using these works need to be boosted for output to feed these billions, which can be done by integrating IoT in farming as shown in Figure 1. Due to several reasons, including insect attacks, plant disease, a lack of knowledge about essential nutrients for crops, and other problems, farmers are no longer able to enjoy the benefits of their work. To eliminate these obstacles and make farming more profitable, smart, and enjoyable for farmers, technological advancement is needed [2]. In

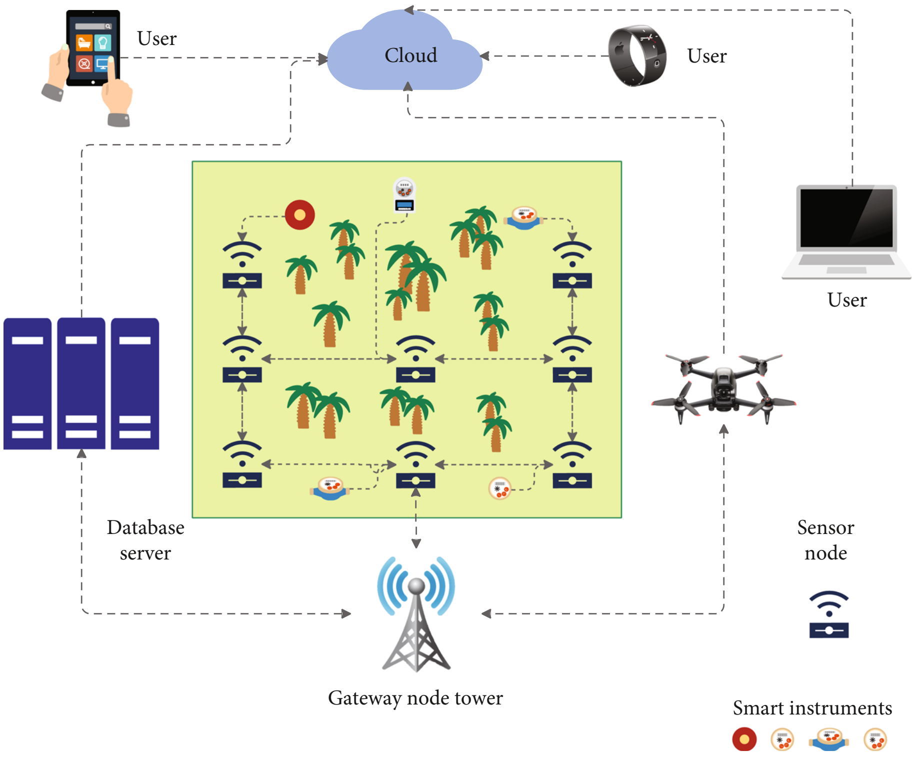


Figure 1: Example of the smart farming application.

every way, smart farming and conventional farming are diametrically opposed. Without regard for market demand, rates, weather predictions, or other variables, traditional farming uses historic and traditional agricultural methods, as well as antiquated equipment for labor and seasonal crop production. Smart farming takes use of modern technology like smart linked devices, Internet of Things sensors, a farmers' chat room, and continuous evaluations of different variables like the optimum circumstances for a plant to develop, the quantity of nutrients needed, soil quality, and water quality monitoring. Smart farming lowers labor costs, boosts crop yields, and enhances production while making farming easy and cheap (cost-effective). Agriculture has progressed to the point where smart farming is the next stage. The use of the Internet of Things (IoT) and unmanned aerial vehicle (UAV) technologies to enhance the efficiency of agricultural operations is referred to as smart farming [3].

A UAV may autonomously and properly reacts to its surroundings based on its context. Agricultural products must be increased mostly because to the rapid expansion of the global population, despite substantial contributions from scientific discoveries in genetics, chemistry, and robotics to the improvement of agricultural technology [4]. At the same time, the agricultural sector is confronted with major challenges such as climate change, land scarcity, and the growing need for freshwater. Information and Communication Technology (ICT) services may be a potential solution to these pressing issues.

UAVs and the IoT are two of the most popular technologies being used for civilian and industrial reasons, as well as to support Industry 4.0. An unmanned aerial vehicle (UAV) is a remotely controlled autonomous vehicle that does not need a human pilot. Unmanned aerial vehicles (UAVs) were originally developed for military purposes, but their growing popularity and technological developments have highlighted their potential for civilian and industrial applications [5, 6]. The Internet of Things (IoT) allows a large number and diversity of linked devices, allowing for remote monitoring and control of various activities [7]. Smart homes or home automation, smart cities, security and surveillance applications, remote patient monitoring, and precision agriculture are all examples of IoT use cases [8–10].

The combination of these two technologies (UAV and IoT) expands the number of options for improving people's lives. Data collection operations in UAV-based applications may be aided by a well-implemented Internet of Things architecture. While UAVs may help gather data from difficult places for IoT applications, the usage of sensor-equipped UAVs in municipal and industrial applications is expanding IoT's power. According to research, UAV-enabled IoT systems might be utilized for a variety of interesting and helpful applications. UAV and IoT integration applications are aimed at smart cities, agriculture, healthcare, disaster management, rescue operations, supply chains, and geoscience [11–13]. Combining UAVs with IoT has a lot of promise, but it also has a lot of technical and legal issues. Examples of applications include air traffic control, obstacle detection, flight schedule and path integrity, the use of different communications designs, data collection via sensors, actual or near real-time data analysis and delivery, and lightweight encryption algorithm to align with restricted onboard resources.

As an emerging technology, blockchain applications are being explored in various industries, including healthcare [14–17], finance [18, 19], real estate [20, 21], agriculture [22, 23], and education [24, 25]. Blockchain implementation is ideal for communication networks, thanks to new improvements in blockchain technology such as decentralization, immutability, security, and transparency. A blockchain is an immutable database that nodes in a distributed and decentralized peer-to-peer network continually update and agree on [26]. Elliptic-curve Public-Key Cryptography is the most prevalent public-key cryptographic technique used in blockchain technology (ECC). This technique has an advantage over Public-Key Cryptography (PKC) in that the authentication and transparency of new transactions are dependent on a widespread agreement among its users. As a result, deploying blockchain technology for distributed UAV networks might provide a slew of security advantages.

The main contribution of this paper is threefold. Firstly, it proposed a novel blockchain-based framework to support Authenticated Wireless Links between a Drone and Sensors. Secondly, it demonstrated a proof-of-concept implementation for the proposed framework by walking through an intelligent farming case study. Lastly, it provided a performance evaluation of the implemented proof of concept.

2. Related Works

The UAV networks' drones can communicate with one another over a wireless link. UAV networks are vulnerable to forgery attacks, man-in-the-middle attacks, and reply to assaults due to their low computing capacity and complicated external environment. Before the drones may communicate with each other, identity authentication is critical, and assuring a legal drone in the network is the top priority of UAV network security. Traditional authentication mechanisms based on username/password or dynamic key have a low level of security. RSA certification necessitates the use of a lengthy session key, which is incompatible with the lightweight requirements of UAV networks. Many security issues are avoided by blockchain's decentralization and secure communications using public cryptography.

This research [9] proposed VAHAK, an Ethereum blockchain-based secure outdoor healthcare medical supply using UAVs. VAHAK enables timely delivery of important medical supplies to critical patients by facilitating decentralized communication between UAVs and entities. In VAHAK, the Ethereum smart contract was utilized to address concerns about security, privacy, and dependability, while the IPFS protocol was used to address storage costs. The VAHAK's security vulnerabilities are tested using the open-source application MyThril. VAHAK is cost-effective in terms of data storage since it uses the InterPlanetary File System (IPFS) for healthcare record storage and 5G-enabled Tactile Internet (TI) for communication. Finally, when compared to conventional systems, VAHAK performance assessment outperformed existing approaches in various performance evaluation parameters such as scalability, latency, and network capacity.

Authors in [27] suggested a blockchain-based intelligent technique for securing the privacy of unmanned aerial vehicles (UAVs) and drones. They presented the hashing process and how it was used in their system, including the creation of a hash code. They created a security mechanism that encrypts information using hashing by combining picture collection and sensing from drones and UAVs with blockchain security. All transactions between the server and the drone, as well as the drone's GPS position, were tracked using the timestamp.

Researchers in [28] built a blockchain-based access management system for the IoD environment, which allows for secure communication between drones and the GSS. Secure data is collected by the GSS in the form of transactions, which are subsequently converted into blocks. Finally, in a peer-to-peer cloud server network, cloud servers connected to the GSS through the Ripple Protocol Consensus Algorithm (RPCA) upload the blocks to the blockchain. After they have been added to the blockchain, the transactions in the blocks cannot be modified, edited, or even removed. They carried out several security analyses, including formal security under the random oracle model, informal security, and simulation-based formal security verification, to ensure that the proposed scheme can withstand a wide range of potential attacks with a high probability, as is required in an IoD environment.

According to Khalifeh et al. [29], a UAV might be used as a data mule to unload sensor nodes and securely transfer monitoring data to a remote control center for further analysis and decision-making. They also spoke about the challenges of putting the proposed framework into reality. Experimenting with their suggested design in the presence of different types of obstructions may be found in typical outdoor fields. During the testing, some differences between the performance metrics provided in the hardware-specific datasheets were uncovered. They uncovered disparities between the declared coverage distance and signal strength via their experiments.

A study in [30] employed a blockchain-enabled identity authentication scheme and a safe data sharing paradigm for drones. Authentication and access control are handled by smart contracts, account creation and security are handled by Public-Key Cryptography, and security auditing is handled by a distributed ledger. To speed up outsourced calculations, ABEM-POC, which is based on the Spark cluster and MapReduce architecture, is presented. The ABE and a modified approach based on ABEM-POC can be used to facilitate parallel outsourced computations. The results showed that both the ABEM-POC and general techniques were successful and straightforward to implement.

ACSUD-IoD, an access control approach for illegal UAV detection and mitigation in an IoD environment, was presented by the authors in [31]. The transactional data to the GSS was stored on a private blockchain, allowing the GSS to identify unauthorized UAVs. They used a range of security tests to show that the suggested system is resilient to a variety of potential attacks that may occur in an IoD environment. Many cryptographic primitives' effectiveness and resilience have been shown in trials.

This study in [32] offered a novel approach for safeguarding communications between unmanned aerial vehicles (UAVs) engaged in various tasks. A one-of-a-kind method for UAVs to enable network transactions without delivering encrypted communications is included in the proposed technique. They also proposed a consensus method based on the proof of communication. They concluded that the suggested approach may be used safely in communication networks.

To mitigate such attacks and achieve trust, this paper [33] proposed a new and systematic framework that combines interest-key-content binding (IKCB), forwarding strategy, and on-demand verification to investigate poisoned content quickly and effectively in NDN-based unmanned aerial vehicles ad hoc networks (UAANETs). They presented a permissioned blockchain network built on top of NDN, as well as a scalable adaptive delegate consensus approach for providing a decentralized IKCB store and detecting internal attackers. Their results show that the suggested architecture may effectively cleanse poisoned material at a low cost and that their techniques performed well enough to be suitable for UAANETs.

In this research [34], the SENTINEL architecture was proposed to facilitate mutual authentication between drones and base stations. SENTINEL generates a flight session key for a drone with a flight plan and registers the flight session key and the drone's flight plan in a centralized database accessible by all ground stations. Ground stations utilize the registered flight session key to authenticate the drone as the MAC key when it is flying. We devised a straightforward certificate format that may be utilized in IoD scenarios. The proposed certificate is designed to convey just the bare minimum of data required to construct public key infrastructure in the Internet of Things (IoT) scenarios. To reduce certificate size even further, they chose a binary format instead of the human-readable text format used in X.509 v3 certificates.

A unique task-oriented authentication method based on blockchain (ToAM) for UAVs was proposed in [35]. They divided UAV authentication into group building authentication and intragroup authentication using a two-stage authentication architecture. They also exhibited a lightweight and cross-domain authentication system based on blockchain that enables for the secure purchase of cross-domain UAVs and task group setup. The job is then performed utilizing a chord ring and a preshared key authentication protocol, which allows for quick and secure authentication inside the UAV group even when the network connection is poor.

The authors of [36] presented a mutual-healing group key distribution technique based on blockchain. The GCS group keys are stored on a private blockchain that is integrated into the Ground Control Station (GCS). Meanwhile, the blockchain was used to manage a dynamic list of UAANET membership certificates. According to different attack situations, a basic mutual-healing protocol and an upgraded one were designed based on the Longest-Lost-Chain approach to recover the node's lost group keys with the aid of its neighbors.

AKMS-AgriIoT, a private blockchain-based system (IPA) for Intelligent Precision Agriculture, was recommended in this study [37]. To confirm and add the blocks created by the encrypted transactions and their accompanying signatures by the GSS to the private blockchain center, the cloud servers mine them. According to extensive security analysis and comparative study, the recommended AKMS-AgriIoT was also given. Table 1 presents a comparison between the existing literature and the proposed framework.

3. Materials and Methods

3.1. Elliptic Curve Cryptographic Protocols (ECC). ECC is a modern family of public-key cryptosystems based on the algebraic structures of elliptic curves over finite fields and the Elliptic Curve Discrete Logarithm Problem's difficulty (ECDLP) [38–40]. ECC provides asymmetric cryptosystem features such as encryption, signatures, and key exchange. ECC is a logical successor to the RSA cryptosystem since it requires fewer keys and signatures to provide the same degree of security as RSA and allows for very fast key generation, key agreement, and signatures. In the ECC, the private keys are integers (typically 256-bit integers in the field size range of the curve). In ECC cryptography, key generation is as easy as dependably generating a random integer inside a given range, making it very quick. A valid ECC private key is an integer within the range.

The ECC's public keys are the EC points, which are pairs of integer coordinates x, y that fall on the curve. Because of their unique properties, EC points may be reduced to only one coordinate plus one bit (odd or even). As a result, the compressed public key is a 257-bit integer, which is equivalent to a 256-bit ECC private key. An ECC public key is an example (corresponding to the above private key, encoded in the Ethereum format, as hex with prefix 02 or 03). The public key needs 33 bytes (66 hex digits) in this format, which may be lowered to roughly 257 bits.

The PKC is a method of generating a pair of keys: public keys that are widely distributed and private keys that are only known by the authorized owner. This serves two purposes: authentication (the public key confirms that the message was sent by the owner of the private key) and encryption (the message can only be decoded by the owner of the associated private key). In this section, ECC has been used for PKC since it provides stronger security with shorter computation time than DSA and RSA. The purpose of Pseudocode 1 is to generate public and private key pairs for authorized users while also sharing data from the smart contract. The computation time for key creation is the same as for symmetric cryptography, but it provides data security and authentication. In the next asymmetric key algorithm of our proposal, we uses private key and public key (private key to $\mathrm{P_r}$ key, public key to $\mathrm{P_u}$) [41].

As shown in Pseudocode 2, the $\mathrm{User_A}$ encrypt message 'M' by the $\mathrm{User_B}$ public key $\mathrm{p_B}$, so that only authorized $\mathrm{User_B}$ can decrypt the message. $\mathrm{User_A}$ encrypts message 'M' using the $\mathrm{User_B}$ public key $\mathrm{P_B}$, as described in Pseudocode 2, so that only authorized $\mathrm{User_B}$ can decode the message.

Table 1: A comparison between existing literature and the proposed framework.

Article ref.	Aerial platform	Hash function cryptography	Blockchain-based	System evaluated?	Security framework	Contribution	Issues
[28]	✓	✗	✗	✗	RPCA protocol	Link between the GSS and cloud servers	High-cost complex
[30]	✓	✗	✓	✗	Public key	ABEM-POC	Retrieve key High cost
[31]	✓	✗	✓	✗	Private key	ACSUD-IoD	High-cost complex
[32]	✓	✓	✓	✗	Public key, OTP encryption	Decentralized securing communications	Not able to authenticate in weak connections
[33]	✓	✗	✗	✗	Systematic key	SDN and NFV-based fleet	Authentication in weak connections
[34]	✓	✗	✗	✗	Public key	Binary format, SENTINEL produces a flight session key	Complex, retrieve key, cost
[36]	✓	✗	✗	✗	Private key	Longest-Lost-Chain method, dynamic list certificates	Complex
[37]	✓	✗	✓	✗	Private key	AKMS-AgriIoT + AI	Cloud server cost
Proposed framework	✓	✓	✓	✓	ECC asymmetric key	SHA256 hashing + IPFS data using a smart contract	Complex

```
Input: $E_q$ (a, b), G, q
Output: generate $P_r$ key, $P_u$ key
  1. Eq (a, b): The parameter of ECC that include a,b,q where q is a prime number and form of $2^m$
  2. G: specific point on curve whose order is big value ‘n’.
  3. generate $User_A$ key and Pr key select $n_A$; $n_A < n$, where n is limitation of curve point.
  4. Calculate $P_u$ key $p_A$; $p_A = n_A * G$
  5. $User_B$ key generate, and Pr key select $n_B$; $n_B < n$, where n is limit of curve point.
  6. $P_u$ keys calculate $p_B$; $p_B = n_B * G$
  7. $User_A$ key calculate $K_A = n_A * p_B$
  8. UserB key calculate KB = nB * pA
```

Pseudocode 1: Generate ECC asymmetric Key exchange.

```
Next steps made by the $User_A$: -1. Suppose message be ‘M’
2. Encoding the message ‘M’ into a point on the elliptic curve
3. Let the point be Pm
4. Choose a random + integer ‘K’ to encrypt the point
5. Cm =K*G*Pm +KpB where G is the base point.
```

Pseudocode 2: Encrypting data using ECC.

```
The $User_B$ recipient does next steps: -1. Multiplication between first point in the pair with $User_B$ secret key
2. Compute $K_B*G*n_B$
3. subtraction it from second point in the pair
$P_m + K*p_B - (K*G*n_B)$
$P_m + K*p_B - (K * p_B) = P_m$, where [ $n_B*G=p_B$]
```

Pseudocode 3: Decrypting data using ECC.

0x428a2f98	0x71374491	0xb5c0fbcf	0xe9b5dba5	0x3956c25b	0x59f111f1	0x923f82a4	0xab1c5ed5
0xd807aa98	0x12835b01	0x243185be	0x550c7dc3	0x72be5d74	0x80deb1fe	0x9bdc06a7	0xc19bf174
0xe49b69c1	0xefbe4786	0x0fc19dc6	0x240ca1cc	0x2de92c6f	0x4a7484aa	0x5cb0a9dc	0x76f988da
0x983e5152	0xa831c66d	0xb00327c8	0xbf597fc7	0xc6e00bf3	0xd5a79147	0x06ca6351	0x14292967
0x27b70a85	0x2e1b2138	0x4d2c6dfc	0x53380d13	0x650a7354	0x766a0abb	0x81c2c92e	0x92722c85
0xa2bfe8a1	0xa81a664b	0xc24b8b70	0xc76c51a3	0xd192e819	0xd6990624	0xf40e3585	0x106aa070
0x19a4c116	0x1e376c08	0x2748774c	0x34b0bcb5	0x391c0cb3	0x4ed8aa4a	0x5b9cca4f	0x682e6ff3
0x748f82ee	0x78a5636f	0x84c87814	0x8cc70208	0x90befffa	0xa4506ceb	0xbef9a3f7	0xc67178f2

Figure 2: A hash function mechanism with length of 256 bits.

In Pseudocode 3 [41], the message was decrypted by User_B using the private key K_B. Because of PKC's secret key mechanism, the message's originality cannot be tampered with.

3.2. Hash Function Cryptography. Several hash functions are widely used in different applications, including MD5, SHA-160, and SHA-256. The MD5 produces a 128-bit hash value, whereas the SHA-160 and SHA-256 produce a 160-bit and a 256-bit hash value, respectively. Some hash functions have demonstrated weaknesses throughout further research, though all are considered adequate for noncryptographic applications. For instance, vulnerabilities were discovered in MD5, and it is no longer recommended for cryptographic applications, but it is still used to validate file transfers and database partitioning [42]. Similarly, vulnerabilities were discovered in SHA-160, which is no longer recommended for cryptographic applications [43]. On the other hand, the SHA-256 is recommended by the National Institute of Standards and Technology (NIST) to use instead of MD5 or SHA-160 for cryptographic applications [44].

SHA-256 (secured hashing, FIPS 182-2) is a 256-bit digest cryptographic algorithm. It is an MDC or a unique hash function (Manipulation Detection Code) [45]. A message is broken down into $512 = 16 \times 32$-bit blocks, with each block taking 64 rounds [46, 47]. The 32 initial bits of the fractions portions of the cube roots of the first 64 prime integers gives us the 64 binary characters Ki as shown in Figure 2.

4. The Proposed Framework

4.1. Framework Design. There are many different interpretations of security. Confidentiality, which prevents unauthorized release of information, integrity, which prevents illegal change or deleting data, and availability, which prevents unauthorized withhold of data, make up security [48].

One of the most important aspects of any information system is data integrity. Protecting data from illegal alteration, deletion, or fabrication is known as data integrity. Managing an entity's access and rights to certain corporate resources helps to guarantee that sensitive data and services are not misused, misappropriated, or stolen. Authorization is a method of restricting data access. It is the method through which a system determines what level of access a certain authorized user should have to the system's secure resources. To establish a solid cryptographic authentication, the authentication technique we used here is an asymmetric key exchange with an ECC authentication algorithm and SHA-256 hashed data within a smart contract. Because the data is hashed using the proposed SHA-256 method before being stored inside a blockchain, this approach ensures data integrity benefits because users can compute the hash data each time, they want to retrieve hashed data.

The proposed model consists of two parts on-chain components and off-chain components to guarantee data availability between users. The components of mentioned On-chain are a smart contract and blockchain. The Interplanetary File System (IPFS) (https://ipfs.io/) is a system that allows IPFS which is a peer-to-peer file-sharing system that authenticates and transports data using cryptographic hash functions.

The primary goal of IPFS is to efficiently store large files. When storing private or secret data on the cloud, data confidentiality is critical. A data confidentiality, authentication, and access control might be solved by improving cloud reliability and trustworthiness so that we propose an on-chain components based blockchain and off-chain components based IPFS. It makes use of a distributed architecture-based file storage system in which each server may save a fraction of the complete data, resulting in a reliable file storage and sharing system. IPFS uses content addresses to name files. Signals, photos, and any other types of data can be saved on an IPFS server in a system context. Figure 3 describes the transfer process between on-chain and off-chain components and presents the communication between user's keys from different users to transmit read/write secured data procedures. An off-chain component is presented to minimize the cost of the model and guarantee the model privacy. To build off-chain storage to our proposal, an IPFS for key management data is presented to make a decentralized system more secured. The off-chain components The off-chain components is commonly used forfor data storage and to ease an increase the communication simplicity. The proposed EEC cryptographic algorithm is used for secure authentication process between sensors, and drones. All user's data are stored in the IPFS data to let the model be integrated and lightweight as the blockchain cannot store a lot of data.

Then smart contract is created and stores all users' cryptographic keys then connected with blockchain as on-chain components stage using Solidity Language for creating the proposed smart contract. Solidity is a high-level object-oriented language for creating smart contracts. Smart contracts are well-known software that able to control of how accounts behave in the Ethereum state.

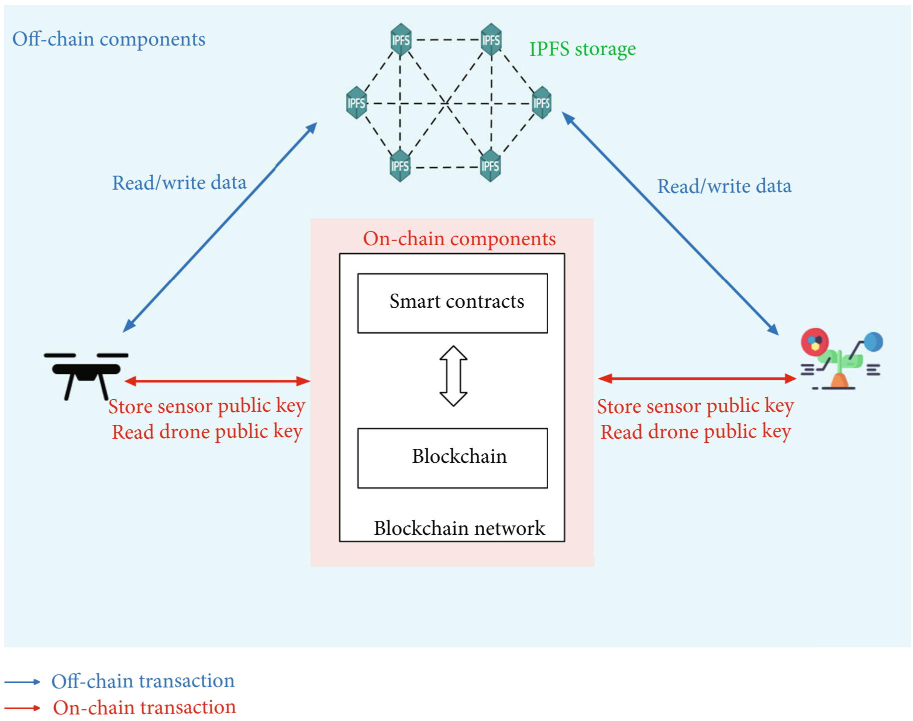


Figure 3: A general diagram of the proposed framework.

Table 2: System's smart contracts main functions.

#	Function	Descriptions
1	*registerNewUser*	This function is responsible for registering a new users in the system
2	*requestData*	This function is responsible for requesting data from a specific user in the system
3	*provideData*	This function is responsible for storing requested data

4.2. Smart Contract. The system smart contract contains three main functions namely *registerNewUser*, *requestData*, and *provideData*. Table 2 provides a description of each function in the smart contract.

4.2.1. The registerNewUser Function. Algorithm 1 describes the process of registering new system users. This function is executed by the user passing the relevant information, including user wallet address, user public key, and role. Based on the user role, smart contracts store user information in the relevant on-chain storage. The smart contract then notifies the system admin, who is responsible for setting up the system, to validate users' information through an off-chain process. After successful validation, the admin executes a specific smart contract function to approve the user registration request.

4.2.2. The registerNewUser Function. Algorithm 2 describes the process of requesting data from a specific user in the system. The function utilizes a mapping data structure for efficient data storing and retrieval. The request information includes the request identification number, sensor wallet address, and drone wallet address. When request information is stored, the smart contract notifies the relevant user to process the request.

```
Input: wallet, publicKey, role
Output: response
User ⟵ $mapping$
if $User[wallet].wallet == null$ then
    User.insert(wallet, [wallet, publicKey, role])
    response: $successful$
else
    response: revert smart contract state
```

Algorithm 1: registerNewUser.

4.2.3. The provideData Function. Algorithm 3 describes the process of providing requested data. To provide data for a specific request, the relevant user, the data owner, prepares the requested data and then executes this function, passing the request identification number and the data. When requested data is available, the smart contract notifies the relevant user to retrieve the data.

5. Results

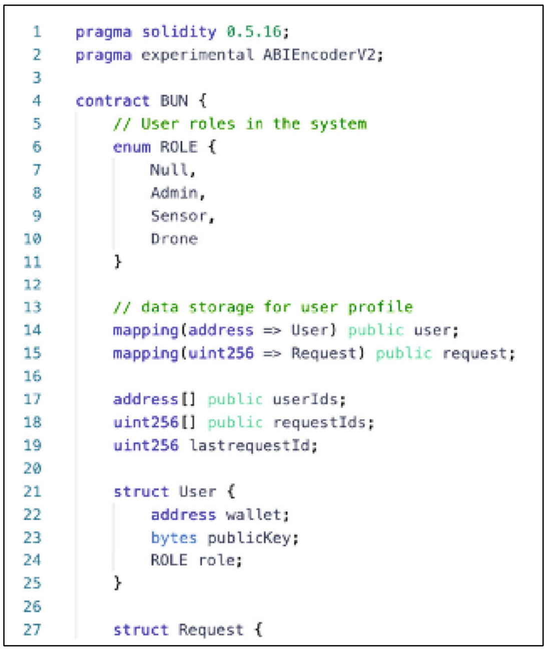
5.1. A Proof of Concept. We implemented a proof of concept to demonstrate the feasibility of the proposed framework. We used Hyperledger Besu (https://www.hyperledger.org/

```
Input: id, semsorWallet, dronerWallet
Output: response
Request ⟵ mapping
if $User[sensorWallet].role == Sensor$ AND $User[dronerWallet].role == droner$ then
    Request.insert($id$,[sensorWallet, dronerWallet])
    response: $successful$
else
  response: revert smart contract state
```

Algorithm 2: : requestData

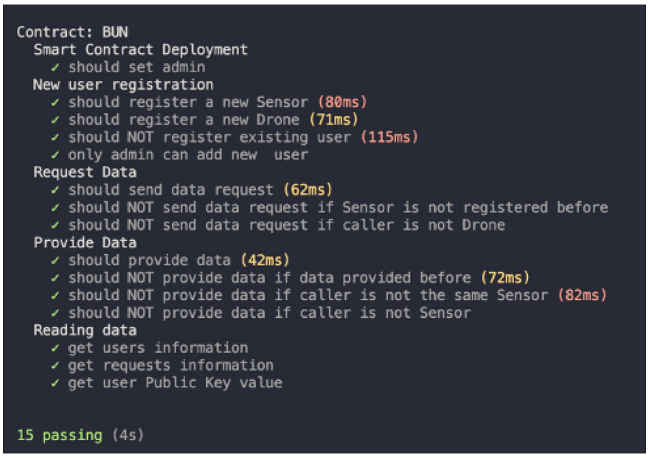
```
Input: id, sensorWallet, encryptedData
Output: response
Request ⟵ mapping
if $User[sensorWallet].role == Sensor$ AND $User[dronerWallet].role == droner$ then
    Request.insert($id$, [encryptedData])
    response: $successful$
else
    response: revert smart contract state
```

Algorithm 3: provideData.



```
pragma solidity 0.5.16;
pragma experimental ABIEncoderV2;

contract BUN {
    // User roles in the system
    enum ROLE {
        Null,
        Admin,
        Sensor,
        Drone
    }

    // data storage for user profile
    mapping(address => User) public user;
    mapping(uint256 => Request) public request;

    address[] public userIds;
    uint256[] public requestIds;
    uint256 lastrequestId;

    struct User {
        address wallet;
        bytes publicKey;
        ROLE role;
    }

    struct Request {
```

Figure 4: A screenshot of the system smart contract.



```
Contract: BUN
  Smart Contract Deployment
    ✓ should set admin
  New user registration
    ✓ should register a new Sensor (80ms)
    ✓ should register a new Drone (71ms)
    ✓ should NOT register existing user (115ms)
    ✓ only admin can add new  user
  Request Data
    ✓ should send data request (62ms)
    ✓ should NOT send data request if Sensor is not registered before
    ✓ should NOT send data request if caller is not Drone
  Provide Data
    ✓ should provide data (42ms)
    ✓ should NOT provide data if data provided before (72ms)
    ✓ should NOT provide data if caller is not the same Sensor (82ms)
    ✓ should NOT provide data if caller is not Sensor
  Reading data
    ✓ get users information
    ✓ get requests information
    ✓ get user Public Key value


15 passing (4s)
```

Figure 5: The result of system smart contract testing.

use/besu) to build a permissioned blockchain. The system smart contract was written using the Solidity programming language, where the Truffle framework (https://www.trufflesuite.com/truffle), an Ethereum smart contracts development tool, was used to test, compile, and deploy the system smart contract. In Figure 4, a screenshot of the system smart contract shows a screenshot of the system smart contract, whereas in Figure 5, the result of system smart contract testing shows the result of smart contract testing undertaken. Lastly, we utilized Node.js and IPFS to develop the off-chain components. The source code is available on Mendeley Data [49] under the CC BY 4.0 license.

The high-level structure of the implemented proof of concept is shown in Figures 6–8. All users in the network save their public keys in the smart contract and the read/write procedures for different user's keys from drones and sensors as explained in Figure 4, for example, a user 1 who wants to send data to user 2 considering the next steps:

(i) User 1, send data to user 2

  (a) The data of the user 1 form is hashed using a hash function, which is passed to the hash value and then stored by smart contract in the blockchain

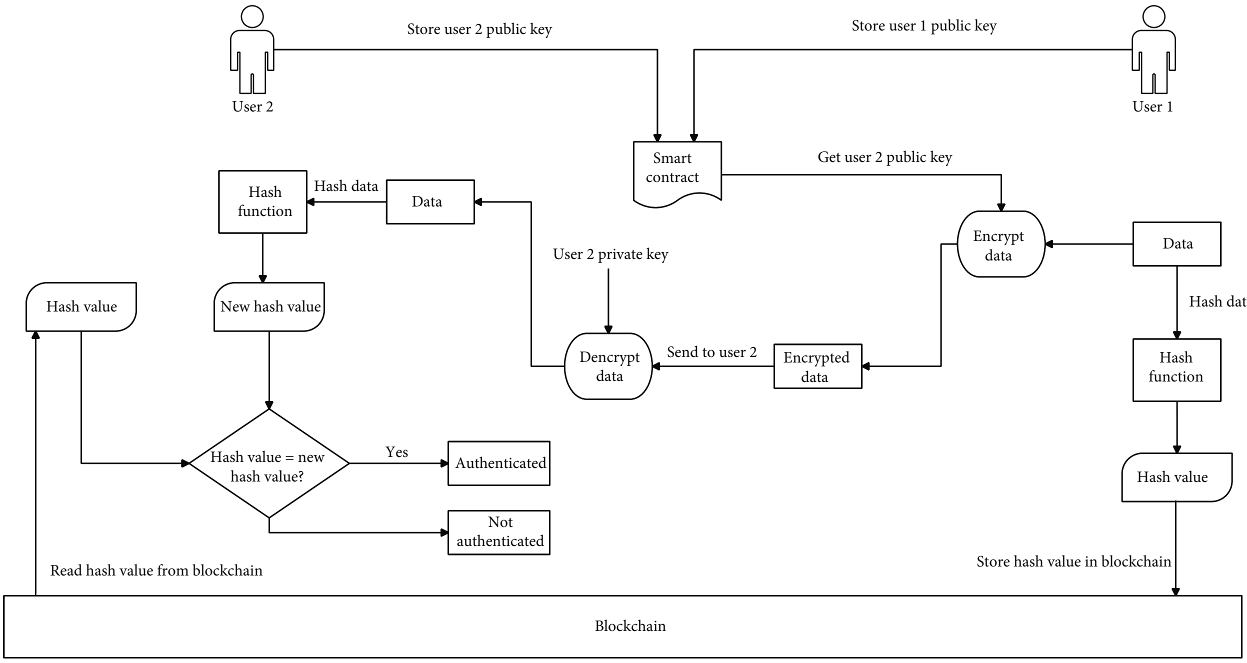


Figure 6: The high-level structure of the proof of concept.

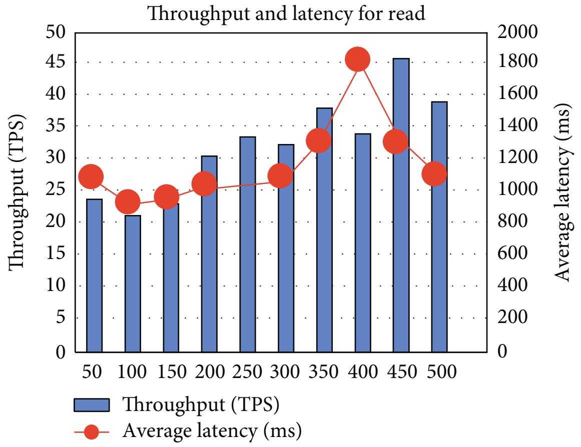


Figure 7: The *throughput* and *latency* results for Read operations.

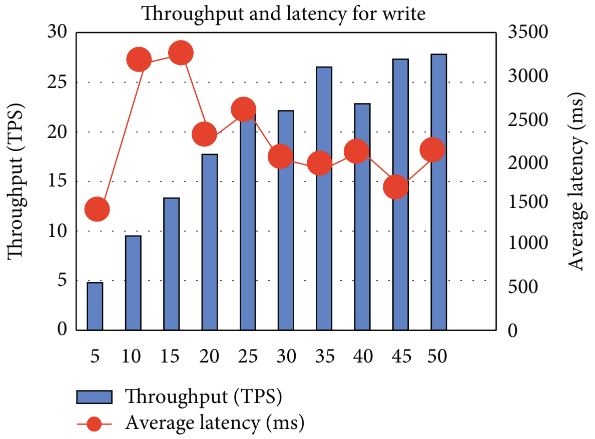


Figure 8: The *throughput* and *latency* results for Write operations.

(ii) User 2 receives data

(a) The data is encrypted using user 2's public key, which is read from user 2's smart contract, and this data is then sent to user 2

(b) With user 2's private key, the data is decrypted, and that data is then hashed into a new hash value

(c) Decision-making occurs when a match is made between two hash values: the hash value is saved in the blockchain and the new hash value. If the hash is matched, authentication occurs, and data is sent

5.2. Performance Evaluation. To evaluate the performance of proposed framework, we utilized an open-source benchmarking tool called Hyperledger Caliper (https://www.hyperledger.org/use/caliper). Table 3 shows the settings of the performance evaluation environment. Two main types of blockchain operations, *Write* and *Read*, were evaluated using four performance indicators, namely, *Write Throughput*, *Read Throughput*, *Write Latency*, and *Read Latency* [50, 51]. In this performance evaluation, we focused on the main smart contract functions that are shown in Table 4.

The performance evaluation was performed in ten rounds with a hundred transactions per round to reduce the likelihood of errors due to network congestion and system overload. Tables 5 and 6 summarize the performance evaluation settings used for the Write and Read operations, respectively.

Table 7 demonstrates the results of the *throughput* and *latency* for Read operations and the *throughput* and *latency*

Table 3: The settings of the performance evaluation environment.

Factor	Setting
Nodes	Four VMs running on Google cloud, where each VM has a 2 GHz 4-core Intel CPU
Peer-to-peer network	Hyperledger Besu v1.4.1, 1 validator node, 3 peer nodes
Consensus protocol	Clique
Smart contract programming language used	Solidity
Benchmarking tool	Hyperledger Caliper v0.4.1

Table 4: Main smart contracts functions of the system.

Main function	Operation type
registerNewUser	*Write*
requestData	*Write*
provideData	*Write*
getUserPublicKey	*Read*
getUsers	*Read*
getRequests	*Read*

Table 5: The performance evaluation settings used for the Write operations.

Test number	1	2	3	4	5	6	7	8	9	10
Functions under test	Write operations									
Worker number	1 worker									
Transaction number	100 transactions									
Type of control rate	Fixed rate									
Send rate (tps)	5	10	15	20	25	30	35	40	45	50

for Write operations, respectively. The results indicate an average *throughput* of 32.54 TPS and an average latency 1166 milliseconds for Read operations. Contrastingly, average Write *throughput* is 19.37 TPS, and the average Write *latency* is 2253 milliseconds. The experimental findings for the *Read* and *Write* operations are shown in Tables 7 and 8.

5.3. Discussion. Blockchain technology has developed as a way to make distributed systems more secure so that we provide security to the network users confidential data over the cloud. Blockchains are digital ledgers that hold explicit and verifiable records of all transactions inside a system. The decentralized blockchain concept has shown to be a reliable technique for resolving trust difficulties in user authentication. As a result of this fact, the specifics of each transaction could be saved in order to ensure that the data transmission is secure. The major objective of this work is to guarantee that the system's authentication is robust and safe against assaults, since each user has their own private and public keys, which were previously issued to them by the system and stored on the smart contract.

To establish a solid cryptographic authentication, the technique uses an asymmetric key exchange with an ECC authentication algorithm and SHA-256 hashed data within a smart contract. The model is built on a private blockchain-based platform that ensures safe connection and secure data transmission through the cloud between sensors and drones in smart farming. This work supports our system by ensuring data integrity since data is hashed before being recorded in a blockchain, and users calculate the hash data each time they want to access hashed data.

The proposed framework utilizes permissioned blockchain where only authorized users can access the system. System user interacts with the system using their wallet accounts which are pseudo-anonymous accounts; therefore, users' privacy is preserved. In addition, multiple pseudo-anonymous accounts can be used by a single user; hence, user transactions cannot be tracked by an adversary. The use of on-chain/off-chain storage in the framework increases data confidentiality and integrity as sensitive data are stored securely off-chain and only the hash value of the data is submitted to the blockchain.

The major goal of this work is to guarantee that the system's authentication is stable and safe against assaults, since each user has their own private and public keys, which were previously issued by the system and stored on the smart contract, before registering and entering the system. To establish a solid cryptographic authentication, the work is based on an asymmetric key exchange within a smart contract utilizing an ECC authentication algorithm and SHA-256 hashed data. The system is built on a private blockchain-based platform that ensures safe connection and secure data transmission through the cloud between sensors and drones in smart farming. Because the data is hashed before being stored within a blockchain, this method protects data integrity, and users may calculate the hash data each time they want to access hashed data.

6. Conclusions

We proposed an asymmetric key cryptography blockchain as an on-chain component for this study, which requires storing data on permission blockchain as the most cost-effective way to keep data decentralized and guarantee model availability. We looked at smart contracts on the blockchain that can be utilized in the realm of the Internet of Things, as well as the benefits and challenges that they bring. We used Ethereum smart contracts, a decentralized and encrypted technology that allows devices to better trust one another and execute peer-to-peer authentication.

The hashed data will simply be transmitted on the on-chain component. No one will be able to access the model's private keys after they are set, which validates the model's privacy.

We used an Ethereum blockchain to test the performance of the proposed approach. Four virtual computers

Table 6: The performance evaluation settings used for the Read operations.

Test number	1	2	3	4	5	6	7	8	9	10
Functions under test	Read operations									
Worker number	1 worker									
Transaction number	100 transactions									
Type of control rate	Fixed rate									
Send rate (tps)	50	100	150	200	250	300	350	400	450	500

Table 7: Experimental findings of *throughput* and *latency* for Read operations.

Test round	Send rate (tps)	Average latency (ms)	Throughput (TPS)
1	50	1050	23.9
2	100	920	21.3
3	150	950	23.2
4	200	1020	30.4
5	250	1040	33.7
6	300	1060	32.6
7	350	1350	40.8
8	400	1800	34.4
9	450	1330	45.7
10	500	1140	39.4

Table 8: Experimental findings of *throughput* and *latency* for Write operations.

Test round	Send rate (tps)	Average latency (ms)	Throughput (WPS)
1	5	1420	4.8
2	10	3180	9.5
3	15	3260	13.3
4	20	2240	17.7
5	25	2500	21.9
6	30	2040	22.1
7	35	1960	26.5
8	40	2100	22.8
9	45	1710	27.3
10	50	2120	27.8

from the Google cloud are used to guarantee the newly added device's security needs while also achieving benefits such as reduced traffic overheads and typifying our solution's high level of intelligence and mobility. We believe that the blockchain solution is a step toward greater data security and privacy and that it has the potential to be employed in a wide range of IoT applications.

This study has several limitations. In the proposed framework, verifying user identity is challenging due to the distributed and the openness nature of blockchain technology. As the proposed framework operates on a permissioned blockchain, the process of verifying user identity is performed by the system owner, who is responsible for setting up the blockchain and inviting users to join the system. This can be mitigated by integrating the system with identity management services such as self-sovereign identity [52–54], identity verification using blockchain [55], and noncustodial login solutions using blockchain [56]. The blockchain's General Data Protection Regulation (GDPR) compliance is another limitation in this study [57–59]. Although utilizing permissioned blockchains might comply with GDPR requirements, determining whether blockchain completely complies with GDPR is challenging [60]. To avoid such limitation, GDPR compliance should be considered during designing the blockchain-based system [61, 62].

Throughout this study, we described an approach on how to achieve trust among smart farming users. We now highlight some future research directions. Firstly, we will improve our work by incorporating an off-chain (IPFS)-based decentralized distributed data storage method to allow for speedy, low-cost, and reliable data access, hence, approving the model's availability. Accessing data across users and networks in a faster, more secure, and network-effective manner is another advantage of using off-chain components. Secondly, a comparative analysis of the proofs of concept implemented with different blockchain frameworks and configurations will be conducted. In this study, the proofs of concept were implemented using the Ethereum blockchain, which was initially designed for developing DApps and services that are open to the public. Ethereum blockchain has several limitations in terms of performance and scalability, such as transaction latency, throughput, and execution time [63]. In future works, an empirical evaluation should be conducted to assess the performance of the proofs-of-concept design under a wide range of blockchain frameworks and configurations other than Ethereum such as Hyperledger Fabric and MultiChain. Finally, more research is required on the framework applicability to explore and assess the sustainability challenges faced by our proposed framework, including its limitations for real-world utilizing.

Data Availability

There are no relevant data to be made available.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] R. Lal, “Feeding 11 billion on 0.5 billion hectare of area under cereal crops,” *Food and Energy Security*, vol. 5, no. 4, pp. 239–251, 2016.

[2] M. Z. Mehmood, M. Ahmed, O. Afzal et al., “Internet of Things (IoT) and sensors technologies in smart agriculture: applications, opportunities, and current trends,” in *Building Climate Resilience in Agriculture*, pp. 339–364, Springer, 2022.

[3] A. Rehman, T. Saba, M. Kashif, S. M. Fati, S. A. Bahaj, and H. Chaudhry, “A revisit of internet of things technologies for monitoring and control strategies in smart agriculture,” *Agronomy*, vol. 12, no. 1, p. 127, 2022.

[4] F. A. Almalki, B. O. Soufiene, S. H. Alsamhi, and H. Sakli, “A low-cost platform for environmental smart farming monitoring system based on IoT and UAVs,” *Sustainability*, vol. 13, no. 11, p. 5908, 2021.

[5] F. A. Almalki and B. O. Soufiene, “Modifying Hata-Davidson propagation model for remote sensing in complex environments using a multifactional drone,” *Sensors*, vol. 22, no. 5, p. 1786, 2022.

[6] F. A. Almalki and M. C. Angelides, “Autonomous flying IoT: a synergy of machine learning, digital elevation, and 3D structure change detection,” *Computer Communications*, vol. 190, pp. 154–165, 2022.

[7] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, “Internet of Things (IoT): a vision, architectural elements, and future directions,” *Future Generation Computer Systems*, vol. 29, no. 7, pp. 1645–1660, 2013.

[8] P. Gope and T. Hwang, “BSN-Care: a secure IoT-based modern healthcare system using body sensor network,” *IEEE Sensors Journal*, vol. 16, no. 5, pp. 1368–1376, 2015.

[9] H. Arasteh, V. Hosseinnezhad, V. Loia et al., “Iot-based smart cities: a survey,” in *2016 IEEE 16th international conference on environment and electrical engineering (EEEIC)*, pp. 1–6, Florence, Italy, 2016.

[10] A. Khanna and S. Kaur, “Evolution of Internet of Things (IoT) and its significant impact in the field of precision agriculture,” *Computers and Electronics in Agriculture*, vol. 157, pp. 218–231, 2019.

[11] W. Ejaz, M. A. Azam, S. Saadat, F. Iqbal, and A. Hanan, “Unmanned aerial vehicles enabled IoT platform for disaster management,” *Energies*, vol. 12, no. 14, p. 2706, 2019.

[12] A. D. Boursianis, M. S. Papadopoulou, P. Diamantoulakis et al., “Internet of things (IoT) and agricultural unmanned aerial vehicles (UAVs) in smart farming: a comprehensive review,” *Internet of Things*, vol. 18, p. 100187, 2022.

[13] F. A. Almalki, M. Aljohani, M. Algethami, and B. O. Soufiene, “Incorporating drone and AI to empower smart journalism via optimizing a propagation model,” *Sustainability*, vol. 14, no. 7, p. 3758, 2022.

[14] F. Albalwy, A. Brass, and A. Davies, “A blockchain-based dynamic consent architecture to support clinical genomic data sharing (ConsentChain): proof-of-concept study,” *JMIR medical informatics*, vol. 9, no. 11, article e27816, 2021.

[15] F. Albalwy, J. H. McDermott, W. G. Newman, A. Brass, and A. Davies, “A blockchain-based framework to support pharmacogenetic data sharing,” *The Pharmacogenomics Journal*, 2022.

[16] U. Chelladurai and S. Pandian, “A novel blockchain based electronic health record automation system for healthcare,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, no. 1, pp. 693–703, 2022.

[17] L. Hang, B. Kim, K. Kim, and D. Kim, “A permissioned blockchain-based clinical trial service platform to improve trial data transparency,” *BioMed Research International*, vol. 2021, Article ID 5554487, 22 pages, 2021.

[18] V. S. Anoop and J. Goldston, “Decentralized finance to hybrid finance through blockchain: a case-study of acala and current,” *Journal of Banking and Financial Technology*, vol. 6, no. 1, pp. 109–115, 2022.

[19] D. Younus, A. Muayad, and M. Abumandil, “Role of smart contract technology blockchain services in finance and banking systems: concept and core values,” *Mohanad, Role of Smart Contract Technology Blockchain Services in Finance and Banking Systems: Concept and Core Values (April 8, 2022)*, 2022.

[20] K. Azari and S. Malek, *Blockchain Applications in Real Estate: Challenges and a Proposed Framework*, 2022.

[21] A. Patil, A. Shinde, A. Panigrahi, A. Arora, D. S. Raviraja, and R. Babu, “The role of blockchain technology in decentralized real estate marketplace: recent findings,” 2022.

[22] F. Jamil, M. Ibrahim, I. Ullah, S. Kim, H. K. Kahng, and D.-H. Kim, “Optimal smart contract for autonomous greenhouse environment based on IoT blockchain network in agriculture,” *Computers and Electronics in Agriculture*, vol. 192, p. 106573, 2022.

[23] L. Hang, I. Ullah, and D.-H. Kim, “A secure fish farm platform based on blockchain for agriculture data integrity,” *Computers and Electronics in Agriculture*, vol. 170, p. 105251, 2020.

[24] M. K. Dash, G. Panda, A. Kumar, and S. Luthra, “Applications of blockchain in government education sector: a comprehensive review and future research potentials,” *Journal of Global Operations and Strategic Sourcing*, vol. 15, no. 3, pp. 449–472, 2022.

[25] A. Garg, P. Kumar, M. Madhukar, O. Loyola-González, and M. Kumar, “Blockchain-based online education content ranking,” *Education and Information Technologies*, vol. 27, no. 4, pp. 4793–4815, 2022.

[26] H. F. Atlam, A. Alenezi, M. O. Alassafi, and G. Wills, “Blockchain with internet of things: benefits, challenges, and future directions,” *International Journal of Intelligent Systems and Applications*, vol. 10, no. 6, pp. 40–48, 2018.

[27] A. Yazdinejad, R. M. Parizi, A. Dehghantanha, H. Karimipour, G. Srivastava, and M. Aledhari, “Enabling drones in the internet of things with decentralized blockchain-based security,” *IEEE Internet of Things Journal*, vol. 8, no. 8, pp. 6406–6415, 2020.

[28] B. Bera, D. Chattaraj, and A. K. Das, “Designing secure blockchain-based access control scheme in IoT-enabled Internet of Drones deployment,” *Computer Communications*, vol. 153, pp. 229–249, 2020.

[29] A. Khalifeh, K. A. Darabkh, A. M. Khasawneh et al., “Wireless sensor networks for smart cities: network design, implementation and performance evaluation,” *Electronics*, vol. 10, no. 2, p. 218, 2021.

[30] C. Feng, K. Yu, A. K. Bashir et al., “Efficient and secure data sharing for 5G flying drones: a blockchain-enabled approach,” *IEEE Network*, vol. 35, no. 1, pp. 130–137, 2021.

[31] B. Bera, A. K. Das, and A. K. Sutrala, “Private blockchain-based access control mechanism for unauthorized UAV detection and mitigation in Internet of Drones environment,” *Computer Communications*, vol. 166, pp. 91–109, 2021.

[32] E. Ghribi, T. T. Khoei, H. T. Gorji, P. Ranganathan, and N. Kaabouch, "A secure blockchain-based communication approach for UAV networks," in *2020 IEEE International Conference on Electro Information Technology (EIT)*, pp. 411–415, Chicago, IL, USA, 2020.

[33] R. L. Kumar, Q.-V. Pham, F. Khan, M. J. Piran, and K. Dev, "Blockchain for securing aerial communications: potentials, solutions, and research directions," *Physical Communication*, vol. 47, p. 101390, 2021.

[34] G. Cho, J. Cho, S. Hyun, and H. Kim, "SENTINEL: a secure and efficient authentication framework for unmanned aerial vehicles," *Applied Sciences*, vol. 10, no. 9, p. 3149, 2020.

[35] A. Chen, K. Peng, Z. Sha, X. Zhou, Z. Yang, and G. Lu, "ToAM: a task-oriented authentication model for UAVs based on blockchain," *EURASIP Journal on Wireless Communications and Networking*, vol. 2021, Article ID 166, 15 pages, 2021.

[36] X. Li, Y. Wang, P. Vijayakumar, D. He, N. Kumar, and J. Ma, "Blockchain-based mutual-healing group key distribution scheme in unmanned aerial vehicles ad-hoc network," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 11, pp. 11309–11322, 2019.

[37] B. Bera, A. Vangala, A. K. Das, P. Lorenz, and M. K. Khan, "Private blockchain-envisioned drones-assisted authentication scheme in IoT- enabled agricultural environment," *Computer Standards & Interfaces*, vol. 80, p. 103567, 2022.

[38] D. Maldonado-Ruiz, J. Torres, and N. El Madhoun, "3BI-ECC: a decentralized identity framework based on blockchain technology and elliptic curve cryptography," in *2020 2nd Conference on Blockchain Research & Applications for Innovative Networks and Services (BRAINS)*, pp. 45-46, Paris, France, 2020.

[39] H. Wang, D. He, and Y. Ji, "Designated-verifier proof of assets for bitcoin exchange using elliptic curve cryptography," *Future Generation Computer Systems*, vol. 107, pp. 854–862, 2020.

[40] A. K. Yadav, "Significance of elliptic curve cryptography in blockchain IoT with comparative analysis of RSA algorithm," in *2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, pp. 256–262, Greater Noida, India, 2021.

[41] R. Kumar and R. Tripathi, "Secure healthcare framework using blockchain and public key cryptography," in *Blockchain Cybersecurity, Trust and Privacy*, pp. 185–202, Springer, 2020.

[42] Z. E. Rasjid, B. Soewito, G. Witjaksono, and E. Abdurachman, "A review of collisions in cryptographic hash function used in digital forensic tools," *Procedia computer science*, vol. 116, pp. 381–392, 2017.

[43] S. Soni and S. P. Singh, "Secure and efficient integrity algorithm based on existing SHA algorithms," *International Journal of Computer Applications*, vol. 113, no. 11, pp. 34–37, 2015.

[44] NIST Policy on Hash FunctionsJanuary 2022, https://csrc.nist.gov/Projects/Hash-Functions/NIST-Policy-on-Hash-Functions.

[45] W. L. Harrison, A. M. Procter, and G. Allwein, "Model-driven design & synthesis of the SHA-256 cryptographic hash function in rewire," in *2016 International Symposium on Rapid System Prototyping (RSP)*, pp. 1–7, Pittsburgh, PA, USA, 2016.

[46] M. Qazi, D. Kulkarni, and M. Nagori, "Proof of authenticity-based electronic medical records storage on blockchain," in *Smart Trends in Computing and Communications*, pp. 297–306, Springer, 2020.

[47] K. Quist-Aphetsi and H. Blankson, "A hybrid data logging system using cryptographic hash blocks based on SHA-256 and MD5 for water treatment plant and distribution line," in *2019 International Conference on Cyber Security and Internet of Things (ICSIoT)*, pp. 15–18, Accra, Ghana, 2019.

[48] A. Ali, M. F. Pasha, J. Ali et al., "Deep learning based homomorphic secure search-able encryption for keyword search in blockchain healthcare system: a novel approach to cryptography," *Sensors*, vol. 22, no. 2, p. 528, 2022.

[49] F. Albalwy, *blockchain-for-uav-networks*, Mendeley Data, 2022.

[50] Z. Lu, Q. Wang, G. Qu, and Z. Liu, "BARS: a blockchain-based anonymous reputation system for trust management in VANETs," in *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*, pp. 98–103, New York, NY, USA, 2018.

[51] Hyperledger, *Hyperledger Blockchain Performance Metrics White Paper*, Hyperledger, 2022, January 2022, https://www.hyperledger.org/learn/publications/blockchain-performance-metrics.

[52] N. Naik and P. Jenkins, "Sovrin Network for decentralized digital identity: analysing a self-sovereign identity system based on distributed ledger technology," in *2021 IEEE International Symposium on Systems Engineering (ISSE)*, pp. 1–7, Vienna, Austria, 2021.

[53] T. Rathee and P. Singh, "A self-sovereign identity management system using blockchain," in *Cyber Security and Digital Forensics*, pp. 371–379, Springer, 2022.

[54] M. Shuaib, N. H. Hassan, S. Usman et al., "Self-sovereign identity solution for blockchain-based land registry system: a comparison," *Mobile Information Systems*, vol. 2022, Article ID 8930472, 17 pages, 2022.

[55] blockpassAugust 2022, https://www.blockpass.org/.

[56] RemmeAugust 2022, https://remme.io/.

[57] M. Berberich and M. Steiner, "Practitioner's Corner Blockchain technology and the GDPR – how to reconcile privacy and distributed ledgers?," *European Data Protection Law Review*, vol. 2, no. 3, pp. 422–426, 2016.

[58] A. V. Humbeeck, "The blockchain-GDPR paradox," *Journal of Data Protection & Privacy*, vol. 2, no. 3, pp. 208–212, 2019.

[59] C. Compert, M. Luinetti, and B. Portier, *Blockchain and GDPR: How Blockchain Could Address Five Areas Associated with GDPR Compliance*, IBM Security, 2018, August 2022, https://iapp.org/media/pdf/resource_center/blockchain_and_gdpr.pdf.

[60] M. Finck, *Blockchain and the General Data Protection Regulation: Can Distributed Ledgers Be Squared with European Data Protection Law?(Study No: PE 634.445)*, European Parliament, Brussels, 2019.

[61] N. Eichler, S. Jongerius, G. McMullen, O. Naegele, L. Steininger, and K. Wagner, *Blockchain, Data Protection, and the GDPR*, Blockchain Bundesverband, 2018, August 2022, https://www.crowdfundinsider.com/wp-content/uploads/2018/06/GDPR_Position_Paper_v1.0.pdf.

[62] A. Rose, "GDPR challenges for blockchain technology," *Interactive Entertainment Law Review*, vol. 2, no. 1, pp. 35–41, 2019.

[63] L. Hang, B. Kim, and D. Kim, "A transaction traffic control approach based on fuzzy logic to improve hyperledger fabric performance," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 2032165, 19 pages, 2022.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 2545958, 9 pages
https://doi.org/10.1155/2022/2545958

Research Article

Transfer Learning-Based Vehicle Collision Prediction

Li Yang,[1] **Zonggao Wang**,[1] **Lijun Ma**,[2] **and Wei Dai**[3]

[1]*School of Information and Safety Engineering, Zhongnan University of Economics and Law, Wuhan 430073, China*
[2]*School of Statistics Mathematics, Zhongnan University of Economics and Law, Wuhan 430073, China*
[3]*Institute of Government Accounting, Zhongnan University of Economics and Law, Wuhan 430073, China*

Correspondence should be addressed to Wei Dai; david@zuel.edu.cn

Received 8 April 2022; Revised 1 August 2022; Accepted 6 August 2022; Published 28 August 2022

Academic Editor: Yin Zhang

Traffic accident is an important problem in modern society. Vehicle collision prediction is one of the key technical points that must be broken through in the future driving system. However, due to the complexity of traffic environment and the difference of emergency ability of drivers, it is very difficult to predict vehicle collision. Although experts and scholars have tried to monitor and predict accidents in real time according to environmental conditions, overly agile warning or inaccurate prediction may cause serious consequences. Therefore, in order to more accurately predict the occurrence of vehicle collision, this paper analyses and models the driving mode of the vehicle based on transfer learning and using the previous performance data of the vehicle, so as to predict the future collision situation and even the collision time of the vehicle. Finally, using a real-world Internet of Vehicles data set, this paper implements a large number of experiments to verify the effectiveness of the proposed model.

1. Introduction

With the development of society, the number of vehicles is gradually increasing, and the problem of traffic safety has attracted more and more people's attention. The frequent occurrence of traffic accidents is worrying. More than 10 million people worldwide are injured in road accidents every year. Among these accidents, vehicle collision is a serious safety problem, accounting for almost 30% of all accidents [1].

However, many accidents are closely related to the improper operation of drivers and the lack of timely and effective response to emergencies. In fact, with the development of technology, especially the development of Internet of Vehicles, automatic driving, and other technologies, the state of various parts of the vehicle can be tracked completely. It is very possible to predict whether the vehicle will collide or even predict the specific collision time in the future based on these data. However, since the implementation and deployment of intelligent transportation system and Internet of Vehicles are still in the initial stage, these data are still very difficult to obtain. At present, aiming at the problem of vehicle collision prediction, some researchers have proposed to monitor the vehicle environment in real time through the radar set on the road, the vehicle's own infrared, camera, and other sensing equipment, so as to use these data to predict the vehicle collision and give timely warnings to the drivers on the vehicle [2–4]. However, these data based on the external environment, such as radar signals and images that are vulnerable to weather, have strong uncertainty. Therefore, some researchers turn their attention to the relatively stable interior of the vehicle, such as assessing the possibility of collision by paying attention to the driver's behaviour [5, 6]. However, vehicle interior data is usually difficult to be effectively dynamically modelled because the driving habits of drivers are very personalized.

Recently, the rapid development of deep learning has brought great technological changes to various fields. Among them, the transfer learning technology, which can make full use of the previously collected data, makes the current model perform better on less data and has been favoured by more and more people [7–9]. Inspired by these works, this paper intends to use transfer learning to realize the complex task of vehicle collision prediction. In fact, compared with complex external data, real-time vehicle operation data from the interior of the vehicle, such as vehicle speed, accelerator pedal position, and brake pedal state, are less vulnerable to environmental interference and are directly related to the driving state of the vehicle. Transfer learning can discover more features related to vehicle collisions from limited vehicle operation data by means of knowledge transfer. In order to explore a more safe and effective vehicle collision prediction method, this paper uses the above data to carry out the research of vehicle collision prediction task. However, these data are still very difficult to obtain. In order to fully mine the correlation law between vehicle running state and vehicle collision from these limited data, this paper uses the efficient modelling method of transfer learning to build a model that can not only accurately predict whether the vehicle has a collision but also clearly point out the possible collision time. Specifically, the contributions of this paper are as follows:

(1) This paper proposes a vehicle collision prediction model based on transfer learning, which is called TLVC. This method explores the new use of Internet of Vehicles data and provides a strong technical support for safe driving in the future

(2) In this paper, a special feature analysis method is developed for the operation data from the inside of the vehicle, which provides a reference for the dynamic behaviour modelling of drivers. Moreover, this method of using vehicle internal data is more reliable than the previous methods based on image and radar signal

(3) Using a small amount of Internet of Vehicles data and EfficientNet, this paper constructs a transfer learning model which is more accurate and clearer than the previous vehicle collision prediction model. Finally, through a large number of experiments, we confirm the effectiveness and accuracy of the proposed model

The remaining chapters of this paper are arranged as follows: the second section introduces the research related to vehicle collision and transfer learning. The third section introduces the vehicle collision prediction model proposed in this paper. The fourth section will introduce the real data set of this paper. The last section will summarize the full text and discuss our future work.

2. Related Work

In this part, this paper will introduce the previous work of vehicle collision prediction in detail and review the previous research on transfer learning.

2.1. Vehicle Collision Prediction. In recent years, many researchers have done research on vehicle avoidance and vehicle collision prediction. For example, Wang et al. [3] based on convolutional neural network and using the collected real trajectory data proposed a set of methods from data collection to preprocessing and then to prediction but did not deeply explore how to use the collected data to make more accurate prediction. Lyu et al. [4] established the lane change intention recognition model by tracking the driving direction of the vehicle and then established the collision early warning model by comprehensively predicting the vehicle trajectory. Candela et al. [10] combined with road layout information, statistical agent dynamics, and discrete Gaussian process for future vehicle position estimation, so as to realize vehicle collision prediction. Peng et al. [5] constructed a comprehensive “driver-vehicle-road” data set for actual driver behaviour evaluation, mainly analysing driver behaviour and relevant factors that significantly affect driving safety in emergency situations. Lee et al. [11] constructed a dynamic riding simulator that can control rolling motion, quantified driving behaviour by using lateral control ability, driver’s head movement, and emotional state, and predicted the overall collision avoidance ability by using multiple regression analysis of driving behaviour. Zhang et al. [12] proposed a multipedestrian collision risk assessment framework according to the motion characteristics of vehicles, including motion prediction module, collision inspection module, and collision risk assessment module. According to Katrakazas et al. [13], under the joint framework of interactive perception motion model and dynamic Bayesian network (DBN), network level collision estimation and vehicle-based risk estimation are integrated in real time, and machine learning classifier is used for real-time network level collision prediction. Wang et al. [14] proposed a collision prediction method based on the bivariate extreme value theory framework, taking into account the driver’s perceived response failure to take appropriate avoidance actions.

To sum up, the current vehicle collision prediction is mainly based on road information, vehicle external motion characteristics, and vehicle trajectory, combined with various roadside sensing units, etc., but there is a lack of attention to the characteristics of the vehicle itself and the driver’s operation state, and the vehicle collision problem is mainly modelled as a classification problem, and there is a lack of prediction of the collision time.

2.2. Transfer Learning. Migration learning improves the performance of the model in the target domain by migrating the knowledge contained in other source domains, which can greatly reduce the dependence of the model on the data of the target domain. Due to its wide application prospects, migration learning has attracted extensive attention recently [15]. For example, Ruder et al. outlined modern transfer learning methods in natural language processing (NLP), how models are pretrained, and what information is captured in their learning representation and reviewed examples and case studies on how these models are integrated and adjusted in downstream NLP tasks [16]. Raghu et al. discussed the characteristics of transfer learning for medical imaging [17]. Through a series of analysis of migrating to block shuffled images, Neyshabur et al. separated the effect of feature reuse from the high-level statistical information of learning data and showed that some benefits of migrating learning came from the latter [18]. Pathak et al. used deep transfer learning technology to classify patients infected with COVID-19 [19]. Wang et al. proposed dynamic distributed adaptation (DDA), which can quantitatively evaluate the relative importance of each distribution to solve the problem of transfer learning [20]. Tammina et al. used one of the pretraining models VGG-16 and deep convolution neural network to classify images [21]. Chen et al. proposed a joint transfer learning framework for wearable healthcare [22]. Rao et al. learned a model that can evaluate protein embedding tasks by migrating five biologically related semisupervised learning tasks [23]. Lotfollahi et al. used transfer learning and parameter optimization to achieve efficient, decentralized, and iterative reference construction and the upper and lower culture of new data sets and existing references [7]. Zhang et al. proposed a transfer learning (TL) technology through domain adaptation to bridge the gap between biased numerical model and real structure, guide Bayesian model update (BMU), and supervise structural damage identification [9].

To sum up, at present, due to some outstanding advantages, transfer learning has been widely concerned in various fields such as medicine and biology, but few people are involved in the emerging field of vehicle network. The work of this paper is to fill the vacancy of this data modelling method in the field of vehicle networking, explore the use of migration learning to model the real-time driving data of vehicles, analyse the vehicle state, learn the vehicle collision prediction model, and finally improve the safety of the driving system.

3. Method

This part will focus on the vehicle collision prediction model based on transfer learning proposed in this paper. The overall structure of the model is shown in Figure 1.

3.1. Problem Definition. Before introducing the model, we first give the definition of the vehicle collision prediction problem solved in this paper.

3.1.1. Definition Vehicle Collision Prediction. Given n vehicle sets $\{c_1, c_2, \cdots c_\mathrm{n}\}$, according to the operation status S_i, $S_i \in R^{h*d}$(h represents the number of historical running state data of the acquired vehicle c_i, and d represents the dimension of monitoring data) of each vehicle c_i obtained from the monitoring system inside the vehicle. The problem of vehicle collision prediction in this paper finally comes down to the training prediction model $\varnothing(\cdot)$, so that it can predict the collision of a given vehicle c_i in the future according to the operation state data S_j of the given vehicle c_i, as follow:

$$\mathrm{T} = \varnothing(\mathrm{S}), \tag{1}$$

where T represents the final prediction result, $T \in [0, 1]$. $T = 0$ means that the modified vehicle will not collide in the future, and $T \in (0, 1]$ represents the specific time when the collision occurred. Figure 2 shows an example of vehicle collision prediction. In this example, the vehicle running state data in the Internet of Vehicles is first used to analyze the vehicle operation law, and the proposed TLVC model is used to simulate the law, and then, it is used to predict whether there will be a collision in the future and the time of the collision.

For the real label G (physical time) of training data, we process it as follows:

$$Y = \begin{cases} 0 & \text{No collision,} \\ \dfrac{86400}{(G - G_{\text{start}}) + 86400} & \text{otherwise.} \end{cases} \tag{2}$$

Y is the training label finally sent into the model; formula (2) indicates the time point when the current vehicle starts monitoring. The above formula makes $\mathrm{Y} \in [0, 1]$. The specific details of the prediction model $\varnothing(\cdot)$ will be described in detail below.

3.2. Preprocessing of Vehicle Operation Status Data. This part corresponds to the left half of Figure 1. The vehicle running data processed in this part include "accelerator pedal position," "collect time," "battery pack main negative relay status," "battery pack main positive relay status," "brake pedal state," "driver leaving prompt," "main driver seat occupancy status," "driver seat belt status," "driver demand torque value," "handbrake status," "vehicle key status," "low voltage battery voltage," "the current gear status of the vehicle," "the current total current of the vehicle," "the current total voltage of the vehicle," "vehicle mileage," "speed," and "steering wheel angle." Among them, features such as "battery pack main positive relay status" and "brake pedal state" are categorical features, while features such as "speed" and "steering wheel angle" are numerical features. For each category of features, this paper adopts three different coding methods to fully mine the relationship between these original features and vehicle collision. The first coding method is simple one hot coding. The final coding result is $H_1, H_1 \in R^{s*\mathrm{d}_1}$, where s is the total number of samples and d_1 represents the dimension of the final one hot code. The other two types of coding are realized by probability distribution. Specifically, for an original feature X, its possible value is $\{x_1, x_2, \cdots, x_k, \cdots x_c\}$, where c is the total number of categories.

The second coding method is realized by using the collinear probability of collision between the feature and the vehicle:

$$P(Y = 1, X = x_k) = \frac{\varphi(Y = 1, X = x_k)}{\varphi_*}, \tag{3}$$

where $\varphi(Y = 1, X = x_k)$ indicates that there is a collision and the value of the last time point of feature X is the current number of vehicles x_k to be coded, where φ_* represents the total number of vehicles in the sample set and P represents a probability value, which is the coding of features $X = x_k$. The final coding result of all category features obtained by this coding method is $H_2, H_2 \in R^{s*d_2}$, where d_2 represents the feature coding dimension using cooccurrence probability for category features.

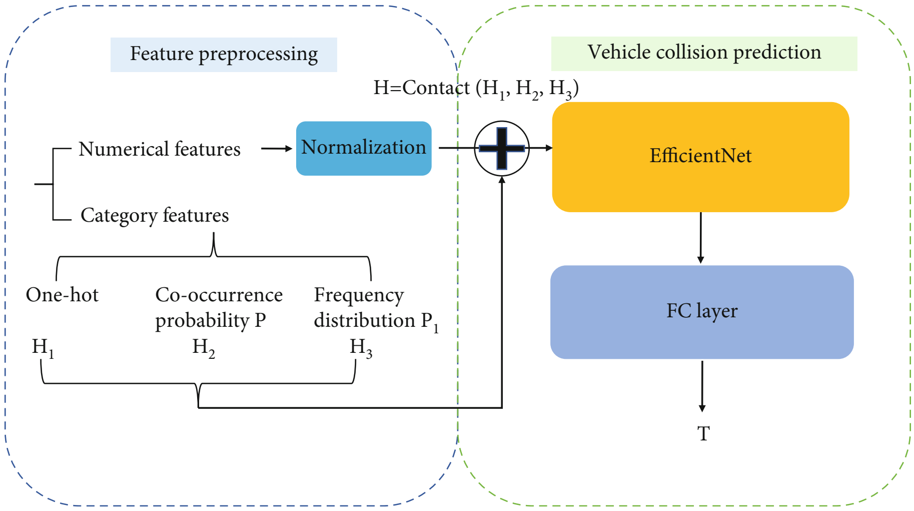
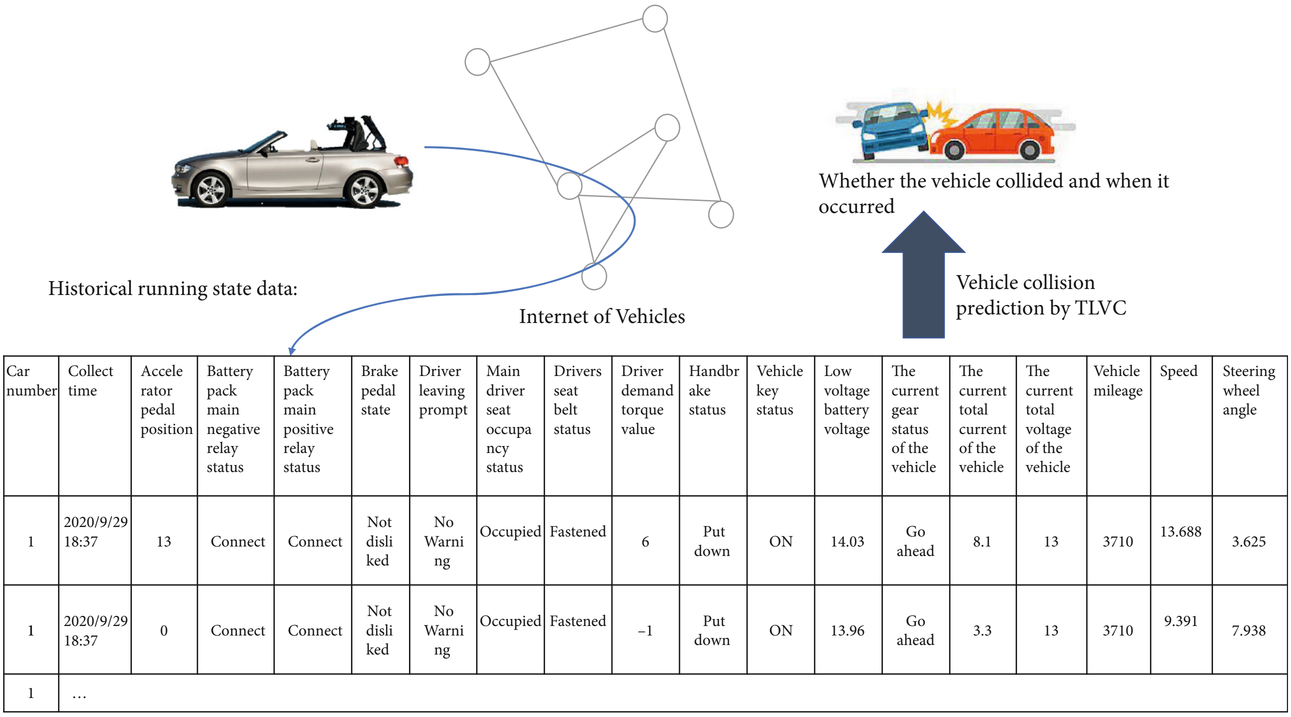

Feature preprocessing
Vehicle collision prediction
H=Contact (H$_1$, H$_2$, H$_3$)
Numerical features
Normalization
EfficientNet
Category features
FC layer
One-hot
H$_1$
Co-occurrence probability P
H$_2$
Frequency distribution P$_1$
H$_3$
T

Figure 1: Framework of vehicle collision prediction model based on transfer learning TLVC.


Historical running state data:
Internet of Vehicles
Whether the vehicle collided and when it occurred
Vehicle collision prediction by TLVC

Car number	Collect time	Accelerator pedal position	Battery pack main negative relay status	Battery pack main positive relay status	Brake pedal state	Driver leaving prompt	Main driver seat occupancy status	Drivers seat belt status	Driver demand torque value	Handbrake status	Vehicle key status	Low voltage battery voltage	The current gear status of the vehicle	The current total current of the vehicle	The current total voltage of the vehicle	Vehicle mileage	Speed	Steering wheel angle
1	2020/9/29 18:37	13	Connect	Connect	Not disliked	No Warning	Occupied	Fastened	6	Put down	ON	14.03	Go ahead	8.1	13	3710	13.688	3.625
1	2020/9/29 18:37	0	Connect	Connect	Not disliked	No Warning	Occupied	Fastened	−1	Put down	ON	13.96	Go ahead	3.3	13	3710	9.391	7.938
1	…																	

Figure 2: Example of vehicle collision prediction.

In the third coding method, we consider the frequency distribution of the vehicle for feature X in the whole time detection window:

$$\mathrm{P}_1 = \frac{F(Y=1, \mathrm{X}=\mathrm{x}_k)}{\sum_{j=1}^{j=c} F(Y=1, \mathrm{X}=\mathrm{x}_j)}, \tag{4}$$

where $F(Y=1, X=x_k)$ represents the number of times that the value of feature X is x_k in the detection window of the vehicle in collision. $\sum_{j=1}^{j=c} F(Y=1, X=x_j)$ represents the total number of samples for all vehicles involved in the collision. The third coding method reflects the occurrence probability of various features in the detection window. The final coding result obtained by this coding method is $H_3, H_3 \in R^{s*d_3}$, where d_3 represents the coding dimension of the frequency distribution used in the window period for category features.

For numerical features, through simple normalization, the final feature code of this part is $H_4, H_4 \in R^{s*d_4}$, where d_4 represents the final dimension of this part of features. Finally, by splicing the above feature code H_1, H_2, H_3, and

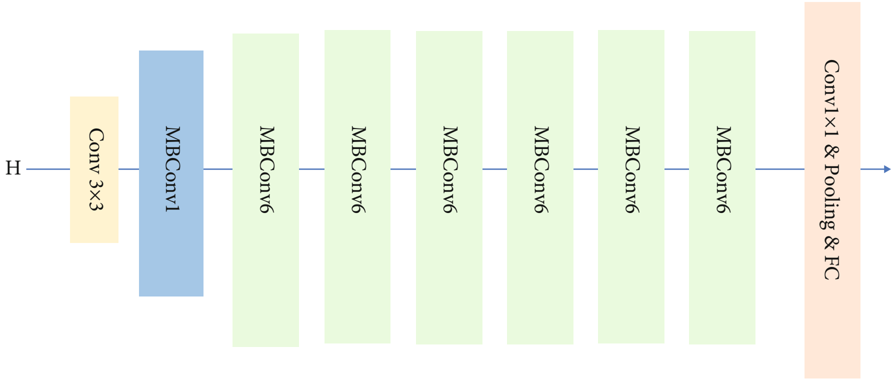


Figure 3: Structure diagram of EfficientNet-B0.

Table 1: Parameters setting.

Parameter	Description	Setting
lr	Learning rate	0.0001
w	Sampling window size	100
Hidden size	The parameter dimension corresponding to the two-layer fully connected neural network in the FC layer	512,256
Dropout	Discard parameter ratio	0.1
Epoch	Model training rounds	30

H_4 as the input h of the collision prediction model, follows $H = \text{contact}(H_1, H_2, H_3, H_4)$.

3.3. Prediction Model. This part corresponds to the right part of Figure 1. In order to obtain better results on the limited vehicle operation data set, this paper uses some parameters of the pretrained EfficientNet [24] to realize vehicle collision prediction through migration learning. In fact, the transfer learning method is widely used in the field of image processing. In this paper, the running state data of two-dimensional vehicles in the time window W is compared with the pictures in image processing, and then, the correlation between vehicle running data and vehicle collision is learned through EfficientNet. Then, make the model learn the dynamics of vehicle operation, so as to realize the modelling of dynamic behaviour habits of drivers.

Figure 3 shows the structure of B0 version of EfficientNet. EfficientNet achieves good results without consuming more computing resources by coordinating and controlling the depth, width, and input data size of the network at the same time. This relatively small and refined model is quite cost-effective for applications in the Internet of Vehicles. EfficientNet has been widely used in transfer learning in recent years and has performed well in various tasks, so this model is used in this paper. In addition, this paper also verified through experiments that compared with other transfer learning models, EfficientNet is a better choice for this task.

In our task, in order to make effective use of the parameters of the pretraining model, we frozen half the parameters in EfficientNet in the training process, and let the other half of the network parameters participate in the training of vehicle collision prediction model, which helps to localize the model parameters as much as possible. That is, on the basis of making full use of the existing network parameters, let the model learn the task characteristics of the current vehicle collision data set. In the EfficientNet model using migration, its input is H. As shown in Figure 1, its output will be input to the final full connection layer FC layer and then output the prediction T.

Finally, the parameters of the whole model are updated by Adam W. Since the two tasks of whether the final vehicle collides and the time of collision are combined into a unified regression problem through formula (2), MSE is used as the final optimization goal in the process of model training.

4. Experiment

In this section, we will use the real vehicle signal data of the Internet of Vehicles to verify the effectiveness and accuracy of the proposed vehicle collision prediction model TLVC based on transfer learning.

4.1. Experimental Settings. This study conducted all experiments on a computer with Intel(R) Core(TM) i7-11700F @ 2.50 GHz and 16 GBDRAM. We implemented these algorithms in Python 3.6. The data used in this paper comes from 2021 Digital China Innovation Contest (https://www.datafountain.cn/competitions/500) and includes a series of vehicle operation data, vehicle collision labels, and collision time. The data includes the operation data of 120 vehicles in 2-5 days in total. The number of detection data of each vehicle is at least 4324 and at most 114460. For building a more accurate vehicle collision prediction model, this paper

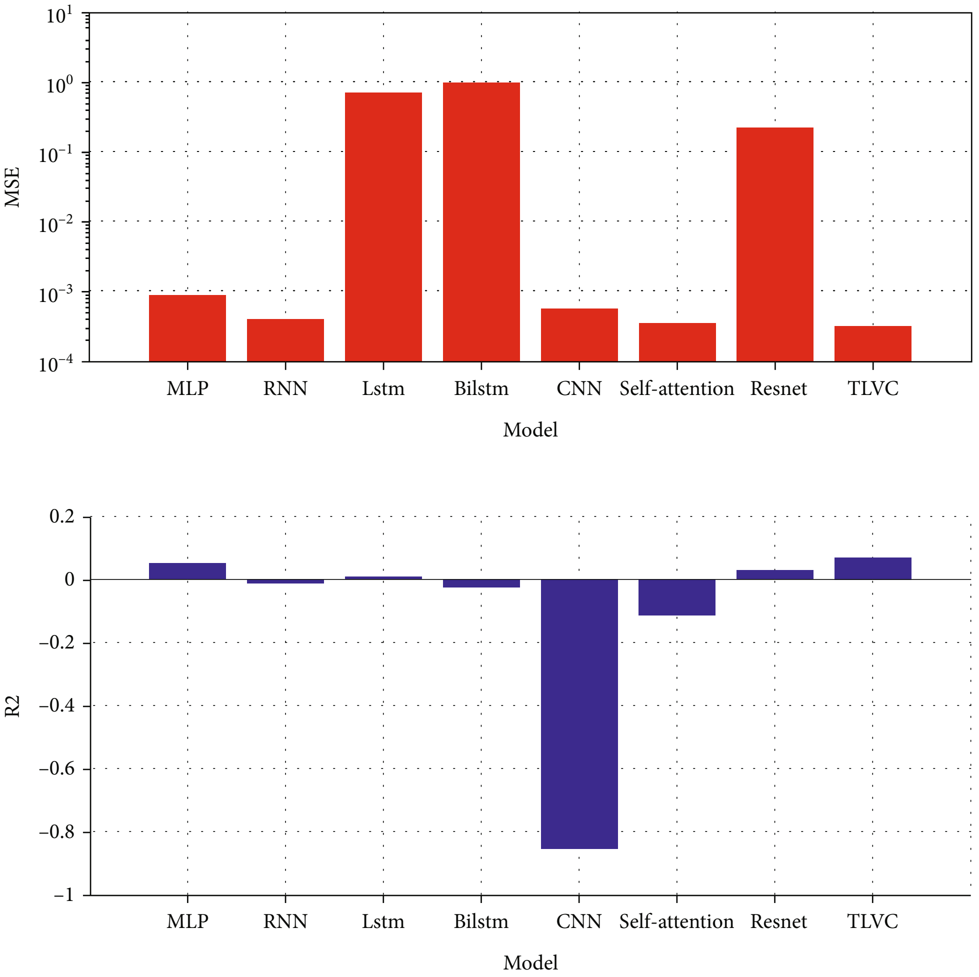


Figure 4: Performance of each model.

Table 2: Ablation analysis.

Model	MSE	R^2
TLVC	0.00019468	0.02050729
$\text{TLVC}_{[f]}$	0.00032577	-0.03514864

takes the data of 10 vehicles as the verification set and the remaining 110 vehicle data as the training set.

To enrich the data set and deal with the long vehicle state data, we truncate all vehicle detection data; that is, take consecutive w records as a sample data. After such processing, we finally obtained 355,509 training samples and 42,058 test samples. The relevant parameter design of this paper is shown in Table 1.

To verify the effectiveness of the proposed TLVC method, this paper compares the proposed model with the following methods:

MLP: multilayer perceptron, a simple neural network model, is used as a comparison method in this paper

RNN: recurrent neural network (RNN) compared with the general neural network; this method can deal with the data with sequence changes

LSTM: long short-term memory; LSTM is a special RNN, which is mainly used to solve the problems of gradient disappearance and gradient explosion in the process of long sequence training. Compared with ordinary RNN, LSTM can perform better in longer sequences

BiLSTM: bidirectional LSTM is an extension of LSTM. Because two LSTM can be trained in two directions, the performance of sequence prediction model can be improved

Self-attention: this method is widely used in the field of natural language processing because of its excellent performance [25]. Vehicle running state data is a time series data, which is very similar to the text in natural language processing. Therefore, this method is used as a comparison method in this paper

CNN: convolutional neural network [26], a method widely used in the field of image processing, is used to process two-dimensional vehicle running state data in this paper

ResNet: deep residual network; ResNet is an excellent model that improves CNN in the field of image processing [27]

To measure the accuracy of the proposed vehicle collision prediction model, MSE is used to measure the accuracy

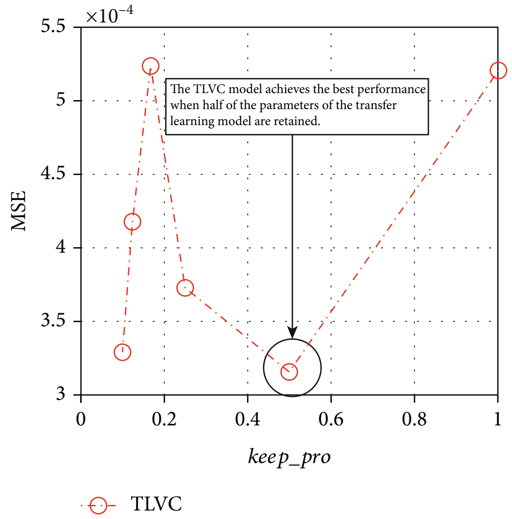


Figure 5: Impact of different keep_pro.

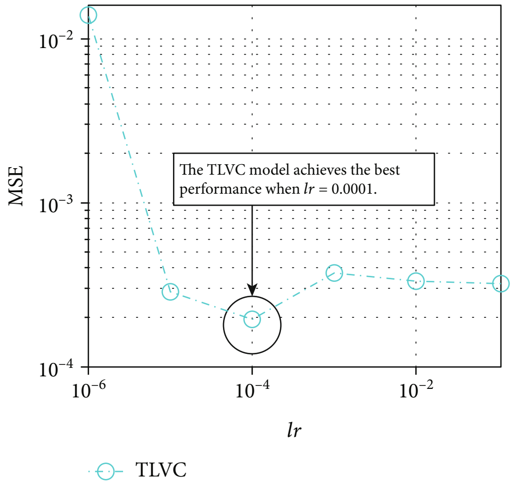


Figure 6: Impact of different lr.

of the prediction model. In addition, in order to evaluate the robustness of the model, another evaluation index R^2 in the regression problem is used to assist the evaluation of the model.

4.2. Comparison Result. In this part, we show the results of the proposed TLVC model compared with various benchmark methods. For the proposed TLVC, we use EfficientNet B4 version. The pretraining model has 1,918,200 parameters in total. The comparison results between the proposed TLVC model and each benchmark model are shown in Figure 4.

As shown in Figure 4, we compared the proposed model with other model benchmark models on the Internet of Vehicles data set from the real world. Among them, MSE reflects the prediction accuracy of the model. It can be seen from the results in Figure 4 that the prediction error of the proposed TLVC model is smaller than that of other benchmark models. As can be seen from Figure 4, using the methods of processing serialized data, LSTM and BiLSTM will obtain large model errors, and the results of these two models are even worse than RNN. This may be because these two models can only learn the characteristics of time series, and this length may lead to the disappearance of gradients in the learning process of these sequence models due to the sampling window $w = 100$. Unexpectedly, the MLP with simpler structure and composed of two-layer fully connected neural network obtained lower prediction error than LSTM and BiLSTM in the experiment. This may be because the fully connected neural network learns more effective features in the whole monitoring window, rather than paying too much attention to the vehicle state transition process in the window as the time series model. The prediction error of CNN model is slightly better than MLP, but its stability is much lower than other models according to the results of R^2. This may be because only the local information of vehicle state change in the monitoring window is extracted through convolution, so it is difficult to infer the final collision of vehicles. The MSE performance of self-attention model is second only to the proposed TLVC model, because it can fully learn the dynamic changes and even correlation of various vehicle characteristic signals in the monitoring window. However, its R^2 value is not high, which may be due to the lack of training data, so the R^2 value of the model is low. The R^2 of ResNet and TLVC with migration model is slightly higher than that of self-attention, mainly because the migration model is less dependent on the amount of data. However, ResNet is similar to CNN model. Because it only focuses on the local features in the monitoring window, the final MSE is large.

On the one hand, the proposed TLVC has better model performance when there is only small training data because the way of transfer learning depends less on data. On the other hand, when using the transfer learning method EfficientNet, we only retain half of the parameters of the original model, and the other half of the model parameters can be learned with the model training, which makes the TLVC model not only do not rely too much on large quantities of data but also carry out effective localization learning.

4.3. Ablation Experiment. In order to verify the effectiveness of the vehicle state preprocessing part proposed in this paper, this paper compares the proposed model TLVC with the preprocessing module of removing early features, and the comparison results are shown in Table 2.

As can be seen from the data in the table, the model accuracy of $\mathrm{TLVC}_{[f]}$ without feature preprocessing is slightly poor. This shows the effectiveness of vehicle running state data processing in this paper. In particular, this paper encodes a series of category data. The results of ablation experiments show the effectiveness of these coding features in the proposed model, which provides ideas for the effective mining and application of Internet of Vehicles data.

4.4. Parameter Learning. In order to make the model as effective as possible, in this part, we analyse the influence of some parameters in the proposed model, such as model learning rate lr and number of reserved copies of migration model parameters keep_pro, on the final effect of the model, as shown in Figures 5 and 6.

As shown in Figure 5, keep_pro is the number of pretrained transfer learning model parameters reserved for the model, keep_pro = 0.5 means that half of the parameters in efficientnet B4 are retained, and the other half of the parameters are trained and learned through the local vehicle collision data set. It can be seen from the figure that when keep_pro = 0.5, the model can achieve low prediction error.

Figure 6 shows our discussion on the learning rate lr of the whole TLVC model. Through experiments, we found that the model can achieve better results when lr = 0.0001.

5. Conclusion

In this paper, we use the vehicle running state data and propose a model TLVC which can predict whether and when the vehicle will collide in the future based on the migration learning model. Compared with the previous methods, this method does not need to rely on external unstable environmental data. It only needs to effectively process the vehicle operation signal by using the proposed preprocessing method and then use the semiparametric migration model for local data training to achieve high accuracy. In particular, because some parameters of TLVC migration model do not need training, learning on less vehicle operation data can achieve the purpose of task localization. Furthermore, by comparing with a series of state-of-the-art benchmark models, we verify the outstanding effect of the proposed method through a large number of experiments.

However, there are still some limitations. For example, the preprocessing of the vehicle operating state also relies on human understanding of the data, which is one of the issues we will further explore later. In addition, traditional transfer learning is usually the transfer of data in the same field, and this paper is limited by the limited sources of Internet of Vehicles data and uses data from different fields. Therefore, if there is a chance to obtain the same type of data in the future, we will discuss more effective transfer model.

Data Availability

The data set used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] National Highway Traffic Safety Administration, “2015 motor vehicle crashes: overview,” *Traffic Safety Facts: Research Note*, vol. 2016, pp. 1–9, 2016.

[2] A. V. Malawade, S. Y. Yu, B. Hsu, D. Muthirayan, P. P. Khargonekar, and M. A. Al Faruque, “Spatio-temporal scene-graph embedding for autonomous vehicle collision prediction,” *IEEE Internet of Things Journal*, vol. 9, no. 12, pp. 9379–9388, 2022.

[3] X. Wang, J. Liu, T. Qiu, C. Mu, C. Chen, and P. Zhou, “A real-time collision prediction mechanism with deep learning for intelligent transportation system,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 9, pp. 9497–9508, 2020.

[4] N. Lyu, J. Wen, Z. Duan, and C. Wu, “Vehicle trajectory prediction and cut-in collision warning model in a connected vehicle environment,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 2, pp. 966–981, 2020.

[5] L. Peng, M. A. Sotelo, Y. He, Y. Ai, and Z. Li, “Rough set based method for vehicle collision risk assessment through inferring driver's braking actions in near-crash situations,” *IEEE Intelligent Transportation Systems Magazine*, vol. 11, no. 2, pp. 54–69, 2019.

[6] Y. Zhang, Y. Li, R. Wang, M. S. Hossain, and H. Lu, “Multi-aspect aware session-based recommendation for intelligent transportation services,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4696–4705, 2021.

[7] M. Lotfollahi, M. Naghipourfar, M. D. Luecken et al., “Mapping single-cell data to reference atlases by transfer learning,” *Nature Biotechnology*, vol. 40, no. 1, pp. 121–130, 2022.

[8] W. Li, R. Huang, J. Li et al., “A perspective survey on deep transfer learning for fault diagnosis in industrial scenarios: theories, applications and challenges,” *Mechanical Systems and Signal Processing*, vol. 167, article 108487, 2022.

[9] Z. Zhang, C. Sun, and B. Guo, “Transfer-learning guided Bayesian model updating for damage identification considering modeling uncertainty,” *Mechanical Systems and Signal Processing*, vol. 166, article 108426, 2022.

[10] E. Candela, Y. Feng, D. Mead, Y. Demiris, and P. Angeloudis, “Fast collision prediction for autonomous vehicles using a stochastic dynamics model,” in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pp. 211–216, Indianapolis, IN, USA, 2021.

[11] J. Lee, S. Kishino, and K. Suzuki, “Prediction of collision avoidance ability of two-wheeled vehicle riders using driving behaviors and emotional states,” *International Journal of Automotive Engineering*, vol. 12, no. 2, pp. 32–40, 2021.

[12] L. Zhang, K. Yuan, H. Chu et al., “Pedestrian collision risk assessment based on state estimation and motion prediction,” *IEEE Transactions on Vehicular Technology*, vol. 71, no. 1, pp. 98–111, 2022.

[13] C. Katrakazas, M. Quddus, and C. W-H, “A new integrated collision risk assessment methodology for autonomous vehicles,” *Accident Analysis & Prevention*, vol. 127, pp. 61–79, 2019.

[14] C. Wang, C. Xu, and Y. Dai, “A crash prediction method based on bivariate extreme value theory and video- based vehicle trajectory data,” *Accident Analysis & Prevention*, vol. 123, pp. 365–373, 2019.

[15] F. Zhuang, Z. Qi, K. Duan et al., “A comprehensive survey on transfer learning,” *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2021.

[16] S. Ruder, M. E. Peters, S. Swayamdipta, and T. Wolf, “Transfer learning in natural language processing,” in *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: Tutorials*, pp. 15–18, Minneapolis, Minnesota, 2019.

[17] M. Raghu, C. Zhang, J. Kleinberg, and S. Bengio, “Transfusion: understanding transfer learning for medical imaging,” *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[18] B. Neyshabur, H. Sedghi, and C. Zhang, “What is being transferred in transfer learning?,” *Advances in Neural Information Processing Systems*, vol. 33, pp. 512–523, 2020.

[19] Y. Pathak, P. K. Shukla, A. Tiwari, S. Stalin, and S. Singh, “Deep transfer learning based classification model for COVID-19 disease,” *IRBM*, vol. 43, no. 2, pp. 87–92, 2020.

[20] J. Wang, Y. Chen, W. Feng, H. Yu, M. Huang, and Q. Yang, “Transfer learning with dynamic distribution adaptation,” *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 11, no. 1, pp. 1–25, 2020.

[21] S. Tammina, “Transfer learning using VGG-16 with deep convolutional neural network for classifying images,” *International Journal of Scientific and Research Publications (IJSRP)*, vol. 9, no. 10, pp. 143–150, 2019.

[22] Y. Chen, X. Qin, J. Wang, C. Yu, and W. Gao, “Fedhealth: a federated transfer learning framework for wearable healthcare,” *IEEE Intelligent Systems*, vol. 35, no. 4, pp. 83–93, 2020.

[23] R. Rao, N. Bhattacharya, N. Thomas et al., “Evaluating protein transfer learning with TAPE,” *Advances in Neural Information Processing Systems*, vol. 32, pp. 9689–9701, 2019.

[24] M. Tan and Q. Le, “Efficientnet: rethinking model scaling for convolutional neural networks,” *International Conference on Machine Learning*, vol. 97, pp. 6105–6114, 2019.

[25] Y. Zhang, Y. Li, R. Wang, J. Lu, X. Ma, and M. Qiu, “PSAC: proactive sequence-aware content caching via deep learning at the network edge,” *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 2145–2154, 2020.

[26] N. Kalchbrenner, E. Grefenstette, and P. Blunsom, “A convolutional neural network for modelling sentences,” https://arxiv.org/abs/1404.2188.

[27] Z. Wu, C. Shen, and A. Van Den Hengel, “Wider or deeper: revisiting the resnet model for visual recognition,” *Pattern Recognition*, vol. 90, pp. 119–133, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 3704720, 12 pages
https://doi.org/10.1155/2022/3704720

Research Article

Trusted Cloud Service System Based on Block Chain Technology

Tilei Gao, Xiaohui Jia, Rong Jiang, Yuanyuan He, Tao Zhang, and Ming Yang

School of Information, Yunnan University of Finance and Economics, Kunming 650221, China

Correspondence should be addressed to Ming Yang; yangming@ynufe.edu.cn

Received 17 May 2022; Revised 5 August 2022; Accepted 9 August 2022; Published 21 August 2022

Academic Editor: Yin Zhang

With the development and popularization of cloud computing technology, more and more users choose cloud services to build their application systems. With the improvement of users' understanding of software systems, in addition to functionality, trustworthiness has become another key issue concerned by users. Based on the existing research on cloud service trustworthiness measurement and evaluation, this paper proposes a cloud service-trusted delivery model based on asymmetric encryption and hash function and a trusted runtime model of cloud service system based on block chain technology. The proposed models will effectively solve the untrusted problems such as denial and tampering in the process of cloud service acquisition, as well as the system anomaly at runtime. Finally, through targeted experiments to verify the effectiveness and feasibility of the proposed models, and through experimental analysis, this paper expounds on the principle and mechanism of model operation and trustworthiness guarantee.

1. Introduction

The popularity of the Internet has had a significant impact on the development of computing mode and has realized a significant change from providing scientific computing power to providing network services [1]. At present, cloud computing has gradually become the main computing mode. According to Flexera's 2020 cloud status report, 59% of enterprises expect cloud usage to exceed previous plans [2]. It can be seen that the demand for cloud services in the global market is gradually increasing, and more and more institutions begin to choose to use cloud services to expand their applications. At the same time, the introduction of new technologies such as software-defined networks, artificial intelligence, and big data also poses a greater challenge to the trustworthiness of cloud computing [3]. With the increase of cloud services and the gradual deepening of users' understanding of software systems, users' needs are no longer limited to the satisfiability of functions but pay more and more attention to the trustworthiness of software systems.

In the cloud computing environment, the trustworthiness of cloud service systems is usually considered from the following three aspects: the selection of trusted cloud services, the trustworthiness of cloud services acquisition or delivery process, and the trustworthiness of cloud service running time. Through the review of the existing research on the trustworthiness of cloud services, it can be seen that most of the existing research results focus on the trustworthiness measurement, evaluation, and selection of cloud services. These results are indeed quite helpful for users to select reliable and appropriate services and build systems. For the cloud service providing platform, it is not difficult to find the services required by users from different cloud service marts. However, due to the different trustworthiness of different service providers in different fairs and the existence of unsafe factors in the network transmission process, how to ensure the trustworthiness of the transmission process of cloud services has become one of the focus issues of cloud platforms. In addition, during the operation of the cloud service system, how to provide continuous services to users without stopping the system is also one of the key concerns of the cloud platform.

The main research content of this paper focuses on the trustworthiness of the cloud service delivery process and the trustworthiness of the system runtime process. In the research of the trustworthiness of cloud service delivery process, a trusted delivery process model based on the hash function and asymmetric encryption technology is proposed. In the research of runtime trustworthiness, a runtime trustworthiness model based on the block chain technology is proposed. Through the design and analysis of simulation experiments, the feasibility and applicability of the models are verified.

The structure of the article is as follows:

Section 1 covers the introduction of this research study

Section 2 introduces the research status and achievements in related fields

Section 3 gives the definition and explanation of relevant basic concepts

Section 4 presents the trusted cloud service delivery process model

Section 5 presents the trusted cloud service system runtime model

Section 6 verifies the feasibility of the model through experimental analysis and expounds on the advantages and disadvantages of the models proposed

Section 7 is the conclusion and future works part

2. Literature Review

2.1. Research Status of Trustworthiness. Trustworthiness in the field of information science is often related to system requirements. In requirements engineering, nonfunctional requirements are usually regarded as the quality attribute of software. The quality of the software is actually an objective evaluation of software system, which will not change due to the difference in environment, personnel, and conditions [4]. Trustworthiness is an attribute with strong subjective preference. Its trustworthiness will show great differences in different application scenarios [5]. This difference is mainly reflected in the attention of users to the subattributes contained in software trustworthiness: some users pay attention to whether the functional requirements can be well realized, some users pay attention to whether the efficiency is high enough, and some users pay attention to whether the reliability is better guaranteed. These are the categories of software trustworthiness.

In terms of the trustworthiness content of software service systems, since the concept of trusted computing was put forward, the research on software trustworthiness [6] and trusted software [7] has gradually become one of the research hotspots in the field of software engineering. Its main research content is the construction and application of the trustworthiness model. According to the definition of TCG, an entity is credible if it always develops towards the expected goal. Shen [8] pointed out that trustworthiness includes reliability and safety. Yang et al. [9] pointed out that trustworthiness includes ability trustworthiness, integrity trustworthiness, predictability, correctness, privacy, and loss cost. In view of the trustworthiness of cloud platforms, Zhao et al. [10] pointed out that the content and analysis evaluation basis of cloud platform trustworthiness is not perfect and there is a lack of analysis and evaluation at the theoretical level. As a kind of evaluation, the trustworthiness evaluation of a cloud platform is inevitably vulnerable to human subjective factors in the evaluation process [11].

Research achievements in trustworthiness measurement and evaluation mainly include demand-driven software trustworthiness evaluation and evolution model [12], software trustworthiness evaluation model based on evidence theory [13], runtime software trustworthiness evidence collection mechanism based on TPM [14], software service trustworthiness evaluation method based on subjective and objective comprehensive weighting [15], evaluation method based on fuzzy theory [16], evaluation method based on D-S evidence theory [17], evaluation method based on risk matrix [18], trusted computing method based on trusted chain [19], and prediction evaluation method based on Bayesian network [20]. At the same time, there are also some general measurement and evaluation methods that can be used in various fields, such as the AHP method and extension method based on AHP [21], analysis method based on the decision tree and its deformation fault tree [22], and method based on information entropy and entropy weight theory [23].

Although there are many existing methods for system trustworthiness and trustworthiness measurement and evaluation, they are all for the measurement and evaluation of the service itself and are all preparations for the selection of services. In the traditional system field, users communicate directly with developers or system providers and need a set of methods to judge the trustworthiness of the system, so as to facilitate users to verify and accept the purchased system. However, under the new computing mode with cloud computing as the main technology, both users and cloud service system construction platforms obtain software services from unknown places to build the system. In this new computing environment, in addition to the traditional measurement and evaluation methods of software services and the software systems themselves, in order to ensure the trustworthiness of the remote systems, the trustworthiness of the service delivery process and service operation process should also be concerned.

2.2. Research Status of Block Chain. Block chain technology originated from bitcoin [24] and was originally designed to solve the problem of overreliance on trusted third parties in electronic payment. Block chain reorganizes mature technologies such as hash function, Merkle tree, and proof of work (POW) [25] and combines cryptography technologies such as public key encryption, digital signature, and zero-knowledge proof to become a new distributed infrastructure and computing paradigm [26]. Block chain solves the problem of consistency in distributed networks and subverts the traditional technical architecture of relying on trusted third parties to realize large-scale organization management and control. Its application has gradually extended to many fields, such as finance [27, 28], digital economy [29, 30], Internet of Things [31, 32], intelligent manufacturing [33, 34], and data security [35, 36], and has become a research hotspot in the global academic community.

Block chain covers a variety of technologies, the related concepts are easy to confuse, and there are many application scenarios. Therefore, relevant reviews have been made to sort out the latest progress, technical differences, and connections of block chain and summarize the technical form and application value from the perspective of technical architecture, technical challenges, and application scenarios. Yuan and Wang [26] gave the basic model of block chain. Taking bitcoin as an example, the unlicensed chain is divided into data layer, network layer, consensus layer, incentive layer, contract layer, and application layer; Shao et al. [37] compared the technical characteristics of various enterprise block chains (license chains) in combination with the details of open source projects; Yang et al. [38] summarized the characteristics, challenges, and development trends of block chain-based network service architecture; Han et al. [39] systematically summarized the research status of block chain security; Ali et al. [40] summarized the application research progress and trend of block chain in the Internet of Things.

Due to the two characteristics of process trustworthiness and decentralization, block chain can build a trusted foundation in a low-cost way in the scenario of multistakeholder participation, aiming to reshape the social credit system. In the cloud computing environment, the sources of cloud services that constitute service systems are diverse, and the existing mechanism is difficult to ensure the trustworthiness of service sources. In addition, the trustworthiness of the operation process of a cloud service system is the basis to ensure the normal operation of user's business. In the cloud environment, systems are far away from users. How to monitor the trustworthiness of the system in real time is also one of the key issues concerned by users and cloud platforms. The use of asymmetric encryption and hash function in block chain technology provides an idea for us to design a trusted service delivery process model. The characteristics of distributed ledger and centralization in the cloud in block chain provide a basis for the trustworthiness of users' real-time monitoring system. Therefore, based on block chain technology, this paper will build the trusted delivery model of cloud services and the runtime trusted verification model of cloud service systems.

3. Concepts

The main research object of this paper is the trustworthiness verification process of cloud service systems. The trustworthiness verification process includes two aspects: the trustworthiness verification of the transmission process and the trustworthiness verification of the running time.

The delivery process of cloud services, that is, the acquisition process of cloud services, refers to the process in which users select cloud services suitable for their application scenarios according to their needs and provide the selection results to the cloud platform and the cloud platform obtains cloud services from the service provider. This process needs to address the following trustworthiness issues:

(1) The confidentiality of the transmission process to avoid unauthorized access or interception

(2) The denial of the sender avoids the phenomenon that the sender sends malicious code and does not admit it

(3) The denial of the receiving party to avoid the phenomenon that the receiving party receives the service without acknowledging it

The main dependent technologies include asymmetric encryption technology in cryptography and hash function.

In terms of trustworthiness verification of cloud service systems at runtime, in the cloud computing environment, most cloud service individuals are function blocks that can run independently. After the cloud platform obtains the cloud services specified by the user, it is gradually integrated into its platform. In the integration process, a status is recorded for each service integrated. Taking the integrated state as the standard, if the state is inconsistent during operation, it indicates that the trustworthiness of the system has been damaged. In this way, the trustworthiness of the system at different times can be detected.

Starting from the first service construction, the status after each service addition, deletion, and modification is regarded as a block. Gradually build a block chain with each state in the construction process of cloud service system as the main body. Finally, the trustworthiness verification model of the operation process of cloud service system is constructed.

In order to realize the above two verification models, it is necessary to formally describe the relevant concepts. The following content of this section will formally describe the concepts used in the modeling process:

Definition 1 (cloud service system). Cloud service system (CSS) is a binary, CSS = (S, Q), during which

(1) S is the collection of cloud services which make up the cloud service system, and $S = \{\mathrm{ser}_1, \mathrm{ser}_2, \cdots, \mathrm{ser}_n\}$

(2) Q is the relationship matrix of CSS

Definition 2 (cloud service). Let ser be cloud service. Ser is a 5-tuples, and ser = (id, name, source, description, codes), during which

(1) Id is the number of the services making up system. And in a system, id is unique and assigned automatically

(2) Name is the name given by the system to be built according to its content specification, not by its developer

(3) Source represents the source of a cloud service

(4) Description represents the overall description of a cloud service

(5) Codes represent the source code package of a cloud service

Definition 3 (relationship matrix). Relationship matrix $Q = (q_{11}, q_{12}, \cdots, q_{ij}, \cdots, q_{nn})$. The value of q_{ij} means whether there is a message communication from ser$_i$ to ser$_j$, and

$$Q = \begin{bmatrix} q_{11} & \cdots & q_{1n} \\ \vdots & q_{ij} & \vdots \\ q_{n1} & \cdots & q_{nn} \end{bmatrix}, \tag{1}$$

$$q_{ij} = \begin{cases} 0, \text{ if } ser_i \longrightarrow ser_j = \varnothing, \\ 1, \text{ if } ser_i \longrightarrow ser_j \neq \varnothing. \end{cases} \tag{2}$$

In order to facilitate the calculation later, let R be the coding form of relation matrix Q, and $R = r_1 r_2 \cdots r_k \cdots r_{n*n}$. For instance, if the relationship matrix

$$Q = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \tag{3}$$

then the value of R is $R = 010001110$.

4. Trusted Cloud Service Delivery Process Model

The delivery process of cloud services is the process that the cloud platform obtains the required cloud services from different cloud service providers. The process includes three stages: the sending stage of cloud services, the transmission stage of cloud services, and the receiving stage of cloud services. The trusted cloud service delivery process should ensure the nonrepudiation of the sender and receiver of the service and the confidentiality and integrity of the transmission process. In order to achieve these trusted goals, this section proposes a trusted cloud service delivery process based on asymmetric encryption and hash function and designs a trusted delivery process from the three stages included in the delivery process.

The delivery process of cloud services includes service providers and service recipients. Assume that the service provider is A and the service recipient is B. The trusted cloud service delivery process must use the asymmetric encryption algorithm twice, which needs to generate the public key and private key of A and B, respectively. Suppose SK$_A$ and PK$_A$ represent the private and public keys of A, respectively, and SK$_B$ and PK$_B$ represent the private and public keys of B, respectively.

4.1. Trusted Sender. The trusted sender can be described as follows:

Definition 4 (trusted sender). Let TS be the trusted sender. TS is a 4-tuples, and TS = (S_S, $S_$ABR, $S_$SIG, S_C), during which

(1) *S_S* is the collection of cloud services defined before

(2) *S_*ABR is the collection of abstracts, and each abstract is a hash of its service, that is,

$$\forall \text{abr}_i \in S\_\text{ABR} : \text{abr}_i \longleftarrow \text{Hash}(\text{ser}_i) \tag{4}$$

(3) *S_*SIG is the collection of signatures, and each signature is obtained by encrypting the abstract, that is,

$$\forall \text{sig}_i \in \text{S}_{\text{SIG}} : \text{sig}_i \longleftarrow E_{\text{SK}_A}(\text{arb}_i) \tag{5}$$

$E_{\text{SK }A}$(arb$_i$) means encrypting arb$_i$ with SK$_A$ as the key. The encryption algorithm is selected according to the specific situation.

(4) *S_C* is the collection of cypher texts, which is the form of transmission on the network

$$\forall c_i \in S\_C : c_i \longleftarrow E_{\text{PK}_B}(\text{ser}_i + \text{sig}_i) \tag{6}$$

E_{PK_B} (ser$_i$ + sig i) means encrypting ser$_i$ and sig$_i$ as a whole, and the key is PK$_B$. The encryption algorithm is selected according to the specific situation. The algorithm can be the same or different with the algorithm in *S_*SIG.

The trusted sender model is shown in Figure 1.

4.2. Trusted Receiver. The trusted receiver can be described as follows:

Definition 5 (trusted receiver). Let TR be the trusted sender. *TR* is a 5-tuples, and TR = (R_S, $R_$ABR, $R_$ABR$^{'}$, $R_$SIG, R_C ; F), during which

(1) *R_S* is the collection of cloud services defined before

(2) *R_*ABR is the collection of abstracts, and each abstract is a hash of its service, that is,

$$\forall \text{abr}_i \in R\_\text{ABR} : \text{abr}_i \longleftarrow \text{Hash}(\text{ser}_i) \tag{7}$$

(3) *R_*ABR$^{'}$ is the collection of abstracts, which obtained by decrypting the signature sig$_i$, that is,

$$\forall \text{abr}_i{}' \in R\_\text{ABR}' : \text{abr}_i{}' \longleftarrow D_{\text{PK}_A}(\text{sig}_i) \tag{8}$$

(4) *R_*SIG is the collection of signatures, and each signature is obtained by encrypting the abstract, that is,

$$\forall \text{sig}_i \in R\_\text{SIG} : \text{sig}_i \longleftarrow D_{\text{SK}_B}(c_i) \tag{9}$$

$D_{SK\ B}(c\ _i)$ means decrypting c_i with SK_B as the key. The decryption algorithm is the inverse of $E_{\text{PK}_B}(\text{ser}_i + \text{sig}_i)$.

(5) *R_C* is the collection of cypher texts, which are contents received from the transmission process

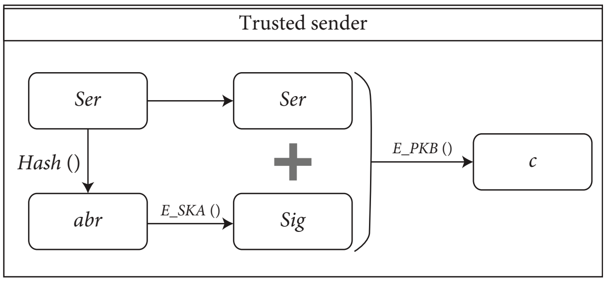


Figure 1: Trusted sender.

(6) F is a matching function, which is used to make sure whether abr_i and $\text{abr}_i^{'}$ are the same, and

$$F : f_i = \begin{cases} 0, \text{if } \text{abr}_i = \text{abr}_i', \\ 1, \text{if } \text{abr}_i \neq \text{abr}_i' \end{cases} \tag{10}$$

The trusted receiver model is shown in Figure 2.

4.3. Trusted Receiver. Based on the models of trusted sender and trusted receiver, the whole trusted cloud service delivery process model is shown in Figure 3.

The sender obtains the digest information of the service to be sold through hash operation and generates a signature from the summary information through the sender's private key SK_A. Package the signature with the cloud service and encrypt it with the public key PK_B of the receiver to get the cypher text packet sent to the receiver. In this process, the digest is used by the receiver to check whether the service has been tampered with during network transmission; the sender's signature generated by the sender's unique private key SK_A can determine the identity of the sender of the cloud service.

After receiving the cypher text packet, the receiver can decrypt the packet by using its unique private key SK_B. If the decryption is successful, the legal identity of the receiver can be confirmed. Split the decrypted packet into service body and signature. The received service body is used to recalculate the message digest, and the signature can be decrypted through the sender's public key PK_A to obtain its original information, that is, the message digest calculated by the sender through the hash function. If the digest calculated by the service is consistent with the digest contained in the signature, it indicates that the service body has not been tampered with in the process of network transmission. Otherwise, the service is tampered with during transmission, and it is not trusted.

5. Trusted Cloud Service System Runtime Model

Most of the existing research results on the trustworthiness of cloud service systems are aimed at the trustworthiness evaluation under the state of system shutdown, that is, the static evaluation results. Once the system runs, it is difficult to obtain real-time, effective, and reliable evaluation results. Based on the previous research results, this paper takes the static evaluation results after system construction as the standard, takes the state of cloud service system at different times as the main block body, and constructs a system state chain based on block chain technology. After the system runs, by calculating the system state at different times and comparing the results with the corresponding blocks in the state chain, the system can be verified whether it has been tampered with, so as to ensure the trustworthiness of the system running time.

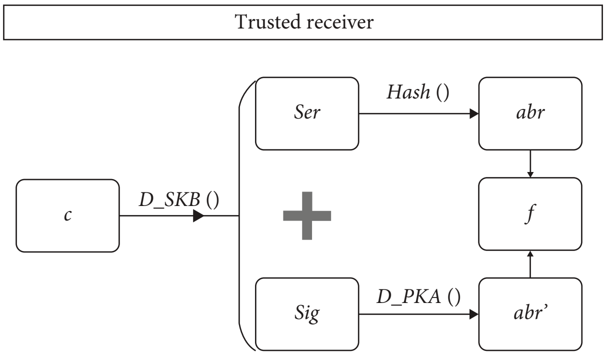


Figure 2: Trusted receiver.

Definition 6 (state chain). Let state chain be SC, and SC is a 3-tuples. $\text{SC} = (\text{st}, B, N)$, during which

(1) st represents the start block of a state chain of a system, and

$$\text{st} = \text{Hash}(\text{sys}_{\text{info}}) \tag{11}$$

In the formula, sys_info represents the overall description of the system and its contents and formats are designed based on the actual needs.

(2) B is the collection of state blocks except the start block st. $B = \{b_1, b_2, \cdots\}$

(3) N is the collection of flows between each blocks

Definition 7 (state block). State block b is a binary, and $b = (\text{Header}, \text{body})$.

Definition 8 (state block body). State block body is a binary, and $\text{Body} = (S, R)$. S is the collection of cloud services in Definition 1 and R is the value of relationship matrix Q in Definition 3.

Definition 9 (state block header). State block header is a binary, and $\text{Header} = (\text{PreHash}, \text{RootHash})$, during which

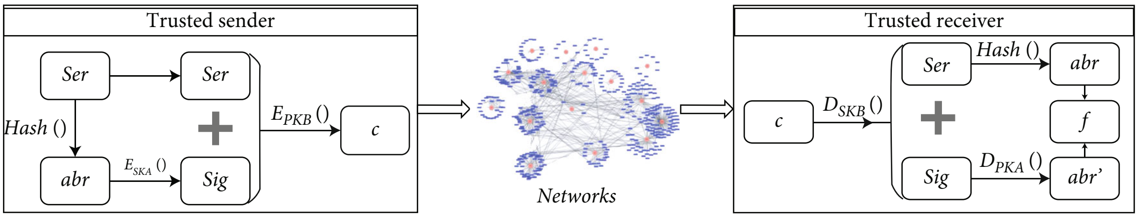


Figure 3: Trusted delivery process.

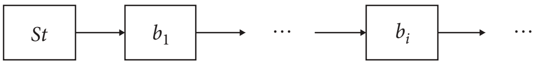


Figure 4: State chain.

Header contains the previous block hash PreHash and current block hash RootHash.

$$\text{PreHash} = \text{Hash}\left(\text{Header}_{i-1}\right), \tag{12}$$

$$\text{RootHash} = \text{Hash}\left(b_i\right). \tag{13}$$

Based on Definitions 6, 7, 8, and 9 and the basic concepts and principles of block chain, the state chain model is shown in Figures 4 and 5.

In Figures 4 and 5, b_i is the abbreviation of body$_i$, representing a certain block. s_i is the abbreviation of service$_i$, representing a certain service. r_i is the abbreviation of relationship$_i$, representing a certain matrix. hashs$_i$ is the hash value of s_i. hashs$_i$ is the hash value of s_i. hashs$_{ij}$ is a combination of multiple hash values from hashs$_i$ to hashs$_j$.

For a new cloud service system, the relevant description of the system itself can be used as the original information of the head node, and the calculated hash value can be used as the head node. For the heritage system, the heritage information to be improved can be packaged and its hash value can be calculated as the head node. The head node is not used as the standard or basis for judging the trustworthiness of the system.

Ordinary nodes are composed of general state blocks. The state block consists of two parts: the head and the main body of the state block. The header contains the hash value of the previous state block, while the main part is composed of the current system composition. Treat each service constituting the system as a transaction record in the block chain and calculate the hash value. Then, find out the interaction information of each service in the system, obtain the structure matrix of the system, take it as the last transaction record of the state block, and calculate the hash value. Then, all hash values calculated by the service and structure are combined to finally obtain the hash value RootHash about the current state block body in the state block header. The size of each state block body varies according to the number of services. However, according to the characteristics of the hash function, no matter how much data is involved in the operation, using the same hash function will get the same output. Therefore, except for the head node, the length information of each block in the state chain is the same regardless of the state of the system.

In principle, the state of the system is unlimited, and services can be added, deleted, and modified at any time. Therefore, the state chain of the cloud service system can be unlimited in principle. However, in the actual use process, during the system construction process, the state blocks will be added more frequently. After the system is stable, there will be relatively few adjustments and modifications to the system. Therefore, the growth of the state chain will slow down significantly. Any software system has its life cycle and will be abandoned. When users give up updating or building a new system, the existing state chain can be ended. The new system rebuilds the new head node and rebuilds the chain according to the corresponding rules. The status chain is a kind of alliance chain and belongs to the cloud service system platform. According to the principle of block chain distributed accounting, users can store the status chain and detect the trustworthiness of the system at the current time according to the status chain and the current running state of the system.

In particular, during the construction and operation of the service system, any active changes are system state changes, which should all be recorded in the state chain. Whether adding, deleting, or replacing services is a change in the overall state of the system. At the same time, only record one state change at a time. That is, if there are multiple services needed to be changed at the same time, treat each service as a separate individual and make changes in turn. Every service change is recorded into the chain as a change of state. The order of modification can be preset according to the actual needs of users or according to the classic principle of first come, first serve. In addition, each state block in the state chain is relatively independent. If there are multiple repeated operations, any variant will be added to the state chain as a new state according to the previous rule. For example, if service s_1 is added at a certain time and deleted at the next time, then the original service s_1 is added. Then, this process will produce three state blocks, namely, adding s_1 block, deleting s_1 block, and adding s_1 block. Due to the different time of adding, the values of these three state blocks will not be the same.

In terms of the trustworthiness verification method of a cloud service system, whether it is a newly built cloud service system or a cloud service system in stable operation, any active modification or adjustment is recorded as a state block and added to the state chain. Before any active modification operation, the new trustworthiness state of the modified system can be predicted by the existing trustworthiness

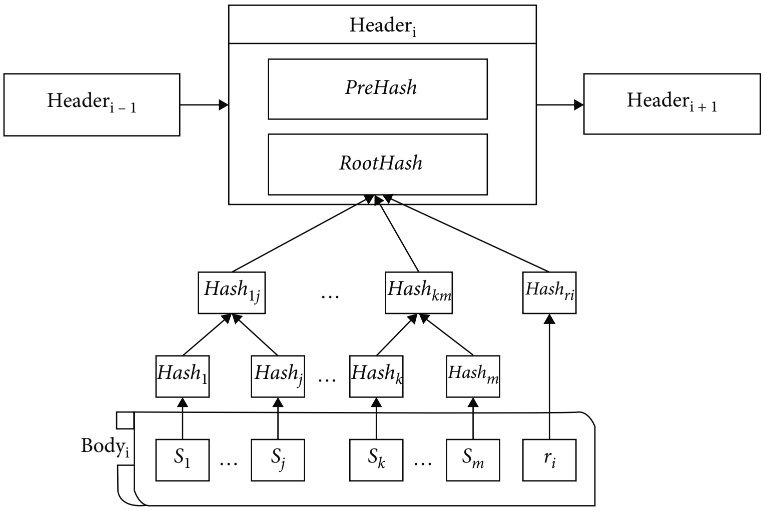


Figure 5: Block body structure of state chain.

measurement and evaluation methods. In the same environment, if the services constituting the system and the interaction relationship between services remain unchanged, its trustworthiness state remains unchanged. Then, any passive change in the system can be regarded as the destruction of trustworthiness. Therefore, the current state information of the service system can be calculated in real time and compared with the information at the initial stage of the state stored in the state chain. If it is the same, it indicates that it has not been invaded and its trustworthiness can be maintained. If it is different, it indicates that there are passively modified contents in the system, which can be traced back to the front block to determine the scope of malicious modification. Therefore, the construction of a state chain can not only determine the trustworthiness of the current system but also help the platform find the tampered services within a certain range.

6. Experiments and Analysis

The purpose of the experiment is to verify the effectiveness of the trusted model of the cloud service delivery process and the trusted model of system runtime. There is a direct correlation between the two models, but they are relatively independent. In terms of relevance, the transmitted result is the input of the system integration process. If the service received by the cloud platform is not trusted, the next process of building the system cannot be carried out. In terms of relative independence, after the cloud platform receives trusted cloud services, the subsequent system and chain-building process will not be affected by the delivery process. In general, the transfer process and the construction process of the state chain also belong to different modules. Therefore, in order to more specifically verify the effectiveness of the model and analyze the feasibility and trustworthiness of the model itself, this section designs experiments for the two models, respectively, and makes an independent analyses for each experiment. If the trustworthiness of each model is verified, the overall trustworthiness can be maintained.

6.1. Experiment and Analysis of Trusted Model of Cloud Service Delivery Process

6.1.1. Experiment Design. Suppose ser is a cloud service. According to Definition 2, ser can be described like this,

The source of the service can be a link or a user, which can be set according to the actual situation. The description of a service can generally be represented by the hash value of the service. The source code of the service or the service itself is generally a package. Here, in order to simplify the operation, the access link of the service is used as the code. The effect is the same as that of the package.

In the verification process, SHA256 and RSA have been selected as the hash function and asymmetric encryption algorithm. Suppose A is the sender and B is the receiver. Then, SK_A presents the private key of A; PK_A presents the public key of A; SK_B presents the private key of B; PK_B presents the public key of B. To simplify the operation, the key length is 512 bits and the format is PKCS#1.

Step 1. According to formula (4), the digest abr of ser can be obtained.

abr =
"6f55d6f3081f8426433d184180056c15789225cc6a7fee23c6049b1fd02bdfb1"

Step 2. According to formula (5), the signature sig of ser can be obtained.

sig = *"k7QvgUZmLjLpjvR+p3/BQNawxLzlj7HtvwKgzFEAJ40/ACDtDSL00G6qcvx2aSVh+g1dJkPnSjz/EiABSvhmfw=="*

Step 3. According to formula (6), the cyphertext can be obtained.

```
ser = (0001,
"Gzipped source tarball",
"https://www.python.org/downloads/release/python-3104/",
"7011fa5e61dc467ac9a98c3d62cfe2be",
"https://www.python.org/ftp/python/3.10.4/Python-3.10.4.tgz")
```

Code 1

```
SK_A =
"———BEGIN RSA PRIVATE KEY———
MIIBOwIBAAJBAOvcQ2C/AAmlAcFs
jZbieobeeB2bFMfS+jsHlhezr344fr7ih
MBinbdWsUUScN9nR0aJxjF3
YKB0CaKfKSF0xMkCAwEAAQJBAJ6q9sjGtQfH8X5l
wHqYsUS5tKR2B2zGCYBcgiQ/xPdrT3sdz
WcyXSppL2uceEwimHtOegl+I6uih4VC
58UNouECIQD27HhcVpUd0i/54
NxRMDunZAyxlNdXYsv360dCr2WuGwIhAPSHs
DSOCX2vIvFt5bz4FYhkVFtn2qEVfErmhmq
acPbrAiAJ2j+nN5E1omhlqRJBbxJCS
Jy1DUJWa0vGNa4fPA5rlwIgInFGSX
DEN3bGtjjjhiVvawGuvB05tzy+gBJOVo+
gX7cCIQDlnx96LQoWPOErGn4
etaM4aJkcHBMSIR6nbVAbaMnQwA==
———END RSA PRIVATE KEY———"PK_A =
"———BEGIN RSA PUBLIC KEY———
MEgCQQDr3ENgvwAJpQHBbI2W4nqG3ngdmxTH0vo7B5YXs69+
OH6+4oTAYp23VrFFEnDfZ0dGicYxd2CgdAminykhdMTJAgMBAAE=
———END RSA PUBLIC KEY———"SK_B =
"———BEGIN RSA PRIVATE KEY———
MIIBOgIBAAJBALJseZoYxQChT8PDv0tgrbkvPx/
ye5nu71Ye5hPvlfgM4VXubos
i119ZYLX0z5FLMqbAco8/Fa3sMRyZGYlxd6cC
AwEAAQJABJ4KB5LchkemaMqICMtXs
5Mlbw43ZKRqTTA/hASPPPwNSJrE61Rhkgnl0vErX/l
YMH40nQ1glcPjGWKJ5O8cAQIh
ANlR/BizAfpJeeq0pPythTsMBz0vWFdBN4A/V
zNmYOU3AiEA0i4z6qjWdq61N7CO3sro
9h6WoFrFSgaYyW1P5QJPORECIQCW8NGnGhYCkCw
kr4l0ktTZuTYB8jNqjzqMUfIwGii
sqwIgI8MGxGOr8g+x9+LLvG7MCqyTtn
8bWIgc0REPagjlj/EC
IGtS7b7QfoGsLY6jtC0S6MC5t
UBJmgArxH0xJnG7hzPZ
———END RSA PRIVATE KEY———"
PK_B =
"———BEGIN RSA PUBLIC KEY———
MEgCQQCybHmaGMUAoU/Dw79LYK25Lz8f8nuZ7u9WHuYT75
X4DOFV7m6LItdfWWC19M+RSzKmwHKPPxWt7DEcmRmJcXenAgMBAAE=
———END RSA PUBLIC KEY———"
```

Code 2

c = "*ixnJEQEQVEzcqZ5Ny1M7ILprCA1CNslfGnBfRGa EPETGrMImSDjqNNZwiWT5LAZTyh3u8KoV9sdro9d1xZy mdA==*

dAulJZqQCRA8ZstTD5hbyF3dow2i2otJDg//hIqV3EXL nSmA8gk50GHo2NuS1tv8ulIvPIDojLyN9ceBcg5BtQ==

DTqG2pgDWyWJ9Q17U9KORn4txRHqn3Pgxpy5S2d3 zdUwP3xKAcZQyMXxONijI32WfAW53CqDyR4d9xsB3W RkJw==

bb2cN1Yy79Mjt2TAvXM2Guj5hKqfU9b2JWJN2su4h/nd T9aN9OKxaV6ZSIDqxtYf2XN1klHTrgkQRhZ2r9UTxg=="

```
ser_1 = (0001,
"Gzipped source tarball",
"https://www.python.org/downloads/release/python-3104/",
"7011fa5e61dc467ac9a98c3d62cfe2be",
"https://www.python.org/ftp/python/3.10.4/Python-3.10.4.tgz").
ser_2 = (0002,
"XZ compressed source tarball",
"https://www.python.org/downloads/release/python-3104/",
"21f2e113e087083a1e8cf10553d93599",
"https://www.python.org/ftp/python/3.10.4/Python-3.10.4.tar.xz").
ser_3 = (0003,
"manpages-2.9.5.tar.xz",
"https:// http://edge.kernel.org/pub/software/scm/git/",
"cb6822a6eedd1682bbe815eb26d9fdbe",
"https:// http://edge.kernel.org/pub/software/scm/git/git-manpages-2.9.5.tar.xz")
ser_4 = (0004,
"B23Downloader-v0.9.5.7",
"https://github.com/vooidzero",
"4192a001fc65a0ad8016bf1328da28d0",
"https://github.com/vooidzero/B23Downloader.git")
```

Code 3

After the cyphertext c obtained, the sender's work has finished, and the cyphertext transmitted on the networks. When the receiver receives the cyphertext, the verification process begins.

Step 4. According to formula (9), cyphertext is decrypted, and the contents of ser and its signature sig can be obtained.

Step 5. According to formula (7), use the same hash function (SHA256) and recalculate the digest abr of ser.

Step 6. According to formula (8), decrypt the obtained signature sig to get a new digest, which is recorded as abr'.

Step 7. According to formula (10), compare the values of abr and abr' to judge the service integrity.

6.1.2. Experiment Analysis

(1) Confirmation of the sender's true identity: in Step 2, the sender encrypts the digest with its unique private key SK_A and gets the sender's signature sig_i, which cannot be forged. The service is delivered together with the signature. Due to the uniqueness of the signature, the sender must be responsible for the content of the service. Thus, the nonrepudiation of the sender can be guaranteed

(2) Maintenance of confidentiality: in Step 3, the cyphertext is encrypted by the receiver's public key PK_B, so only the receiver's private key can decrypt the cyphertext. Even if others intercept the cyphertext, they cannot know the content of the cyphertext. This ensures the confidentiality of the service

(3) Confirmation of the receiver's true identity: in Step 4, only the receiver's private key SK_B can decrypt the cyphertext of the receiver's public key PK_B. Once the cyphertext is decrypted, it must be the cyphertext received by the receiver to perform the decryption operation. Therefore, the decryption process itself can confirm the true identity of the receiver

(4) Integrity detection of service: the digest calculated by the service and the digest information saved in the signature is essentially calculated from the service. If the content of the service is tampered with during transmission, any small change will result in a completely different digest. Then, the new digest calculated by the receiver through the hash function must be inconsistent with the digest decrypted in the signature. Therefore, by comparing the calculated digest with the digest in the signature, whether the service has been tampered with during transmission can be confirmed, so as to the trustworthiness of the service

6.2. Experiment and Analysis of Trusted Model of Cloud Service System Runtime

6.2.1. Experiment Design. In order to verify the trustworthiness of cloud service systems based on block chain technology, this experiment is designed. This experiment focuses on the process of constructing the state chain, but there are no specific requirements for the service itself. Therefore, in order to more specifically verify the runtime trustworthiness model, our paper designed four simple services and randomly generated the message communication relationship between services, that is,

$$S = \{\mathrm{ser}_1, \mathrm{ser}_2, \mathrm{ser}_3, \mathrm{ser}_4\}, \tag{14}$$

```
st = Hash(sys_info) = "52a6cbc3f20ce4ebb293e7fcdb2722c3cbd3
04c6c4e99d41b72b87f4d2a922f4"
```

Code 4

Table 1: State chain.

State chain	Services and relationships contained	Contents of header
st	—	*52a6cbc3f20ce4ebb293e7fcdb2722c3cbd* *304c6c4e99d41b72b87f4d2a922f4*
st $\longrightarrow$ Header_1	ser_1, r_1	*816c02329c388c3012961c809a69911310636* *7d324a0551cb56aa4013382fee4* *020b9d64225f909183d641322dbf901cef11* *fc015156987e092c9f1f3445d214*
st $\longrightarrow$ Header_1 $\longrightarrow$ Header_2	ser $_1$, ser_2, r_2	*ef3daff6529f1874fbdd0a8219f5c246* *2ef658baaba3f645d6ff4ca82b67c2cb* *e625ac37260f7653cf7bb2c361b81b2ce1f* *77f03d1095e75ad3bbcb346003e80*
st $\longrightarrow$ Header_1 $\longrightarrow$ Header_2 $\longrightarrow$ Header_3	ser_1, ser_2, ser_3, r_3	*c52c489ca818292a6b2738efd17b83d208* *7f646d48aeb1acd28890a66058b541* *6df28493e4d015884fac6ca944be6f42* *8981c4d4ec8cd6598a339ff3ca154734*
st $\longrightarrow$ Header_1 $\longrightarrow$ Header_2 $\longrightarrow$ Header_3 $\longrightarrow$ Header_4	ser_1, ser_2, ser_3, ser_4, r_4	*72ed72575903894c75228f7ff294ec72* *2e2d57980295fe1340b7d306447240e2* *853f02d65c1f026c2fc40f627fb6ebfe* *4e231852e9bc583ec372f2f58071bafd*

$$Q = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \tag{15}$$

According to formula (2) and relationship matrix Q, $R = 0101001111000100$.

According to Definition 2, each service can be described as follows:

Hash function SHA256 is used during the chain building process. Suppose that the contents of sys_info is "This is the first nodes!", then, according to formula (11),

Then, the head of the state chain is built. Next, add the four services to the system in turn, and each addition or modification recreates the state block and adds it to the state chain. Then, the chain is established, which is shown in Table 1.

6.2.2. Experiment Analysis. Based on block chain technology, this paper constructs the state chain of cloud service systems. Take the construction process and modification process of cloud service systems as state information. Whether adding, deleting, or modifying system services and structures, it will actively trigger the chain building mechanism, establish new state blocks, and add them to the existing state chain. During the system runtime, any unauthorized modification in the system, whether caused by intrusion or system change caused by other reasons, will not trigger the chain building mechanism. Assuming that the status after each authorized addition, deletion, and modification is normal, the real-time status block information of the cloud service system will not match the status block modification in the chain after unauthorized modification, so the trusted status of the system can be determined. At the same time, by tracing back to the previous state in the state chain, the scope of unauthorized modification can be determined to a certain extent, so as to facilitate further error correction and improve the trustworthiness of the system.

```
"72ed72575903894c75228f7ff294e
c722e2d57980295fe1340b7d306447240e2
2efe352f56f83b3f8d7ac75af5ba555111908df6
6a0cfee65c29347f604e37aa"
```

Code 5

For example, in the current state, if the number of service 3 was tampered with to "003", by recalculating the status of the current system, the Header contents will change to

The PreHash will not change but the RootHash changed completely. By constantly deleting the services in the currently modified system, when the third service is deleted, the current system state is consistent with that in the normal chain. In this way, it can be judged that the problem lies in the third and fourth services.

7. Conclusion and Future Works

The trustworthiness of cloud service systems not only depends on each cloud service constituting the systems but also depends on the trustworthiness of the service acquisition process and running time. The trustworthiness of the acquisition process determines whether the real cloud service is obtained and whether it can be held accountable in case of service problems. The trustworthiness of running time determines whether the system maintains its normal running state during operation. Most of the tampered systems can also run, but they will get completely wrong results, which are difficult to find, resulting in huge losses. Aiming at the above two problems, this paper proposes a cloud service-trusted delivery process model based on asymmetric encryption and hash function and a cloud service system runtime trustworthiness model based on block chain technology. By using digital signature, asymmetric encryption, hash function, block chain, and other technologies, the trustworthiness of the service delivery process and trustworthiness detection at runtime are solved. Through the targeted experimental design, on the one hand, the feasibility and effectiveness of the model are verified. On the other hand, the principle and mechanism of ensuring trustworthiness are analyzed, which provides theoretical support for protecting the system rented by users on the cloud platform.

One of the deficiencies and shortcomings is that this paper focuses on the establishment of the trustworthiness models. The formal definitions of services on cloud platforms are slightly simple in content, which are difficult to cover various types of cloud services, and need to be optimized and improved in follow-up research. In addition, the existing trusted runtime model will be further studied in the future to strengthen the fault location ability of the model, so as to improve the weakness in the existing model that can only determine the approximate range of tampering.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the Yunnan Fundamental Research Projects (grant nos. 202201AT070142 and 202101AT070211), the Talent Introduction Projects of Yunnan University of Finance and Economics (grant no. 2021D16), and the Scientific Research Foundation of Yunnan Education Department (grant no. 2022J0475).

References

[1] D. Guo and H. Wang, “Development review and prospect of typical forms of network computing,” *Communications of the CCF*, vol. 18, no. 2, pp. 39–45, 2022.

[2] Flexera, *2020 State of the Cloud Report*, Flexera, America, 2020.

[3] X. Rong, “Research on information security in cloud computing network environment,” *Network Security Technology Surgery and Application*, vol. 7, no. 7, pp. 83-84, 2021.

[4] L. Li, C. Chen, Y. Li, and J. Li, “Overview of software fault localizaion technology,” *Computer Measurement and Control*, vol. 27, no. 5, pp. 1-4–121, 2019.

[5] H. Hu, D. Liu, and S. Wang, “Web ontology language OWL,” *Computer Engineering and Design*, vol. 30, no. 12, pp. 1-2–147, 2004.

[6] H. Chen, J. Wang, and W. Dong, “Highly trusted software engineering technology,” *Journal of Electronics*, vol. 31, no. 12, pp. 1933–1938, 2003.

[7] X. He, J. Tian, and F. Liu, “Overview of trusted cloud platform technology,” *Journal of Communications*, vol. 40, no. 2, pp. 154–163, 2019.

[8] C. Shen, “Scientific concept of network security and trusted computing,” *Information Technology and Network Security*, vol. 37, no. 1, p. 105, 2018.

[9] X. Yang, P. Luo, and G. Jabeen, “The concept model of software trustworthiness based on trust-theory of sociology,” *Acta Electronica Sinica*, vol. 47, no. 11, pp. 2344–2353, 2019.

[10] B. Zhao, Z. Dai, S. Xiang, and W. Tao, “A method of establishing cloud platform trustworthiness analysis model,” *Journal Of Software*, vol. 6, p. 17, 2016.

[11] T. Zhang, K. Zhao, M. Yang, T. Gao, and W. Xie, “Research on privacy security risk assessment method of mobile commerce based on information entropy and Markov,” *Wireless Communications and Mobile Computing*, vol. 2020, Article ID 8888296, 11 pages, 2020.

[12] S. Ding, F. Lu, S. Yang, and C. Xia, “A requirement-driven software trustworthiness evaluation and evolution model,” *Journal of Computer Research and Development*, vol. 48, no. 4, pp. 647–655, 2011.

[13] S. Yang, S. Ding, and C. Fu, “Software trustworthiness evaluation model considering information source correlation,” *Chinese Management Science*, vol. 17, no. 6, pp. 163–169, 2009.

[14] L. Gu, Y. Guo, H. Wang, Y. Zou, and B. Xie, “Runtime software trustworthiness evidence collection mechanism based on TPM,” *Journal of Software*, vol. 21, no. 2, pp. 373–387, 2010.

[15] Y. Zhang, Y. Yuan, X. Liu, and X. Sun, “Evaluation method of software service trustworthiness of E-commerce website,” *Computer Application Research*, vol. 37, no. 1, pp. 244-246–244-263, 2020.

[16] X. Hu, R. Jiang, M. Shi, and J. Shang, “A privacy protection model for health care big data based on trust evaluation access control in cloud service environment,” *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 3, pp. 3167–3178, 2020.

[17] D. Wang and Q. Wang, “Trustworthiness evidence supporting evaluation of software process trustworthiness,” *Journal of Software*, vol. 29, no. 11, pp. 178–200, 2018.

[18] R. M. C. Ratnayake and K. Antosz, “Development of a risk matrix and extending the risk-based maintenance analysis with fuzzy logic,” *Procedia Engineering*, vol. 182, pp. 602–610, 2017.

[19] W. Shang and X. Xing, “ICS software trust measurement method based on dynamic length trust chain,” *Scientific Programming*, vol. 2021, Article ID 6691696, 11 pages, 2021.

[20] P. Chen, X. Wang, and D. Dang, “Construction of model based on petri net and reliability analysis based on Bayes net of web

service transaction," *Journal on Communications*, vol. 39, no. S1, pp. 99–104, 2018.

[21] L. Shao, J. Zhang, Y. Wei, J. Zhao, B. Xie, and H. Mei, "Personalized QoS prediction forweb services via collaborative filtering," in *Presented at the IEEE International Conference on Web Services (ICWS 2007)*, Salt Lake City, UT, USA, 2007.

[22] L. Yao, Q. Z. Sheng, A. Segev, and J. Yu, "Recommending web services via combining collaborative filtering with content-based features," in *Presented at the 2013 IEEE 20th International Conference on Web Services*, Santa Clara, CA, USA, 2013.

[23] Y. Zhang, C. Yin, Z. Lu, and D. Yan, "Recurrent tensor factorization for time-aware service recommendation," *Applied Soft Computing*, vol. 85, no. 6, p. 105762, 2019.

[24] S. Nakamoto, "Bitcoin: a peer-to-peer electronic cash system," *Decentralized Business Review*, vol. 2009, Article ID 21260, 2009.

[25] C. Dwork and M. Naor, "Pricing via processing or combatting junk mail, in the 12th Annual International Cryptology Conference, California, USA, 2001, pp. 139-147: Springer, Berlin," in *Annual International Cryptology Conference CRYPTO 1992: Advances in Cryptology — CRYPTO' 92*, Lecture Notes in Computer Science book series (LNCS,volume 740), Springer, Berlin Heidelberg, 1993.

[26] Y. Yuan and F. Wang, "Blockchain: the state of the art and future trends," *Acta Automatica Sinica*, vol. 42, no. 4, pp. 481–494, 2016.

[27] R. Huang, "Research on supervision of financial blockchain technology," *Academic Forum*, vol. 39, no. 10, pp. 53–59, 2016.

[28] Y. Zhu, W. Song, D. Wang, D. Ma, and W. C.-C. Chu, "TA-SPESC: toward asset-driven smart contract language supporting ownership transaction and rule-based generation on blockchain," *IEEE Transactions on Reliability*, vol. 70, no. 3, pp. 1255–1270, 2021.

[29] L. Zhuang and C. Zhao, "Research on the evolution of digital currency under blockchain technological innovation: theory and framework," *The Economist*, vol. 5, no. 5, pp. 76–83, 2017.

[30] K. Li, Y. Liu, H. Wan, and Y. Huang, "A discrete-event simulation model for the bitcoin blockchain network with strategic miners and mining pool managers," *Computers & Operations Research*, vol. 134, pp. 105365–105365, 2021.

[31] J. Shi and R. Li, "Overview of blockchain access control under the Internet of Things," *Journal of Software*, vol. 30, no. 6, pp. 1632–1648, 2019.

[32] Z. Abubaker, N. Javaid, A. Almogren, M. Akbar, M. Zuair, and J. Ben-Othman, "Blockchained service provisioning and malicious node detection via federated learning in scalable Internet of Sensor Things networks," *Computer Networks*, vol. 204, pp. 108691–108691, 2022.

[33] M. Suvarna, K. S. Yap, W. Yang, J. Li, Y. T. Ng, and X. Wang, "Cyber-physical production systems for data-driven, decentralized, and secure manufacturing–a perspective," *Engineering*, vol. 7, no. 9, pp. 1212–1223, 2021.

[34] Y. Chengyue, M. Prabhu, M. Goli, and A. K. Sahu, "Factors affecting the adoption of blockchain technology in the complex industrial systems: data modeling," *Complexity*, vol. 2021, 10 pages, 2021.

[35] M. Liu, Z. Chen, Y. Shi, L. Tang, and D. Cao, "Research progress of blockchain in data security," *Chinese Journal of Computers*, vol. 44, no. 1, pp. 1–27, 2021.

[36] N. Deb, M. A. Elashiri, T. Veeramakali, A. W. Rahmani, and S. Degadwala, "A metaheuristic approach for encrypting blockchain data attributes using ciphertext policy technique," *Mathematical Problems in Engineering*, vol. 2022, 10 pages, 2022.

[37] Q. Shao, Z. Zhang, Y. Zhu, and A. Zhou, "Survey of enterprise blockchains," *Journal of Software*, vol. 30, no. 9, pp. 2571–2592, 2019.

[38] W. Yang, E. Aghasian, S. Garg, D. Herbert, L. Disiuta, and B. Kang, "A survey on blockchain-based Internet service architecture: requirements, challenges, trends, and future," *IEEE Access*, vol. 7, pp. 75845–75872, 2019.

[39] X. Han, Y. Yuan, and F. Wang, "Security problems on blockchain: the state of the art and future trends," *Acta Automatica Sinica*, vol. 45, no. 1, p. 206, 2019.

[40] M. S. Ali, M. Vecchio, M. Pincheira, K. Dolui, F. Antonelli, and M. H. Rehmani, "Applications of blockchains in the Internet of Things: a comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1676–1717, 2019.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 8021267, 14 pages
https://doi.org/10.1155/2022/8021267

Research Article

Revocable One-Time Ring Signature from Pairings

Xu Han, [1,2] **Dawei Zhang**, [1,2] **Zongmin Huang**, [1,2] **Shuang Yao**, [1,2] **and Zuodong Wu** [1,2]

[1] *School of Computer and Information Technology, Beijing Jiaotong University, Beijing, Beijing 100044, China*
[2] *Beijing Key Laboratory of Security and Privacy in Intelligent Transportation, Beijing, Beijing 100044, China*

Correspondence should be addressed to Dawei Zhang; dwzhang@bjtu.edu.cn

Received 5 April 2022; Accepted 22 July 2022; Published 18 August 2022

Academic Editor: Yin Zhang

Ring signature is an anonymous signature that allows a person to sign a message on behalf of a self-formed group while concealing the identification of the signer. However, due to its anonymity and unlinkability, malicious or irresponsible signers can easily attack the signature without any responsibility in some scenarios. In this paper, we propose a novel revocable one-time ring signature (roRS) scheme from bilinear pairings, which introduces linkability and mandatory revocability into ring signature. In particular, linkability can resist the double-signing attack and mandatory revocability guarantees that a revocation authority can identify the actual signer when a suspicious signer appears in any situation. The computational complexity of pairing computations is constant, and the time of the revocation phase is more efficient than previous schemes. Furthermore, our scheme is provable secure in the random oracle model, using DL, CDH, and DBDH assumptions.

1. Introduction

Ring signature, initially proposed in 2001 by Rivest [1], is a variant of digital signature, which can prove that one among a set of spontaneous parties has already signed a message, without revealing the actual signer. And these spontaneous parties compose a particular set called a "ring." More specifically, a ring member can sign the signature without reveal any identity information, namely, a verifier who uses ring members' public keys only know whether the signature is true or not and cannot find out the actual signer, and the verifier has no clue who the signer is. As shown in Figure 1, first step, the actual signer uses private key sk_j and randomly chooses $r_j \in \mathbb{Z}_p^*$ to generate L_j by using the commit function $C(sk_j, r_j)$; then, the signer uses L_j to compute the (j+1)-th challenge c_{j+1} by hash function H; signer randomly picks a response z_{j+1} and the (j+1)-th user's public key pk_{j+1} to reconstruct the L_{j+1} by the verify function Ver and generates c_{j+2} by hash function H; then, the ring is formed sequentially; finally, the signer uses sk_j, c_j, r_j to compute z_j by the response function Z. In the whole process of generating a ring, we only need the actual signer's private key sk_j and a set of users' public keys which contains pk_j. In the view of the actual signer, users except the actual signer can be seen as decoys. When the verifier does the verifications, he/she does not know any information about the knowledge of the actual signer. As for security, ring signature not only provides regular properties, such as correctness and unforgeability which any signature schemes must possess, but also has the special feature, anonymity. Correctness requires a ring member who represents a ring to sign a message and unforgeability demands that an efficient adversary cannot forge a signature on behalf of a ring which the adversary knows nothing about one secret key of ring members. As for anonymity, it allows that ring signature schemes cannot leak any information about the identity of the actual signer, that is, no one can tell which key was used to produce a signature.

As an extension of ring signature, one-time ring signature can be known as linkable ring signature; the slight difference between these two kinds of ring signature is that

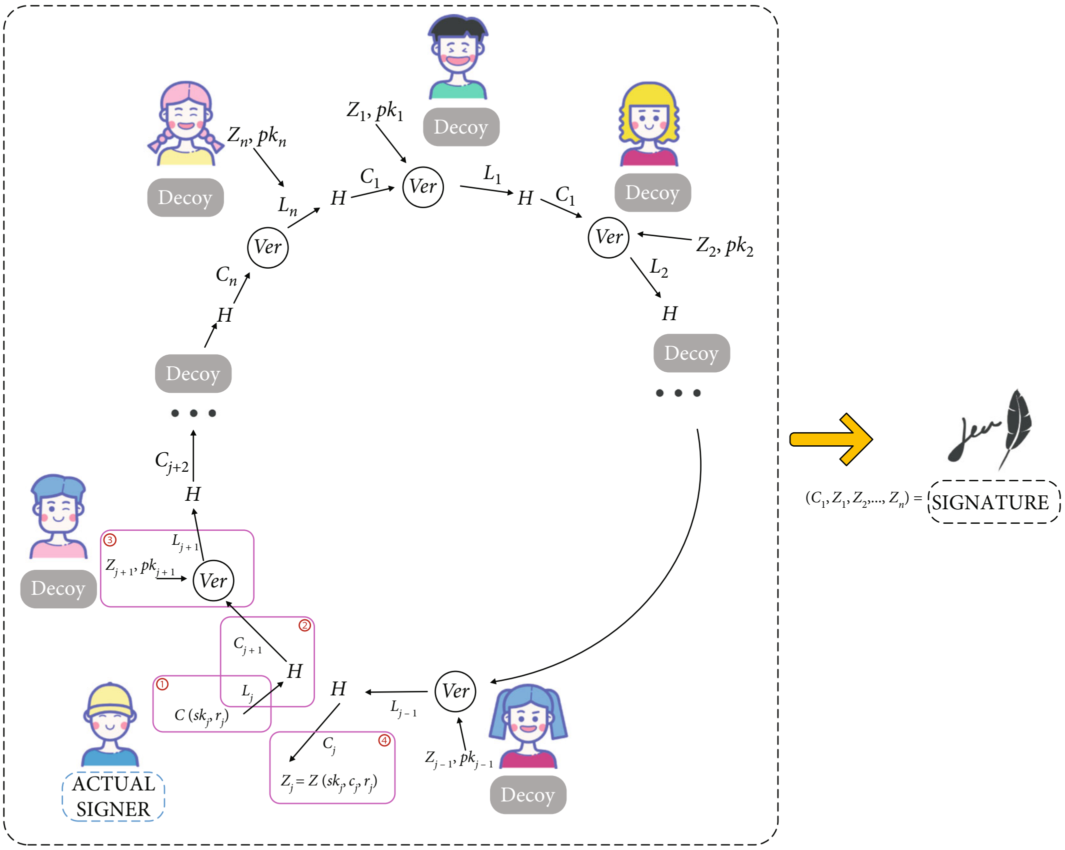


Figure 1: Ring signature.

signers in one-time ring signature use one-time key images to sign a signature, while signers in linkable ring signature take static ones. So we just need to introduce linkable ring signature in this section. Liu et al. [2] first put forward the concept of linkable ring signature (lRS). Beside the regular properties of ring signature, lRS provides two more special properties: non-slanderability and linkability. Non-slanderability guarantees that a ring member should not be entrapped that he has signed twice. Linkability requires that two signatures with the same ring on random messages must be linked if signed by the identical signer; thus, it can defeat the double-signing (double-spending) attack. This property is suitable in some practical applications; one scenario is on detecting double-voting in e-voting [3] systems. At the beginning of e-voting systems, we use ring signature schemes to implement the systems for its spontaneity, and there is no registration phase. The only requirement is that everyone has a public key pair which is considered as a well-known assumption in a ring. However, using classic ring signature as e-voting has a main problem. Anyone can vote more than once without being detected as ring signature schemes are unlinkable and anonymous. Thus, using lRS can solve this problem as double voting (double signing) can be detected easily, and anyone only can vote once in the system. Beside e-voting systems, lRS can also apply in other actual scenarios, such as ad hoc network authentication [4], blockchain-based applications [5, 6], and cryptocurrencies (Monero [7]). But in some actual transactions based on ring signature, when a ring signer has committed an offence, such as money laundering, online extortion, and terrorist financing, the authority needs to find out who is the actual signer among the ring members. Since lRS cannot let the actual signer be identified, the revocability of ring signature becomes necessary. Revocability requires that the authority can revoke the anonymity of ring signers when a suspicious signer does a transaction.

In order to solve the above problem, we propose a novel revocable one-time ring signature (roRS) scheme. Our scheme can be applied in some blockchain transactions. For example, by using the functionality of ring signature, Monero [7] protects the privacy of the signer's identity and provides autonomous mixing in transactions, but the unconditional anonymity of the ring signature makes difficult to regulate transactions for authorities. As for our scheme, a verifier can prevent the user's double-spending behavior according to the linkability during the transactions, and a revocation authority can recover the public key of the transaction user by using the revocability of our scheme, thus restoring the transaction user's identity.

1.1. Related Work. Rivest et al. [1] in 2001 proposed the first ring signature scheme, using trapdoor permutation based on the discrete logarithm problem assumption. Hereafter, many of schemes [8–11] followed this idea, but using different techniques has come out. For instance, Boneh et al. [12] first proposed a ring signature using bilinear pairings based on co-computational Diffie-Hellman (co-CDH) assumption. Cayrel et al. [13] presented a lattice-based ring signature scheme with modifying Melchor's code-based method [14] to make the short integer solution (SIS) problem as a security assumption. As for proving the membership problem in ring signature, most of these schemes use non-interactive witness indistinguishable (NIWI) proofs [15] or dynamic accumulator [16] to be more efficient, and readers who want to learn more refer to [17–21].

Then, Liu et al. [2] first proposed a linkable ring signature (lRS) scheme in 2004. This scheme inherits the anonymity of ring signature and provides a new property called linkability to resist double-spending attempts in real transactions, and it is proven to be secured in the random oracle model. Tsang et al. [22] constructed the first separable linkable ring signature scheme with introducing the security notions of accusatory linkability and non-slanderability. Liu and Wong [23] enhanced the security model for adapting to new attacking scenarios, and proposed two polynomial-structured lRS schemes based on zero knowledge proof. In 2007, Zheng [24] designed an lRS scheme based on linear feedback shift register under discrete logarithm assumption. In 2014, Liu et al. [25] put forward the first unconditional anonymous lRS scheme and provide mandatory linkability. Recently, Noether [26] proposed a dual lRS scheme which key images are tied to both output one-time public keys in a dual, and this can be considered using in non-interactive refund transactions in Monero. Tang et al. [27] presented an identity-based lRS scheme by employing trapdoor generation and rejection sampling as the basic building tool under the SIS problem on NTRU lattice. Hu et al. [28] designed a lattice-based lRS scheme under the well-studied standard lattice assumptions (SIS and LWE) in the standard model.

Revocable ring signature, also called traceable ring signature, is presented to reduce and even revoke the anonymity of the signers mainly. In 2007, Liu et al. [29] first proposed a revocable ring signature that authorities can mandatory revoke the anonymity of the actual signer when authorities need in some scenarios, but this scheme cannot provide linkability against the double-signing attack. Fujisaki et al. [30] put forward a traceable ring signature which only can trace a signer who was double-signing, that is, the traceability is not mandatory. The similar constructions can be found in [17–21]. In [31], Fujisaki presented the first secure traceable ring signature scheme without random oracles in the common reference string model, and the signature size grows linearly with $\sqrt{n}$ where n is the number of users in the ring. Au et al. [32] adapted traceable ring signature to the identify-based setting with constant signature size and enhanced privacy. Recently, Feng et al. [33] designed a logarithmic-size traceable ring signature scheme from lattices which proved to be secure in the quantum random oracle model.

1.2. Motivation and Contributions. In this section, to be more concrete, we summarize that the contribution of our paper is as follows:

(i) We present a novel revocable one-time ring signature scheme and define a perfect security model which provides the security properties: unforgeability, anonymity, linkability, non-slanderability, and revocability. And revocability of our scheme is mandatory

(ii) We show that our scheme is provable secure in the random oracle model under the assumptions that discrete logarithm (DL) problem, computational Diffie-Hellman (CDH) problem, and decisional bilinear Diffie-Hellman (DBDH) problem are intractable

(iii) We compare the efficiency of our scheme and previous schemes. Our scheme requires 4 times pairing computations which is independent of the size in the ring. Besides, the computational complexity in revocation part is more efficient than previous ones, and it only requires one scalar multiplication computation and one additional computation

2. Preliminaries

In this section, we introduce bilinear pairing and complex assumptions. They are utilized in the construction and provable security for our scheme. The notations used throughout the paper are described in Table 1.

2.1. Bilinear Pairing. Let $\mathbb{G}_1$ and $\mathbb{G}_T$ be cyclic groups of a large prime order p. We write $\mathbb{G}_1$ additively and $\mathbb{G}_T$ multiplicatively. We assume that the discrete logarithm problems in $\mathbb{G}_1$ and $\mathbb{G}_T$ are intractable.

Let G be a bilinear group generator that, on input of a security parameter κ, outputs a description of bilinear groups $(\mathbb{G}_1,\mathbb{G}_T,e,P)$ such that $\mathbb{G}_1$ and $\mathbb{G}_T$ are cyclic groups of prime order p, P is a generator of $\mathbb{G}_1$, and a map $e : \mathbb{G}_1 \times \mathbb{G}_1 \longrightarrow \mathbb{G}_T$ satisfies the following properties:

(i) Bilinear: $\forall P \in \mathbb{G}_1$ and $a, b \in \mathbb{Z}_p^*$: $e(aP, bP) = e(P, P)^{ab}$

(ii) Non-degenerate: There exists $\forall P \in \mathbb{G}_1$, such that $e(P, P) \neq 1$

(iii) Computability: The map $e(P, P)$ is efficiently computability for any $P \in \mathbb{G}_1$

2.2. Complexity Assumptions

Definition 1 Discrete logarithm (DL) assumption. We say that the DL assumption holds if for any polynomial-time adversary $\mathscr{A}$, the following advantage ε^{DL} is negligible function in κ:

$$\varepsilon^{DL} \coloneqq \boldsymbol{Pr}\left[\mathscr{A}_{DL}(P, aP) = a : \begin{array}{c} (p, \mathbb{G}_1) \longleftarrow G(\kappa)\,; \\ P \longleftarrow \mathbb{G}_1\,; \\ a \longleftarrow \mathbb{Z}_p^*\,; \end{array}\right] = \mathrm{negl}(\kappa). \tag{1}$$

Table 1: The symbol description.

Symbol	Description
κ	A security parameter
p	A large prime number
$\mathbb{Z}_p^*$	The set consisting of positive integers less than p
$\mathbb{G}_1$,P	The additive group with p order and a generator
$\mathbb{G}_T$	The multiplicative group with prime number p order
e	A bilinear pairing, where $e : \mathbb{G}_1 \times \mathbb{G}_1 \longrightarrow \mathbb{G}_T$
pk_{revoke}	A revocation authority's public key, where $pk_{revoke} = \hat{y} = sk_{revoke}P = \hat{x}P$
H	A cryptographic hash function, where $H : \{0,1\}^* \longrightarrow \mathbb{Z}_p^*$
H_1	A deterministic hash function, where $H_1 : E(\mathbb{G}_1) \longrightarrow E(\mathbb{G}_1)$
H_2	A cryptographic hash function, where $H_2 : \mathbb{G}_1 \longrightarrow \mathbb{Z}_p^*$

Table 2: Comparison of ring signature schemes.

Scheme	Signature size	Sign	Verify	Assumption	Security model
Zhang et al. [35]	$n\|\mathbb{G}_1\|$	$(2n-1)\mathcal{S}m$	$(n+1)\mathcal{P}air$	q_s-CAA	ROM
Schäge et al. [36]	$(n+1)\|\mathbb{G}_1\|$	$(n+2)Sm + nadd$	$(n+2)Pair + nadd$	CDH	StanM
Liu et al. [25]	$(2n+3)\|\mathbf{G}_T\|$	$(3n+4)Sm + (4n+3)add$	$(2n+5)Pair + (3n+1)add$	CDH	StanM
roRS	$n\|G_1\| + n\left\|Z_p^*\right\|$	$(n+1)Sm + Pair$	$4Pair + nSm$	DL, DBD, CDH	ROM

Table 3: The notions in comparison.

Notion	Description
$\left\|\mathbb{Z}_p^*\right\|$	The length of the elements in $\mathbb{Z}_p^*$
$\|\mathbb{G}_1\|$	The size of the underlying group in $\mathbb{G}_1$
$\|\mathbb{G}_2\|$	The size of the underlying group in $\mathbb{G}_2$
$\|\mathbb{G}_T\|$	The size of the underlying group in $\mathbb{G}_T$
add	The time required for an addition computation
δm	The time required for a scalar multiplication computation
$\mathcal{P}air$	The time required for a pairing computation
ROM	The abbreviation for random Oracle model
StanM	The abbreviation for standard model

Table 4: Comparison in [29, 30] and roRS.

Scheme	Signature size	Revoke	Assumption	Linkability	Mandatory Revocability
Liu et al. [29]	$(2n+2)\left\|Z_p^*\right\|$	$nPair$	DBDH	✘	√
Fujisaki et al. [30]	$(2n+1)\left\|Z_p^*\right\|^{``}$	$2nadd + 2n\mathcal{S}m$	Dl, DDH	√	✘
roRS	$n\|G_1\| + n\left\|Z_p^*\right\|$	$\mathcal{S}m + add$	Dl, DBDH, CDH	√	√

Definition 2 Computational Diffie-Hellman (CDH) assumption. Let G be a group generator that, on input of a security parameter κ, outputs a cyclic group. We say that the CDH assumption holds if for any polynomial-time adversary $\mathscr{A}$, the following advantage ε^{CDH} is negligible function in κ:

$$\varepsilon^{CDH} \coloneqq \boldsymbol{Pr}\left[\mathscr{A}_{CDH}(P, aP, bP) = abP : \begin{array}{c} (p, \mathbb{G}_1) \longleftarrow G(\kappa) \,; \\ P \longleftarrow \mathbb{G}_1 \,; \\ a, b \longleftarrow \mathbb{Z}_p^* \,; \end{array}\right] = \operatorname{negl}(\kappa). \tag{2}$$

Definition 3 Decisional bilinear Diffie-Hellman (DBDH) assumption. Let $e : \mathbb{G}_1 \times \mathbb{G}_1 \longrightarrow \mathbb{G}_T$ be a bilinear pairing, $P \in \mathbb{G}_1$. For $Z \in \mathbb{G}_T$, given the tuples (P, aP, bP, cP, Z), we say that the DBDH assumption holds if for any polynomial-time adversary $\mathscr{A}$, the following advantage ε^{DBDH} is negligible function in κ:

$$\begin{aligned} \varepsilon^{DBDH} \coloneqq \Big| \Pr\Big[\mathscr{A}_{DBDH}\Big(P, aP, bP, cP, e(P, P)^{abc}\Big) = 1\Big] \Big| \\ - \Pr[\mathscr{A}_{DBDH}(P, aP, bP, cP, Z) = 1] = \operatorname{negl}(\kappa). \end{aligned} \tag{3}$$

3. Security Model

In this section, we give the security model and the security notions of our revocable one-time ring signature.

3.1. Definition of Revocable One-Time Ring Signature

3.1.1. Revocable One-time Ring Signature. Revocable one-time ring signature (roRS) scheme is the tuples (Setup, KeyGen, Sign, Verify, Link, and Revoke).

(i) $pp \longleftarrow Setup(\kappa)$: The setup algorithm is a probabilistic polynomial time algorithm which takes as input a security parameter $\kappa \in N$ and outputs a set of public parameters pp

(ii) $(sk_i, pk_i) \longleftarrow KeyGen(pp)$: The key generation algorithm is a probabilistic polynomial time algorithm which takes as input public parameters pp and outputs a private/public key pair (sk_i, pk_i). Respectively, we denote SK and PK as the domain of possible private keys and possible public keys

(iii) $\sigma \longleftarrow \mathrm{Sign}(I, sk, Y, pk_{revoke}, M)$: The signing algorithm is a probabilistic polynomial time algorithm which takes as input a key image I, a private key sk, a message M, a revocation authority's public key $pk_{revoke} \in PK$, and a list Y of public keys in PK which includes the one corresponding to sk and produces a signature σ

(iv) $accept/reject \longleftarrow Verify(I, Y, M, pk_{revoke}, \sigma)$: The signature verification algorithm is a probabilistic polynomial time algorithm which takes as input the key image I, a set Y of public keys in PK, a revocation authority's public key $pk_{revoke} \in PK$, a message M, and a signature σ and returns accept or reject

(v) $linked/unlinked \longleftarrow Link(I_1, I_2, Y_1, Y_2, M_1, M_2, \sigma_1, \sigma_2)$: The linking algorithm which takes as input key images I_1, I_2, a set Y_1 of public keys in PK, and a set Y_2 of public keys in PK, messages M_1 and M_2, and signatures σ_1 and σ_2, such that $Verify(I_1, Y_1, M_1, \sigma_1) = accept$ and $Verify(I_2, Y_2, M_2, \sigma_2) = accept$ and returns linked or unlinked

(vi) $pk \longleftarrow Revoke(Y, \sigma, sk_{revoke})$: The revoking algorithm is a probabilistic polynomial time algorithm which takes as input a set Y of public keys in PK, a valid signature σ, and a secret key sk_{revoke} of the revocation authority and returns a public key pk in Y

3.1.2. Correctness. A revocable one-time ring signature scheme should satisfy the following:

(i) Verification correctness: A signature signed by honest signers is verified to be valid as follows:

$$\Pr\left[Verify(I, Y, M, pk_{revoke}, \sigma^*) = accept : \begin{array}{c} pp \longleftarrow Setup(\kappa) \,; \\ (sk_i, pk_i) \longleftarrow KeyGen(pp) \,; \\ \sigma \longleftarrow \mathrm{Sign}(I, sk, Y, pk_{revoke}, M) \,; \end{array}\right] = 1. \tag{4}$$

(ii) Linking correctness: Two signatures with the same event description generated by the same secret key of the identical signer must be linkable

(iii) Revocation Correctness: The revocation authority can reveal an honest signer's public key with overwhelming probability

3.2. Notions of Security. Security of our roRS scheme has five aspects: unforgeability, anonymity, linkability, non-slanderability, and revocability.

Formally, we capture attack behaviors as adversarial queries to oracles implemented by a challenger S. We provide adversary $\mathscr{A}$ the following oracles.

(i) JO (joining oracle). O_{join}: $pk_i \longleftarrow JO(\perp)$. $\mathscr{A}$ queries this oracle for adding a new user to the system. S keeps track of this type of queries by maintaining a list T_{join}, which is initially empty. Upon receiving a fresh query, S responds as below: picks random public parameters pp, runs $(sk_i, pk_i) \longleftarrow KeyGen(pp)$ to obtain (sk_i, pk_i). S records (sk_i, pk_i) in T_{join}, and then returns $pk_i (pk_i \in PK)$ to $\mathscr{A}$. This type of oracle captures $\mathscr{A}$ can observe the public keys of honest users in the system

(ii) CO (corruption oracle). $O_{corrupt}$: $sk_i \longleftarrow CO(pk_i)$. $\mathscr{A}$ queries this oracle with a public key $pk_i \in PK$ in

T_{join}. S keeps track of this type of queries by maintaining a list $T_{corrupt}$, which is initially empty. Upon receiving a fresh query, S records pk_i in $T_{corrupt}$. S returns the associated $sk_i \in SK$ to $\mathscr{A}$ and moves this entry to $T_{corrupt}$. This oracle captures $\mathscr{A}$ can corrupt some honest users and return private key sk_i

(iii) SO (signing oracle). O_{sign}: $\sigma' \longleftarrow SO(I, Y, pk_{revoke}, M)$. $\mathscr{A}$ queries this oracle with (I, Y, pk_{revoke}, M) (a key image I, a set Y of public keys, a revocation authority's public key $pk_{revoke} \in PK$, and a message M) and subjects to the restriction that $pk_s \in T_{join}(s = 1, \cdots, n)$. S keeps track of this type of queries by maintaining a list T_{sign}, which is initially empty. Upon receiving a fresh query, S responds as below: S runs $\sigma' \longleftarrow \text{Sign}(I, sk, pk_{revoke}, Y, M)$, records σ' on the list T_{sign}, and then sends σ' to $\mathscr{A}$. This oracle captures $\mathscr{A}$ can generate a signature itself

3.2.1. Unforgeability. We define unforgeability via the following security experiment $Expt_{\mathscr{A}}^{unforge}$ between $\mathscr{A}$ and S:

$$\mathscr{A}dv_{\mathscr{A}}^{unforge}(\kappa) = \Pr\left[\begin{array}{ccc} Verify(I, Y, pk_{revoke}, M, \sigma^*) = accept \wedge & & pp \longleftarrow Setup(\kappa)\,; \\ & : & \\ pk_i \in T_{join} \wedge \sigma^* \notin T_{\text{sign}} & & (I, Y, M, \sigma^*) \longleftarrow \mathscr{A}^{JO,CO,SO}(pp)\,; \end{array}\right]. \quad (5)$$

Here, $n \in \mathbb{N}$, $Y = \{pk_1, pk_2, \cdots, pk_n\} \in PK$, and I is a key image. pk_{revoke} is a revocation authority's public key, and all public keys are in T_{join}. No public keys are in $T_{corrupt}$. σ^* is not in T_{sign}. $\mathscr{A}dv_{\mathscr{A}}^{unforge}(\kappa)$ is the successful probability of adversary $\mathscr{A}$ who wins the unforgeability security experiment.

The process of unforgeability security experiment $Expt_{\mathscr{A}}^{unforge}$ is briefly described as follow:

(1) $\mathscr{A}$ runs $Setup(\kappa)$ with security parameter κ and sends the public parameter pp to $\mathscr{A}$

(2) $\mathscr{A}$ is allowed to make queries to O_{join}, $O_{corrupt}$, and O_{sign} according to any adaptive strategy

(3) $\mathscr{A}$ outputs a forgery signature σ^*

The conditions that $\mathscr{A}$ wins the experiment are as follows:

(1) All public keys are outputs of JO

(2) $Verify(I, Y, pk_{revoke}, M, \sigma^*)$=accept, and σ is not the output of SO

(3) No public keys have been queried to CO

Our roRS satisfies unforgeability if no PPT adversary has non-negligible advantage in the above experiment.

Definition 4 Unforgeability. A revocable one-time ring signature scheme is unforgeable, if for every probabilistic polynomial-time adversary $\mathscr{A}$, the advantage $\mathscr{A}dv_{\mathscr{A}}^{unforge}(\kappa)$ is negligible in κ.

3.2.2. Anonymity. We define anonymity via the following security experiment $Expt_{\mathscr{A}}^{anonymous}$ between $\mathscr{A}$ and S:

$$\mathscr{A}dv_{\mathscr{A}}^{anonymous}(\kappa) = \left| \Pr\left[s = s' : \begin{array}{c} pp \longleftarrow Setup(\kappa)\,; \\ (I, Y, M) \longleftarrow \mathscr{A}^{JO}(pp)\,; \\ s \overset{R}{\longleftarrow} \{1, \cdots, n\}\,; \\ \sigma_s \longleftarrow \text{Sign}(I, sk, Y, pk_{revoke}, M)\,; \\ s' \longleftarrow \mathscr{A}(\sigma_s)\,; \end{array} \right] - \frac{1}{n} \right|. \quad (6)$$

Here, $n \in \mathbb{N}$, $Y = \{pk_1, pk_2, \cdots, pk_n\} \in PK$, and I is a key image. pk_{revoke} is a revocation authority's public key, and pk_i is chosen by $\mathscr{A}$ in T_{join}. sk_s is a corresponding private key of pk_s. $\mathscr{A}dv_{\mathscr{A}}^{anonymous}(\kappa)$ is the successful probability of adversary $\mathscr{A}$ who wins the anonymity security experiment.

The process of anonymity security experiment $Expt_{\mathscr{A}}^{anonymous}$ is briefly described as follow:

(1) $\mathscr{A}$ runs $Setup(\kappa)$ with security parameter κ and sends the public parameter pp to $\mathscr{A}$

(2) $\mathscr{A}$ is allowed to make queries to O_{join} according to any adaptive strategy

(3) $\mathscr{A}$ gives S a set Y of public keys in PK such that all of the public keys in Y are query outputs of JO, a key image I and a message M. Parse the set Y as $\{pk_1, pk_2, \cdots, pk_n\}$. S randomly picks $s \in \{1, \cdots, n\}$ and computes $\sigma_s = \text{Sign}(I, Y, sk_s, M)$, where sk_s is a corresponding private key of pk_s. σ_s is given to $\mathscr{A}$

(4) $\mathscr{A}$ outputs a guess $s' \in \{1, \cdots, n\}$

So our roRS satisfies anonymity if no PPT adversary has non-negligible advantage in the above experiment.

Definition 5 Anonymity. A revocable one-time ring signature scheme is anonymous, if for every probabilistic polynomial-time adversary $\mathcal{A}$, the advantage $Adv^{anonymous}\mathcal{A}(\kappa)$ is negligible in κ.

3.2.3. Linkability. We define linkability via the following security experiment $Expt_{\mathcal{A}}^{linkable}$ between $\mathcal{A}$ and S:

$$\mathcal{A}dv_{\mathcal{A}}^{linkable}(\kappa) = \Pr\left[\begin{array}{c} Verify(I, Y, pk_{revoke}, M, \sigma^*) = accept\wedge \\ pk_i \in T_{join}\wedge\sigma_i \notin T_{\text{sign}}\wedge \\ Link(\sigma_1, \sigma_2) = unlinked, i = 1, 2 \end{array} : \begin{array}{c} pp \longleftarrow Setup(\kappa)\,; \\ (I, Y, M, \sigma^*) \longleftarrow \mathcal{A}^{JO,CO,SO}(pp)\,; \\ \end{array}\right]. \tag{7}$$

Here, $n_i \in \mathbb{N}$, $i = 1, 2$; $Y_i = \{pk_1, pk_2, \cdots, pk_{n_i}\} \in PK$, a message M and signature σ_i, $i = 1, 2$; I is a key image, and pk_{revoke} is a revocation authority's public key. All pk_i chosen by $\mathcal{A}$ are in T_{join}. σ_i are not in T_{sign}, CO has been queried less than 2 times. $\mathcal{A}dv_{\mathcal{A}}^{linkable}(\kappa)$ is the successful probability of adversary $\mathcal{A}$ who wins the linkability security experiment.

The process of linkability security experiment $Expt_{\mathcal{A}}^{linkable}$ is briefly described as follows:

(1) $\mathcal{A}$ runs $Setup(\kappa)$ with security parameter κ and sends the public parameter pp to $\mathcal{A}$

(2) $\mathcal{A}$ is allowed to make queries to O_{join}, $O_{corrupt}$, and O_{sign} according to any adaptive strategy

(3) $\mathcal{A}$ outputs a forgery signature σ

The conditions that $\mathcal{A}$ wins the experiment are as follows:

(1) All public keys are outputs of JO

(2) $Verify(I, Y, pk_{revoke}, M, \sigma)$=accept, and σ is not the output of SO

(3) $\mathcal{A}$ can only at most queried 1 time and at most have one user's private key

(4) $Link(\sigma_1, \sigma_2)$= unlinked

So our roRS satisfies linkability if no PPT adversary has non-negligible advantage in the above experiment.

Definition 6 Linkability. A revocable one-time ring signature scheme is linkable, if for every probabilistic polynomial-time adversary $\mathcal{A}$, the advantage $\mathcal{A}dv_{\mathcal{A}}^{linkable}(\kappa)$ is negligible in κ.

3.2.4. Non-slanderability. We define non-slanderability via the following security experiment $Expt_{\mathcal{A}}^{\text{non}-slanderous}$ between $\mathcal{A}$ and S:

$$\mathcal{A}dv_{\mathcal{A}}^{\text{non}-slanderous}(\kappa) = \Pr\left[\begin{array}{c} Verify(I, Y^*, pk_{revoke}, M^*, \sigma^*) = accept\wedge \\ \sigma^* \neq \sigma'\wedge\sigma^* \notin T_{\text{sign}}\wedge \\ pk_s \notin T_{corrupt}\wedge pk_s \notin T_{\text{sign}}\wedge \\ Link\left(\sigma^*, \sigma'\right) = linked \end{array} : \begin{array}{c} pp \longleftarrow Setup(\kappa)\,; \\ (I, Y, M, pk_s) \longleftarrow \mathcal{A}^{JO,CO,SO}(pp)\,; \\ \sigma' \longleftarrow \text{Sign}(sk_s, Y, pk_{revoke}, M)\,; \\ (I, Y^*, M^*, \sigma^*) \longleftarrow \mathcal{A}^{JO,CO,SO}\left(\sigma'\right); \end{array}\right]. \tag{8}$$

Here, $Y, Y^* \in PK$, message M, M^*, and signature σ', σ^*; pk_{revoke} is a revocation authority's public key, and I is a key image. $\mathcal{A}$ is subject to the restriction that pk_s chosen by $\mathcal{A}$ is not allowed to be either in $T_{corrupt}$ or T_{sign}. All of the public keys in Y^* and Y are in T_{join}. $\sigma^* \neq \sigma'$ and σ^* are not in T_{sign}. $\mathcal{A}dv_{\mathcal{A}}^{\text{non}-slanderous}(\kappa)$ is the successful probability of adversary $\mathcal{A}$ who wins the non-slanderability security experiment.

The process of non-slanderability security experiment $Expt_{\mathcal{A}}^{\text{non}-slanderous}$ is briefly described as follows:

(1) $\mathcal{A}$ runs $Setup(\kappa)$ with security parameter κ and sends the public parameter pp to $\mathcal{A}$

(2) $\mathcal{A}$ is allowed to make queries to O_{join}, $O_{corrupt}$, and $O_{\boldsymbol{sign}}$ according to any adaptive strategy

(3) $\mathcal{A}$ outputs a forgery signature σ^*

The conditions that $\mathcal{A}$ wins the experiment are as follows:

(1) $Verify(I, Y^*, pk_{revoke}, M^*, \sigma^*)$=accept

(2) σ^* is not an output of SO

(3) pk_s has not been queried to CO

(4) All public keys are in Y^*; Y are query outputs of JO

(5) $Link(\sigma', \sigma^*)$=linked

So our roRS satisfies non-slanderability if no PPT adversary has non-negligible advantage in the above experiment.

Definition 7 Non-slanderability. A revocable one-time ring signature scheme is non-slanderous, if for every probabilistic polynomial-time adversary $\mathscr{A}$, the advantage $\mathscr{A}dv_{\mathscr{A}}^{\text{non}-slanderous}(\kappa)$ is negligible in κ.

3.2.5. Revocability. We define revocability via the following security experiment $Expt_{\mathscr{A}}^{revocable}$ between $\mathscr{A}$ and S:

$$\mathscr{A}dv_{\mathscr{A}}^{revocable}(\kappa) = \Pr\left[\begin{array}{ccc} Verify(I, Y, pk_{revoke}, M, \sigma) = accept\wedge & & pp \longleftarrow Setup(\kappa)\,; \\ pk_i \in T_{join}\wedge\sigma \notin T_{\boldsymbol{sign}}\wedge & : & sk_s \longleftarrow \mathscr{A}^{CO}(pp)\,; \\ pk_j = Revoke(Y, \sigma, sk_{revoke}), j \neq s & & (I, Y, M, \sigma) \longleftarrow \mathscr{A}^{JO,CO,SO}(pp)\,; \end{array}\right]. \tag{9}$$

Here, $n \in \mathbb{N}$, $Y = \{pk_1, pk_2,\cdots,pk_n\} \in PK$, signature σ, and pk_{revoke} is a revocation authority's public key, and I is a key image. All pk_i chosen by $\mathscr{A}$ are in T_{join}. σ are not in $T_{\boldsymbol{sign}}$, CO has been queried less than 2 times, that is, $\mathscr{A}$ can only obtain at most one private key denotes as sk_s, and pk_{revoke} is the corresponding public key of sk_{revoke}. $\mathscr{A}dv^{revocable}\mathscr{A}(\kappa)$ is the successful probability of adversary $\mathscr{A}$ who wins the revocability security experiment.

The process of revocability security experiment $Expt_{\mathscr{A}}^{revocable}$ is briefly described as follows:

(1) $\mathscr{A}$ runs $Setup(\kappa)$ with security parameter κ and sends the public parameter pp to $\mathscr{A}$

(2) $\mathscr{A}$ is allowed to make queries to O_{join}, $O_{corrupt}$, and $O_{\boldsymbol{sign}}$ according to any adaptive strategy

(3) $\mathscr{A}$ outputs a forgery signature σ

The conditions that $\mathscr{A}$ wins the experiment are as follows:

(1) $Verify(I, Y, pk_{revoke}, M, \sigma)$=accept

(2) σ^* is not an output of SO

(3) All public keys are query outputs of JO

(4) CO has been queried less than 2 times, that is, $\mathscr{A}$ can only obtain at most one private key denotes as sk_s

(5) $y_j = Revoke(Y, \sigma, sk_{revoke})$ for $j \neq s$

So our roRS satisfies revocability if no PPT adversary has non-negligible advantage in the above experiment.

Definition 8 Revocability. A revocable one-time ring signature scheme is revocable, if for every probabilistic polynomial-time adversary $\mathscr{A}$, the advantage $\mathscr{A}dv_{\mathscr{A}}^{revocable}(\kappa)$ is negligible in κ.

4. Construction

In this section, we propose our new revocable one-time ring signature (roRS) scheme. Our scheme is constructed as follows:

Setup: Let $\mathbb{G}_1$ and $\mathbb{G}_T$ be two groups with prime order $p > 2^\kappa$; $\mathbb{G}_1$ is an additive group and $\mathbb{G}_T$ is a multiplicative group. P is the generator of $\mathbb{G}_1$ and a bilinear pairing $e : \mathbb{G}_1 \times \mathbb{G}_1 \longrightarrow \mathbb{G}_T$. Let $H : \{0,1\}^* \longrightarrow \mathbb{Z}_p^*$ and $H_2 : \mathbb{G}_1 \longrightarrow \mathbb{Z}_p^*$ be two cryptographic hash functions and $H_1 : E(\mathbb{G}_1) \longrightarrow E(\mathbb{G}_1)$ be a deterministic hash function. And the public parameters $params = (\mathbb{G}_1, \mathbb{G}_T, e, P, H, H_1, H_2, p)$..

KeyGen: A ring user i randomly picks $x_i \in \mathbb{Z}_p^*$ and computes $pk_i = x_i P \in \mathbb{G}_1$. So the user i has secret key and public key pair $(sk_i, pk_i) = (x_i, x_i P)$. The secret key sk_{revoke} of the revocation authority is $\widehat{x}$, and the corresponding public key pk_{revoke} is $\widehat{y} = \widehat{x}P$.

Sign: Let $Y = \{pk_1, pk_2, \cdots, pk_n\}$ be a set of users' public keys in the ring. So a ring user $s(1 \leq s \leq n)$ has his own key pair $(sk_s, pk_s) = (x_s, x_s P)$. Additionally, the user has given a message $M \in \{0,1\}^*$; then, the user s with the knowledge of x_s computes a signature of knowledge as follows:

(1) Compute the key image I: First, use a hash function with pk_s to make one signer only have the corresponding one key image.

$K = H_1(pk_s)$; then, compute the key image I with the signer's private key x_s and K,$I = x_s K$.

(2) Randomly choose $\omega \in \mathbb{Z}_p^*$, and compute: First make proof of knowledge for private key x_s, $B_1 = \omega P$ and then construct B_2 for the revoking phase and B_3 for the verification phase:

$$\begin{aligned} B_2 &= \omega\widehat{y} + pk_s, \\ B_3 &= e(\omega pk_s, K). \end{aligned} \tag{10}$$

(3) Randomly choose $c_i \in \mathbb{Z}_p^*$, where $i = \{1, 2, \cdots, n\}$; and randomly choose $z_i \in \mathbb{G}_1$, where $i = \{1, 2, \cdots, s-1, s+1, \cdots, n\}(i \neq s)$. Then, set the following transformations for all users in the ring:

$$C_i = \begin{cases} c_iP, & if\, i = s \\ c_iP + pk_sH_2(z_i), & if\, i \neq s \end{cases} \text{and}\, Z_i = \begin{cases} c_iK, & if\, i = s \\ c_iK + IH_2(z_i), & if\, i \neq s \end{cases}. \tag{11}$$

(4) Get the non-interactive challenge:

$$z = H(M, C_1, \cdots, C_n, Z_1, \cdots, Z_n). \tag{12}$$

(5) Randomly choose $\Phi_i \in \mathbb{G}_1$, where $i = \{1, 2, \cdots, s-1, s+1, \cdots, n\}$ and $i \neq s$. Then when $i \neq s$, make a hash function h_i to be a random number for verifiers. Compute $h_i = H(\Phi_i, Y, I, M, B_3)$, where $i \neq s$

(6) Then, when $i = s$, use notions above and compute:

Compute Φ_s as a random number:

$$\Phi_s = zP - \sum_{i=1, i\neq s}^{n} (h_ipk_i + \Phi_i) \in \mathbb{G}_1, \tag{13}$$

then make a hash function h_s as a random number for $i = s$:

$$h_s = H(\Phi_s, Y, I, M, B_3) \in \mathbb{Z}_p^*, \tag{14}$$

use private key x_s to make the response V when $i = s$:

$$V = (z + h_sx_s) \cdot B_1 \in \mathbb{G}_1. \tag{15}$$

(7) To close the ring, set $l_i = \begin{cases} c_s - x_sH_2(\Phi_s), & if\, i = s \\ c_i, & if\, i \neq s \end{cases}$

(8) Output the signature $\sigma = (l_1, \cdots, l_n, \Phi_1, \cdots, \Phi_n, V, I, B_1, B_2)$

Verify: Given the tuples (σ, Y, M), the verifier carries out the following steps:

(1) Compute $C_i' = l_iP + pk_sH_2(\Phi_i)$ and $Z_i' = l_iH_1(pk_s) + IH_2(\Phi_i)$, $B_3 = e(B_1, I)$, and $h_i = H(\Phi_i, Y, I, M, B_3) \in \mathbb{Z}_p^*$, where $i = \{1, 2, \cdots, n\}$

(2) Check whether $\sum_{i=1}^{n}(h_ipk_i + \Phi_i) = H(M, C_1', \cdots, C_n', Z_1', \cdots, Z_n') \cdot P$ and $e(P, V) = e(\sum_{i=1}^{n}[h_ipk_i + \Phi_i], B_1)$. If two equalities hold, accept the signature; otherwise, reject it

Link: On receive the tuples:

$$\begin{aligned} &(\sigma_1 = (\cdot, I_1), Y_1, M_1), \\ &(\sigma_2 = (\cdot, I_2), Y_2, M_2). \end{aligned} \tag{16}$$

The verifier checks if both σ_1 and σ_2 are two valid signatures. If yes, then outputs link if $I_1 = I_2$. Otherwise, reject.

Revoke: On receive the tuples (Y, σ, sk_{revoke}). The revocation authority first checks whether σ is a valid signature. If it holds, continue; otherwise, reject. In order to revoke the anonymity of the actual signer, the revocation authority makes as follows:

If there exists $y_s \in Y$, such that $y_s = B_2 - \hat{x}B_1$, where y_s is the public key of the actual signer.

5. Correctness Analysis

5.1. Verification Correctness. On correctness, using the bilinearity and nondegeneracy of the pairing e, a signature is correctly verified by the Verify algorithm as follows:

$$\begin{gathered} B_3 = e(B_1, I) = e(\omega P, x_sK) = e(\omega x_sP, K) = e(\omega pk_s, K), \\ e\left(\sum_{i=1}^{n}[h_i \cdot pk_i + \Phi_i], B_1\right) = e\left(\sum_{i=1, i\neq s}^{n}[h_i \cdot pk_i + \Phi_i] + h_s \cdot pk_s + \Phi_s, B_1\right) = e(z \cdot P + h_s \cdot pk_s, B_1) = e(P, V). \end{gathered} \tag{17}$$

5.1.1. Linking Correctness. On linking correctness, the signer computes the key image as follows:

$$I = x_iH_1(pk_i). \tag{18}$$

Therefore, the user can only compute the key image once with the same event description.

5.1.2. Revoking Correctness. On revoking correctness, the revocation authority can successfully identify the actual signer's public key as follows:

$$y_s = B_2 - \hat{x}B_1 = \omega\hat{y} + pk_s - \hat{x}\omega P = pk_s, \tag{19}$$

where $\hat{x}$ is the revocation authority's private key and y_s is the actual signer's public key.

6. Security Analysis

In this section, the security proofs of the proposed scheme are given.

Theorem 9 Unforgeability. *Our proposed scheme is unforgeable in the random oracle model with the CDH assumption.*

Proof. Suppose there exists a PPT adversary $\mathscr{A}$ with advantage ε, which means that $\mathscr{A}$ can forge a valid signature with probability ε. Then, we use a simulator S to solve the CDH problem. Let (P, aP, bP) be a given instance, where a, b is randomly picked in $\mathbb{Z}_p^*$ and $P \in \mathbb{G}_1$. Through the adversary $\mathscr{A}$, we use the simulator S which outputs the CDH solution abP. S randomly selects $j \in \{1, \cdots, n\}$ as the actual signer in the simulation. So the simulator S simulates the oracles by interacting with the adversary $\mathscr{A}$ in the nature way as follows:

(1) Setup: S sets $pk_j = aP$ and $B_1 = bP$. S randomly chooses $u \in \mathbb{Z}_p^*$ and sets $\hat{y} = uP$. $\mathscr{A}$ is given the public parameters $(\mathbb{G}_1, \mathbb{G}_T, p, e, P, H, H_1, H_2)$, and the public keys $Y = \{pk_1, pk_2, \cdots, pk_n\}$

(2) Queries: $\mathscr{A}$ queries the oracles H, H_1, H_2, JO, CO, and SO in this phase, and S sets several lists to record the queries and answers. These lists are initially empty

(i) H-queries: S maintains and checks a corresponding list L_H as $\mathscr{A}$ queries hash values. If an entry for the query is found in L_H, S returns the same answer to $\mathscr{A}$. Otherwise, S generates a random value as an answer to $\mathscr{A}$, and then the query and the answer are added in L_H

(ii) H_1-queries: S maintains a list L_{H_1}. When $\mathscr{A}$ issues a query $H_1(\alpha_i)$, S randomly chooses $f_1 \in \mathbb{G}_1$ and sets $H_1(\alpha_i) = f_1$ as the answer. The query and the answer then are stored in list L_{H_1}

(iii) H_2-queries: S maintains a list L_{H_2}. When $\mathscr{A}$ issues a query $H_1(\beta_i)$, S randomly chooses $f_2 \in \mathbb{Z}_p^*$ and sets $H_2(\beta_i) = f_2$ as the answer. The query and the answer then are stored in list L_{H_2}

(iv) JO-queries: When $\mathscr{A}$ queries JO at the jth query, S returns the corresponding public key $pk_j = aP$ to $\mathscr{A}$. When $\mathscr{A}$ queries JO at the ith query$(i \neq j)$, S randomly chooses $d_i \in \mathbb{Z}_p^*$ and returns d_iP as the corresponding public key. The query and the answer then are stored in list L_{JO}

(v) CO-queries: When $\mathscr{A}$ queries CO with the public key pk_i which is an output of JO, S first checks if $i = j$. If yes, S fails and stops. Otherwise, S returns the corresponding d_i as the corresponding secret key. The query and the answer then are stored in list L_{CO}

(vi) SO-queries: When $\mathscr{A}$ queries SO with a tuple $(event, M, Y, pk_j)$. If $pk_i \neq pk_j$, S outputs a signature σ by Sign algorithm. Otherwise, S maintains a list L_{sign} with a tuple $(event, pk_j, I)$ and performs as follows:

(1) S retrieves the list $L_{\boldsymbol{sign}}$, if (pk_j, I) is found, and then takes the value I. If not, randomly selects a new value $I \in \mathbb{G}_1$ and adds (pk_j, I) to the list L_{sign}

(2) S computes $B_1 = bP, B_2 = b(uP) + aP, B_3 = e(B_1, I)$

(3) S chooses a random $j \in \{1, \cdots, n\}$

(4) S chooses $l_i \in \mathbb{Z}_p^*$ and $\Phi_i \in_R \mathbb{G}_1$ randomly, where

$i = \{1, 2, \cdots, j-1, j+1, \cdots, n\}$ and $i \neq j$; and S computes $C_i = l_iP + f_2(aP)$ and $Z_i = l_if_1 + f_2I$ and then selects $h_i \in \mathbb{Z}_p^*$ in L_H such that $H(\Phi_i, M, Y, I, B_3) = h_i$

(5) S randomly picks $z, h_j \in \mathbb{Z}_p^*$ and computes

$$\Phi_j = z \cdot P - h_j \cdot pk_j + \left[\sum_{i=1, i \neq j}^{n} h_i pk_i + \Phi_i \right], \tag{20}$$

and stores the value $h_j = H(\Phi_j, M, Y, I, B_3)$ to L_H. If there has a collision with hash values in L_H, do Step (5) again until no collision happen

(6) S computes $V = zB_1$ and then outputs $\sigma = (l_1, \cdots, l_n, \Phi_1, \cdots, \Phi_n, V, I, B_1, B_2)$ as a response to $\mathscr{A}$

Since each response is independently and uniformly distributed, all responses in the simulation are as in the real attack. Besides, all responses to SO are valid, and the output $\sigma = (l_1, \cdots, l_n, \Phi_1, \cdots, \Phi_n, V, I, B_1, B_2)$ in signing query can be verified with the Verify algorithm. Therefore, from $\mathscr{A}$'s view, the simulation is indistinguishable from the real attack. Now, $\mathscr{A}$ has the tuple $(\sigma = (l_1, \cdots, l_n, \Phi_1, \cdots, \Phi_n, V, I, B_1, B_2), h_1, \cdots, h_n)$, then by using Forking Lemma for ring signature [34], S rewinds the same random tape to let $\mathscr{A}$ obtain another tuple $(\sigma' = (l_1, \cdots, l_n, \Phi_1, \cdots, \Phi_n, V', I, B_1, B_2), h_1', \cdots, h_n')$ such that $h_j \neq h_j'$ for all $i \in \{1, \cdots, n\}$ and $i \neq j$. Then, there exists $V - V' = (h_j - h_j') \cdot a \cdot B_1 = (h_j - h_j') \cdot a \cdot (bP)$, that is,

$$\frac{V - V'}{\left(h_j - h_j'\right)} = abP. \tag{21}$$

Therefore, S can compute abP as a solution to solve CDH problem. Due to the Forking Lemma, the probability of successful rewind simulation is at least $\varepsilon/4$. Then, S can solve the CDH problem with probability $\varepsilon/4$ at least. □

Theorem 10 Anonymity. *Our proposed scheme is anonymous in the random oracle model with the DBDH assumption.*

Proof. Suppose there exists a PPT adversary $\mathscr{A}$ with advantage ε. Then, we use a simulator S to solve the DBDH problem. Let (aP, bP, cP, Z) be a given instance, where a, b, c is random picked in Z_p^*, $Z \in \mathbb{G}_T$, and $P \in \mathbb{G}_1$. Through the adversary $\mathscr{A}$, the simulator S's objective is to determine whether $Z = e(P, P)^{abc}$.

So the simulator S simulates the oracles by interacting with the adversary $\mathscr{A}$ as follows:

(1) Setup: The challenge signature is created using the randomly picked public key in Y. S randomly chooses $u \in \mathbb{Z}_p^*$ and sets $\hat{y} = uP$. In addition, S sets $pk_j = aP$ and $K = bP$. $\mathscr{A}$ is given the system parameters P and the public keys $Y = \{pk_1, pk_2, \cdots, pk_n\}$ of signers

(2) Queries: $\mathscr{A}$ does the same queries with the oracles (H, H_1, H_2, JO and CO) as Theorem 9

(3) Then, $\mathscr{A}$ queries SO, and S performs the steps as follows:

(1) S retrieves the list $L_{\boldsymbol{sign}}$, if (pk_j, I) is found, and then takes the value I. If not, randomly selects a new value $I \in \mathbb{G}_1$ and adds (pk_j, I) to the list $L_{\boldsymbol{sign}}$

(2) S sets $B_1 = cP$ and $B_3 = Z \in \mathbb{G}_T$ and computes $B_2 = c(uP) + bP$

(3) S chooses a random $j \in \{1, \cdots, n\}$

(4) S chooses $l_i \in \mathbb{Z}_p^*$ and $\Phi_i \in_R \mathbb{G}_1$ randomly, where $i = \{1, 2, \cdots, j-1, j+1, \cdots, n\}$ and $i \neq j$; and S computes $C_i = l_i P + f_2(aP)$ and $Z_i = l_i f_1 + f_2 I$ and then selects $h_i \in \mathbb{Z}_p^*$ in L_H such that

$$H(\Phi_i, M, Y, I, Z) = h_i \tag{22}$$

(5) S randomly picks $z, h_j \in \mathbb{Z}_p^*$ and computes

$$\Phi_j = z \cdot P - h_j \cdot pk_j + \left[\sum_{i=1, i \neq j}^{n} h_i pk_i + \Phi_i\right], \tag{23}$$

and stores the value $h_j = H(\Phi_j, M, Y, I, Z)$ to L_H. If there has a collision with hash values in L_H, do this step again until no collision happens.

(6) S computes $V = zB_1$ and then outputs$\sigma = (l_1, \cdots, l_n, \Phi_1, \cdots, \Phi_n, V, I, B_1, B_2)$ as a response to $\mathscr{A}$

S gives σ to $\mathscr{A}$, and $\mathscr{A}$ can query the random oracles adaptively and returns a bit $\eta \in \{0, 1\}$. Suppose $\mathscr{A}$ guesses that the signer's index is $j \in [1, n]$. If $\mathscr{A}$ cannot identify a signer, S returns 0. If $j = s$, it returns 1; if $j = 0$, it returns 0; otherwise, it returns 1/0 with equal probability. If $Z = e(P, P)^{bcd}$, then $B_3 = e(\omega pk_j, K) = e(c(aP), bP) = e(P, P)^{abc} = Z$. And when S returns 0, from $\mathscr{A}$'s view, all signers has equal probability to sign the signature. Suppose $\mathscr{A}$ has advantage ε in the simulation, then we set $\Pr[Z = e(P, P)^{abc}] = 1/2 + \varepsilon$. The probability of the right choice is computed as

$$\boldsymbol{Pr}\left[Z = e(P, P)^{bcd}\right] \geq \frac{1}{2}\left(\boldsymbol{Pr}\left[Z = e(P, P)^{bcd} | S \longleftarrow 1\right] + \boldsymbol{Pr}\left[Z = e(P, P)^{bcd} | S \longleftarrow 0\right]\right) \geq \frac{1}{2}\left[\left(\frac{1}{2} + \frac{1}{2n} + \frac{\varepsilon}{2}\right) + \left(\frac{1}{2} - \frac{1}{2n}\right)\right] = \frac{1}{2} + \frac{\varepsilon}{4}. \tag{24}$$

Therefore, S can determine whether $Z = e(P, P)^{abc}$ with the probability than 1/2 if $\mathscr{A}$ can win, contradiction occurs. □

Theorem 11 Linkability. *Our proposed scheme is linkable in the random oracle model with the discrete logarithm assumption.*

Proof. In order to prove linkability of our roRS scheme, we perform the same setting of oracle queries as the proof in Theorem 9, and we allow S to give $\mathscr{A}$ the public parameters, where S has at most one private key sk_s, and this private key can correspond to two different keys in ring group Y for $i = 1, 2$. (When $\mathscr{A}$ queries the CO, $\mathscr{A}$ can only get one private key. $\mathscr{A}$ can be allowed to get only one private key.)

So through the queries, $\mathscr{A}$ can output two valid signatures:

$$\begin{aligned} \sigma^{(1)} &= \left(\cdot, \Phi_1^{(1)}, \cdots, \Phi_n^{(1)}, V^{(1)}, I^{(1)}\right), \\ \sigma^{(2)} &= \left(\cdot, \Phi_1^{(2)}, \cdots, \Phi_n^{(2)}, V^{(2)}, I^{(2)}\right), \end{aligned} \tag{25}$$

and the key image of $\sigma^{(1)}$ is $I_s^{(1)} = x_s H_1(pk_s)$, and the key image of $\sigma^{(2)}$ is $I_s^{(2)} = x_s' H_1(pk_s)$.

For $\sigma^{(1)}$, S rewinds the same tape with a different value to obtain another valid signature $\tilde{\sigma}^{(1)}$. Then, we obtain

$$\tilde{\sigma}^{(1)} = \left(\cdot, \tilde{\Phi}_1^{(1)}, \cdots, \tilde{\Phi}_n^{(1)}, \tilde{V}^{(1)}, I^{(1)}\right). \tag{26}$$

If $\Phi_s^{(1)} = \tilde{\Phi}_s^{(1)}$, abort. If $\Phi_s^{(1)} \neq \tilde{\Phi}_s^{(1)}$, we have

$$\begin{aligned}\tilde{V}^{(1)} - V^{(1)} &= \left(z + \tilde{h}_s^{(1)} x_s\right)P - \left(z + h_s^{(1)} x_s\right)P,\\ x_s &= \log_P\left(\frac{\tilde{V}^{(1)} - V^{(1)}}{\tilde{h}_s^{(1)} - h_s^{(1)}}\right),\end{aligned} \tag{27}$$

where $I^{(1)} = x_s H_1(pk_s)$ and $y_s = x_s P = pk_s$.

For $\sigma^{(2)}$, S rewinds the same tape with a different value to obtain another valid signature $\tilde{\sigma}^{(2)}$. Then, we obtain

$$\tilde{\sigma}^{(2)} = \left(\cdot, \tilde{\Phi}_1^{(2)}, \cdots, \tilde{\Phi}_n^{(2)}, \tilde{V}^{(2)}, I^{(2)}, B_1, B_2\right). \tag{28}$$

If $\Phi_s^{(2)} = \tilde{\Phi}_s^{(2)}$, abort. If $\Phi_s^{(2)} \neq \tilde{\Phi}_s^{(2)}$, we have

$$\begin{aligned}\tilde{V}^{(2)} - V^{(2)} &= \left(z + \tilde{h}_s^{(2)} x_s'\right)P - \left(z + h_s^{(2)} x_s'\right)P,\\ x_s' &= \log_P\left(\frac{\tilde{V}^{(2)} - V^{(2)}}{\tilde{h}_s^{(2)} - h_s^{(2)}}\right).\end{aligned} \tag{29}$$

Therefore, $x_s = x_s'$ and $I^{(1)} = I^{(2)}$. Two signatures $(\sigma^{(1)}, \sigma^{(2)})$ are linked. S can break DLP, and the advantage of $\mathscr{A}$ is negligible since the rewind simulation is successful. □

Theorem 12 Non-slanderability. *Our proposed scheme is non-slanderable in the random oracle model with the discrete logarithm assumption.*

Proof. In order to prove non-slanderability of our roRS scheme, we perform the same setting of oracle queries as the proof in Theorem 9; the adversary $\mathscr{A}$ can query CO to get any public key, but $\mathscr{A}$ cannot be allowed to query the signer's public key pk_s. But $\mathscr{A}$ can give simulator S the tuples $(pk_s, Y, M, pk_{revoke})$. S uses the tuple to generate a valid signature $\sigma = (\cdot, I)$ which I is the key image computed using x_s. ($\mathscr{A}$ can keep querying oracles with the restriction of submitting pk_s to CO.)

Suppose $\mathscr{A}$ generates another valid signature $\sigma'' = (\cdot, I'')$ which I'' is computed by x_s'' and σ'' is not an output of SO. Additionally, σ and σ'' are linked. Therefore, $I = I''$ which means $I = x_s H_1(pk_s) = I'' = x_s'' H_1(pk_s)$. So $x_s = x_s''$, it can imply that $\mathscr{A}$ knows the secret key x_s corresponding to pk_s. This is opposite to our assumption that $\mathscr{A}$ cannot query C O to get the secret key of pk_s. □

Theorem 13 Revocability. *Our proposed scheme is revocable in the random oracle model if the construction is unforgeable.*

Proof. In order to prove revocability of our roRS scheme, we perform the same setting of oracle queries as the proof in Theorem 9, and we allow S give $\mathscr{A}$ the public parameters, where $\mathscr{A}$ can get one private key denoted as $sk_s = x_s$ corresponding to $pk_s = y_s$ in Y from CO. Due to the fact that $\{pk_1, \cdots, pk_{s-1}, pk_{s+1}, \cdots, pk_n\}$ is randomly and independently distributed, $\mathscr{A}$ cannot tell out the corresponding secret keys according to our assumption. Then, we suppose that $\mathscr{A}$ can generate a valid signature $\sigma = (l_1, \cdots, l_n, \Phi_1, \cdots, \Phi_n, V, I, B_1, B_2)$ successfully for contradiction. The valid signature must be generated by $sk_s = x_s$ because our proposed scheme is unforgeable. There exists a case that can break revocability in our scheme. We consider one case that can break revocability of our scheme. $\mathscr{A}$ randomly selects $j \in \{1, \cdots, s-1, s+1, \cdots, n\} (j \neq s)$ and performs the Sign algorithm; but on behalf of closing the ring, $\mathscr{A}$ must know the secret key x_j $(x_j \neq x_s)$. Due to our assumption that $\mathscr{A}$ can only get one private key, contradiction occurs. □

7. Efficiency Analysis

In this section, performance analysis is shown in Table 2. The efficiency computational cost and signature size between our proposed scheme and several signature schemes such as revocable ring signature [29], traceable ring signature [30] and so on. Then, we have several computational notions need to define as follows in Table 3.

We do the comparison about computation cost on the size of signature schemes and the timings of signing and verifying with [25, 29, 30, 35, 36], where n is the number users included in the group.

From Table 2, we compare our scheme with [25, 35, 36] in signature size, signing, verifying, assumption and security model. We can see that roRS and [25, 35, 36] have the same level of computational complexity in terms of signature size and signature signing. As for verifying, we only need the constant pairing computations which show more efficient than [25, 35, 36] constructed from bilinear pairings similarly. And our security model is built in the random oracle model that is different from [25, 36] built in the standard model.

In Table 4, in relation to the revocation phase, we only need one scalar multiplication computation and one addition computation compared with n pairing computations in [29] and $2n$ addition computations and $2n$ scalar multiplication computations in [30]. As for the functionality of roRS and [29, 30], [29] cannot achieve the linkability, and [30] cannot provide mandatory revocability; nevertheless, we have accomplished the linkability and the mandatory revocation simultaneously in the random oracle.

8. Conclusion

In this literature, we proposed a novel revocable one-time ring signature scheme based on bilinear pairings, which is provable secure under the DL, DBDH, and CDH assumptions in the random oracle model. In our scheme, we have simultaneously introduced linkability and mandatory revocability to distinguish from other ring signature schemes. In particular, linkability can prevent the double-signing attack, and mandatory revocability guarantees that a revocation authority can identify the actual signer when the actual signer commits a crime in transactions. The scheme about revocation phase requires only one scalar multiplication computation and one additional computation, and the pairing computations in the timings of verifying phase is constant. And constructing provable security revocable one-time ring signature schemes from lattices to resist quantum attackers is an interesting problem that we leave open for further research.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This paper was supported by the National Natural Science Foundation of China (Grant no. U21A20463). We thank Zhang et al. [37] who have provided the inspiration for our scheme.

References

[1] R. L. Rivest, A. Shamir, and Y. Tauman, "How to leak a secret," in *Advances in Cryptology-ASIACRYPT 2001*, C. Boyd, Ed., Springer Berlin Heidelberg, Berlin, Heidelberg, 2001.

[2] J. K. Liu, M. H. Au, W. Susilo, and J. Zhou, "Linkable ring signature with unconditional anonymity," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 1, pp. 157–165, 2014.

[3] P. P. Tsang and V. K. Wei, "Short Linkable Ring Signatures for E-Voting, E-Cash and Attestation," in *Information Security Practice and Experience*, R. H. Deng, F. Bao, H. H. Pang, and J. Zhou, Eds., Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.

[4] Y. Dodis, A. Kiayias, A. Nicolosi, and V. Shoup, "Anonymous Identification in Ad Hoc Groups," in *Advances in Cryptology-EUROCRYPT 2004*, C. Cachin and J. L. Camenisch, Eds., pp. 609–626, Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[5] K. Kajita, K. Ogawa, and E. Fujisaki, "A constant-size signature scheme with a tighter reduction from the cdh assumption," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E103.A, no. 1, pp. 141–149, 2020.

[6] W. Wang, J. Song, G. Xu, Y. Li, H. Wang, and C. Su, "Contractward: automated vulnerability detection models for ethereum smart contracts," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 2, pp. 1133–1144, 2021.

[7] S. Noether, A. Mackenzie, and T. M. Research Lab, "Ring confidential transactions," *Ledger*, vol. 1, pp. 1–18, 2016.

[8] M. Abe, M. Ohkubo, and K. Suzuki, "1-out-of-n signatures from a variety of keys," in *Advances in Cryptology-ASIACRYPT 2002*, YC. Zheng, Ed., pp. 415–432, 2002.

[9] A. Bender, J. Katz, and R. Morselli, "Ring signatures: stronger definitions, and constructions without random oracles," in *Theory of Cryptography*, S. Halevi and T. Rabin, Eds., pp. 60–79, Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[10] R. L. Rivest, A. Shamir, and Y. Tauman, "How to leak a secret: theory and applications of ring signatures," in *Theoretical Computer Science*, pp. 164–186, Springer, 2006.

[11] H. Shacham and B. Waters, "Efficient ring signatures without random oracles," in *International Workshop on Public Key Cryptography*, pp. 166–180, Springer, 2007.

[12] D. Boneh, C. Gentry, B. Lynn, and H. Shacham, "Aggregate and verifiably encrypted signatures from bilinear maps," in *Advances in Cryptology — EUROCRYPT 2003*, E. Biham, Ed., pp. 416–432, Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.

[13] P.-L. Cayrel, R. Lindner, M. Rückert, and R. Silva, "A lattice-based threshold ring signature scheme," in *Progress in Cryptology-LATINCRYPT 2010*, M. Abdalla and P. S. L. M. Barreto, Eds., pp. 255–272, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[14] J. K. Liu, T. H. Yuen, and J. Zhou, "Forward secure ring signature without random oracles," in *Information and Communications Security*, S. Qing, W. Susilo, G. Wang, and D. Liu, Eds., pp. 1–14, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[15] R. Cramer, I. Damgård, and B. Schoenmakers, "Proofs of partial knowledge and simplified design of witness hiding protocols," in *Advances in Cryptology-CRYPTO '94*, Y. G. Desmedt, Ed., pp. 174–187, Springer Berlin Heidelberg, Berlin, Heidelberg, 1994.

[16] J. Camenisch and A. Lysyanskaya, "Dynamic accumulators and application to efficient revocation of anonymous credentials," in *Advances in Cryptology-CRYPTO 2002*, M. Yung, Ed., pp. 61–76, Springer Berlin Heidelberg, Berlin, Heidelberg, 2002.

[17] M. H. Au, J. K. Liu, W. Susilo, and T. H. Yuen, "Constant-size id-based linkable and revocable-iff-linked ring signature," in *Progress in Cryptology-INDOCRYPT 2006*, R. Barua and T. Lange, Eds., pp. 364–378, Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[18] I. R. Jeong, J. O. Kwon, and D. H. Lee, "Ring signature with weak linkability and its applications," *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 8, pp. 1145–1148, 2008.

[19] J. K. Liu and D. S. Wong, "Linkable ring signatures: security models and new schemes," in *Computational Science and Its Applications-ICCSA 2005*, O. Gervasi, M. L. Gavrilova, V. Kumar, A. Laganá, H. P. Lee, Y. S. Mun, D. Taniar, and C. J. K. Tan, Eds., pp. 614–623, Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.

[20] M.-J. Qin, Y.-L. Zhao, and Z.-J. Ma, "Practical constant-size ring signature," *Journal of Computer Science and Technology*, vol. 33, no. 3, pp. 533–541, 2018.

[21] T. H. Yuen, J. K. Liu, M. H. Au, W. Susilo, and J. Zhou, "Efficient linkable and/or threshold ring signature without random oracles," *The Computer Journal*, vol. 56, no. 4, pp. 407–421, 2013.

[22] P. P. Tsang, V. K. Wei, T. K. Chan, M. H. Au, J. K. Liu, and D. S. Wong, "Separable linkable threshold ring signatures," in *Progress in Cryptology-INDOCRYPT 2004*, A. Canteaut and K. Viswanathan, Eds., pp. 384–398, Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.

[23] J. K. Liu, V. K. Wei, and D. S. Wong, "Linkable spontaneous anonymous group signature for ad hoc groups," in *Information Security and Privacy*, H. Wang, J. Pieprzyk, and V.

Varadharajan, Eds., pp. 325–335, Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[24] D. Zheng, X. Li, K. Chen, and J. Li, “Linkable ring signatures from linear feedback shift register,” in *Emerging Directions in Embedded and Ubiquitous Computing*, M. K. Denko, C. S. Shih, K. C. Li, S. L. Tsao, Q. A. Zeng, S. H. Park, Y. B. Ko, S. H. Hung, and J. H. Park, Eds., pp. 716–727, Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.

[25] D. Y. Liu, J. K. Liu, Y. Mu, W. Susilo, and D. S. Wong, “Revocable ring signature,” *Journal of Computer Science and Technology*, vol. 22, no. 6, pp. 785–794, 2007.

[26] C. A. Melchor, P.-L. Cayrel, P. Gaborit, and F. Laguillaumie, “A new efficient threshold ring signature scheme based on coding theory,” *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4833–4842, 2011.

[27] Y. Tang, F. Xia, Q. Ye, M. Wang, R. Mu, and X. Zhang, “Identity-based linkable ring signature on ntru lattice,” *Security and Communication Networks*, vol. 2021, Article ID 9992414, 17 pages, 2021.

[28] M. Hu and Z. Liu, *Lattice-Based Linkable Ring Signature in the Standard Model*, Cryptology ePrint Archive, 2022.

[29] L. Li, J. Liu, L. Cheng et al., “Creditcoin: a privacy-preserving blockchain-based incentive announcement network for communications of smart vehicles,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 7, pp. 2204–2220, 2018.

[30] E. Fujisaki and K. Suzuki, “Traceable ring signature,” in *International Workshop on Public Key Cryptography*, pp. 181–200, Springer, 2007.

[31] E. Fujisaki, “Sub-linear size traceable ring signatures without random oracles,” in *Topics in Cryptology-CT-RSA 2011*, A. Kiayias, Ed., pp. 393–415, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[32] M. H. Au, J. K. Liu, W. Susilo, and T. H. Yuen, “Secure id-based linkable and revocable-iff-linked ring signature with constant-size construction,” *Theoretical Computer Science*, vol. 469, pp. 1–14, 2013.

[33] H. Feng, J. Liu, D. Li, Y.-N. Li, and Q. Wu, “Traceable ring signatures: general framework and post-quantum security,” *Designs, Codes and Cryptography*, vol. 89, no. 6, pp. 1111–1145, 2021.

[34] J. Herranz and G. Sáez, “Forking lemmas for ring signature schemes,” in *Progress in Cryptology - INDOCRYPT 2003*, T. Johansson and S. Maitra, Eds., pp. 266–279, Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.

[35] F. Zhang, R. Safavi-Naini, and W. Susilo, “An efficient signature scheme from bilinear pairings and its applications,” in *International workshop on public key cryptography*, pp. 277–290, Springer, 2004.

[36] S. Schäge and J. Schwenk, “A cdh-based ring signature scheme with short signatures and public keys,” in *Financial Cryptography and Data Security*, R. Sion, Ed., pp. 129–142, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[37] X. Zhang, J. K. Liu, R. Steinfeld, V. Kuchta, and J. Yu, “Revocable and linkable ring signature,” in *Information Security and Cryptology*, Z. Liu and M. Yung, Eds., pp. 3–27, Springer International Publishing, Cham, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 2795004, 11 pages
https://doi.org/10.1155/2022/2795004

Research Article

Detectable, Traceable, and Manageable Blockchain Technologies BHE: An Attack Scheme against Bitcoin P2P Network

Jiale Yang,[1] **Guozi Sun**,[1,2] **Rongyu Xiao**,[1] **and Hansen He**[3]

[1]*School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China*
[2]*Key Laboratory of Urban Land Resources Monitoring and Simulation, MNR, Shenzhen 518000, China*
[3]*Nanjing Jiangshipin Data Technology co. LTD, Nanjing 210019, China*

Correspondence should be addressed to Guozi Sun; sun@njupt.edu.cn

Received 27 March 2022; Accepted 11 July 2022; Published 2 August 2022

Academic Editor: Yin Zhang

As the most successful cryptocurrency, bitcoin has become the primary target of attackers. The security risks existing in bitcoin network (P2P networks) may pose serious threats to itself. It has been proved that network attackers of the autonomous system level could isolate a specific set of bitcoin nodes using prefix hijacking attacks; since this attack achieves bitcoin partition by deleting all data packets of the victim node, it is easy to be discovered by the victim node, and cannot maintain a long-term connection (the partition will disappear after canceling the BGP hijacking) (Apostolaki M et al. (2017)). This paper proposes a new attack scheme—eclipse attack method based on BGP hijacking (BHE). The attack can occupy the network connection of the victim node, and only need to delete part of the TCP handshaking packets of the victim node during the attack, and it makes the attack more hidden and can occupy the network connection of the victim node for a long time. The innovation of the BHE attack is that it can control the peering decision of the victim node by controlling the victim node's internal peer database (new table and tried table) and preventing the victim node from establishing a good connection. It enables the attacker to occupy all network connections of the victim node and become its natural network middleman. We verify the feasibility of the BHE attack through experimental evaluation and demonstrate that an attacker who can launch BGP hijacking may occupy all connections of the victim node within 20 minutes (ignoring the time of traffic diversion). To reduce the attack's impact, the paper provides some countermeasures that can use in practice according to the basic characteristics of the attack.

1. Introduction

The essence of the bitcoin system is a decentralized ledger based on the Internet, and the blockchain is the name of this ledger. The bitcoin system does not depend on a centralized entity. All nodes in the system have equal identities and are connected to form a huge p2p network and share the mission of providing network services. Anyone can join this decentralized network through a bitcoin client [1]. Satoshi Nakamoto pointed out that the bitcoin system uses a proof-of-work mechanism to ensure the consistency of the entire blockchain state [2]. Any attacker who attempts to destroy the bitcoin system needs to control more than half of the computing power in the entire network to break this consistency, which undoubtedly guarantees the security of the bitcoin. However, this security is based on the consistency of information. All Bitcoin nodes have the same view of the blockchain and can always receive the same blocks and transactions within a certain period, which requires the bitcoin network to be safe and reliable.

The bitcoin network is a typical P2P network. Each node in the network maintains a long-term connection with multiple peer nodes. Through these connections, nodes exchange blockchain views to synchronize information and maintain the consistency of the blockchain state. The purpose of bitcoin network-level attacks is to control the network connection of the victim node as much as possible so that the victim node cannot receive the latest or even receive the wrong blockchain view [3]. The bitcoin partition attack showed that an autonomous system (AS) with a large number of IP sources can intercept Bitcoin traffic by hijacking interdomain routing and then isolate the selected victim node set from the bitcoin network [4]. Due to the characteristics of BGP hijacking and the way the attack is implemented (dropping the traffic of the victim node), the attack has the defects of not being able to maintain a long-term connection with the victim node and being easy to detect.

Based on the BGP hijacking mechanism and the Bitcoin network mechanism, this paper proposes a new attack method against the bitcoin network, which can occupy the peer-to-peer connection of the victim node, and only needs to delete some TCP handshake packets of the victim node during the attack. Therefore, compared with the Bitcoin hijacking attack, this attack is more difficult to detect and can maintain a long-term connection with the victim node (even if the BGP hijacking is canceled, the partition will not disappear).

Figure 1 provides a general overview of BHE, mainly describing how the attack occupies the peering connections of the victim node (in AS A) (only two connections are shown here). After hijacking the traffic of the victim node, the attacker (AS D) does not indiscriminately delete all packets of the victim node but monitors the network activity of the victim node. When the victim node establishes a new connection (blue dotted line), the attacker blocks the formation of a good connection (A to E) and forces the victim node to the bitcoin nodes (in AS D and F) (the path from the victim node to these nodes contains attacker D) and establishes the connection (solid red line). Eventually, the attacker will occupy all peer connections of the victim node. In this way, D can control the network view of the victim node through ordinary traffic interception and Bitcoin message forgery.

The rest of the paper is organized as follows. The second chapter gives an overview of the bitcoin P2P network and typical Bitcoin network attacks. The third chapter describes the current research status at home and abroad. The paper detailed introduces the BHE attack in the fourth chapter and evaluates BHE attacks in chapter 5. Then, the sixth chapter puts forward some countermeasures to BHE in a targeted manner. Finally, the seventh chapter concludes this article.

2. Background

This section first introduces the bitcoin network and then reviews two typical Bitcoin network attacks.

2.1. Bitcoin Peer-to-Peer Network. The bitcoin network is a vast p2p network mounted on the Internet, and all nodes in the network are identified by IP addresses. Each Bitcoin node can select up to 8 remote nodes to connect. If the node has a public IP, it can also accept up to 117 incoming connections [5, 6].

During the running of the node, it will store and broadcast the verified node information to maintain the stability of the network. Below, we describe the most noteworthy part of this section: how a node obtains and stores network information and how to select a node to connect.

2.1.1. Node Information Propagation. Bitcoin provides two ways to spread node information: DNS seed and addr message. DNS seed is a DNS server that can return the address information of full nodes on the bitcoin network, and it is hardcoded into the source code to help the node joining the network for the first time find the full node [7]. Addr message is a list containing no more than 1000 node information used to relay node information on the network.

2.1.2. Node Information Storage. The IP address of the public node is stored in the tried table and new table of the bitcoin node.

The tired table contains 64 buckets. Each bucket stores the nodes that have established outgoing connections and can store up to 64 IP addresses. New table contains 1024 buckets. Each bucket stores the IP addresses that have not successfully established a connection and can store up to 64 IP addresses.

2.1.3. Peer-to-Peer Connection. Bitcoin nodes will maintain at most 8 outgoing connections by default. When the node restarts or an outgoing connection is disconnected, the node will establish a new connection. When establishing a new connection, the node will select the new table or tried table with a probability of 1/2 and then randomly select an IP whose group is distinct from other outgoing connections from the selected table to connect.

2.2. BGP. BGP (border gateway protocol) is a network protocol used to exchange routing information between networks on the Internet [8, 9]. Generally, it is used to determine the best path to route data between independent networks or autonomous systems. Different autonomous systems need to exchange routing information and inform each other of their IP prefix. When an autonomous system obtains a new IP prefix, it needs to broadcast the routing information to its neighbors, who will further spread the information until the whole network receives and stores the routing information.

In BGP protocol, the authenticity of routing information will not be checked, which means that any autonomous system can publish false routing information, causing other autonomous systems to send traffic to the wrong location [10]. This attack is called BGP hijacking.

2.3. Bitcoin Eclipse Attack. The purpose of the eclipse attack is to monopolize the peer-to-peer connection of the victim node and partition the victim node [11]. The early Bitcoin address manager had some vulnerabilities in the new table and the tried table. For example, tried table stores the IP address of the incoming connection, and the node does not check the validity of the IP before inserting the IP into the new table. It allows the attacker to control a botnet or the basic organization to fill the new table and tried table of the victim node with invalid IP and malicious IP, respectively, and finally, occupy all connections of the victim node when the victim node restarts. At present, the bitcoin community has fixed these vulnerabilities.

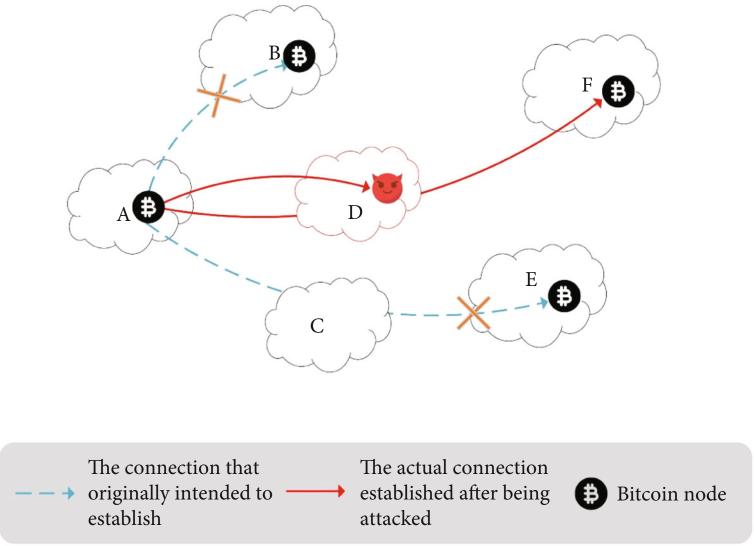

B
F
A
D
E
C
The connection that originally intended to establish
The actual connection established after being attacked
Bitcoin node

Figure 1: Adversary D used the bitcoin partition to prevent the A's legal network connections from forming and then force A to establish connections to D or through D.

2.4. Bitcoin Partition Attack. The principle of partition attack is to use the well-known BGP prefix hijacking vulnerability to redirect the outgoing and incoming connections of the victim node set to the attacker's autonomous system (AS), and then control the communication between the victim node set and its neighbor [4]. Attackers generally take the country or nation as the background and rely on their network topology advantages to hijack the traffic of the victim node. If the attacker can hijack all the connections of the victim node, he can isolate the victim node from the bitcoin network.

3. Related Work

3.1. Attacks on Bitcoin Peer-to-Peer Networks. Chapter 2 describes the bitcoin eclipse attack [11] and the bitcoin partition attack [4], which are most similar to BHE. Among them, the BHE attack and the eclipse attack have similar attack effects (both can monopolize the network connection of the victim node, making the attacker become the natural middleman of the victim node) [11], but BHE does not take advantage of Bitcoin's own vulnerabilities but relies on the attacker's network topology advantage to hijack the connection of victim nodes, so there is currently no patch that can be applied to BHE attacks. Although both BHE and partition attacks utilize the BGP prefix hijacking mechanism, their attack strategies are different. The bitcoin partition attack achieves the partition by deleting the data packets of the victim node, so the attack will form a black hole, which is easy to be discovered by the victim node, and the partition will disappear when the attacker cancels the BGP hijacking; that is, the attack cannot maintain a long-term connection [4]. BHE implements partitioning by controlling the peering decision of the victim node. During this process, only a part of the TCP handshake packets are deleted, which has higher concealment than the partition attack [4], and after the attack is successful, the partitioning will not disappear even if the BGP hijacking is canceled.

Gervais et al. took advantage of the vulnerability that bitcoin nodes only request blocks from the same neighbor each time and successfully delayed the time for the victim node to receive blocks by 20 minutes; the vulnerability exploited by the attack has now been patched [12]. Walck et al. successfully used one full node and two light nodes to carry out delay attacks on the victim node by taking advantage of the vulnerability of the new Bitcoin propagation protocol [13]; compared with BHE and eclipse attacks, the delay attack can only prevent the victim from receiving the latest block and does not allow the attacker to send the wrong block to the victim node. Yves Christian et al. used the defects of the bitcoin's behavior mechanism to realize an eclipse attack on the victim node with a small amount of IP [14], and the attack is based on an eclipse attack, so this attack is difficult to implement in the latest Bitcoin. Muoi tran et al. proposed an attack on the data level of the bitcoin network [15]. The attack does not apply any routing operation. It uses the attacker's network topology advantage to fill the malicious IP to the victim node, slowly affecting the routing table of the victim node. Although the attack can influence the peering decision of the target node, it requires the routing table of the victim node to be heavily populated with malicious IPs, which often takes weeks.

3.2. Defensive Measures for the Bitcoin Network Attack. At present, the research results at home and abroad mainly resist network attacks by optimizing the bitcoin network structure. A more efficient network structure means higher network connectivity, which makes it difficult for attackers to control the network view of nodes.

Marcal et al. proposed an adaptive network mechanism that can reduce bandwidth [16]. And they showed that this mechanism would reduce bandwidth consumption by 10.2%, reduce the number of exchanged messages by 41.5%, and not harm transaction submission. Gleb Naumenko et al. proposed a more effective bitcoin transaction relay protocol [17]. The protocol abandoned the original messages flooding transmission mode and adopted the collective coordination method. Otsuki el at. reduce the degree of data redundancy and improve network connectivity by adjusting the ratio of relay nodes in the bitcoin [18]. Bin Zhang proposes an eclipse attack traffic detection method based in a custom combination of features and deep learning [19] and a distributed DDoS-attack traffic detection method based on a cross multilayer convolutional neural network model in the blockchain network layer [20].

4. The BHE Attack

In this section, we first introduce the attack model of BHE based on the threat model considered in this paper. Then, we describe the attack process of BHE in detail; in this part, we propose an attack strategy that can control the peering decision of the victim node. The strategy is to control the internal routing table of the victim node and prevent the victim node from establishing a good connection. Finally, we analyze the possible harm caused by BHE.

4.1. Threat Model. Similar to the bitcoin partition attack [4], our attacker is a network adversary that controls a single AS, and the attacker's goal is to control all network connections of the victim node. Our victim node is a Bitcoin node with a public IP. Besides, we assume that during the attack, the attacker can hijack all the traffic of the victim node through the BGP prefix hijacking attack and leak point deletion algorithm [4].

4.2. Attack Model. Aiming at the problem that Bitcoin hijacking attacks are easy to be detected and cannot maintain long-term malicious connections [4], we propose a new Bitcoin network attack method—BHE based on BGP prefix hijacking mechanism and Bitcoin network mechanism, which can control the peering decision of the victim node, allowing the attacker to occupy all the peering connections of the victim node in a short time, becoming the natural middleman of the victim node.

Our attacker is a malicious AS (AS D in Figure 1), and the victim node is a bitcoin node with a public IP (node in AS A). Since the original peer-to-peer connection of the victim node does not necessarily pass through the attacker, the attacker's attack goal is to force the victim node to establish peer-to-peer connections to some special nodes (nodes in D or F) in order to manipulate these connections.

BHE attacks are mainly divided into two phases: attack preparation and attack execution.

4.2.1. Attack Preparation. During the attack preparation phase, the attacker's goal is to collect IP addresses that can be used for the attack, and these IPs have special IP prefixes, and when the victim node establishes outgoing connections to these IPs, the route passes through the attacker. We call these IPs as malicious IPs.

4.2.2. Attack Execution. During the attack execution phase, the attacker's goal is to force the victim node to establish an outgoing connection to the malicious IP. The traditional eclipse attack has been proved to be unable to affect the peering decision of the node [11], and the attacker of BHE takes advantage of its network topology, that is, our attacker attackers can simulate different malicious IPs to slowly fill up the victim node's internal peer database, and more importantly, our attackers capture the relevant data packets of the victim node through the BGP hijacking attack and perform operations such as deletion and modification, which can prevent the victim node from establishing a good connection and learning a good IP. As shown in Figure 2, this stage consists of 3 steps: (1) Hijack the victim node's traffic. (2) Fill the victim node's internal peer database with malicious IPs. This step is similar to the Bitcoin eclipse attack [11, 15], and the purpose is to increase the probability of the victim node establishing an outgoing connection to malicious IPs. (3) Observe the network activity of the victim node through hijacked packets, wait for it to establish a new connection, and prevent it from establishing a good outgoing connection (e.g., AS1 to AS2), and to speed up the process, the attacker may force the victim node to restart.

4.2.3. Attack Properties. BHE attacks have the following three properties:

(1) Stealthiness: Unlike the Bitcoin partition attack, which indiscriminately discards the data packets of the victim node [4], the attacker of BHE will only delete part of the TCP handshake packets of the victim node to prevent it from establishing a good connection. During the attack, the victim node can communicate normally, so it is difficult to detect that it is being attacked in time

(2) Persistence: The attacker of the BHE attack will eventually occupy all network connections of the victim node (similar to the eclipse attack). After the attack is successful, even if the BGP hijacking attack is canceled, the attacker can still operate the network view of the victim node. And if the attacker has enough computing power, he can launch n confirmed double-spend attacks to the victim node

(3) Efficiency: BHE attackers mainly control the peering decision of the victim node through route hijacking, without waiting for the internal routing table of the victim node to be filled with malicious IP in a large area. Therefore, compared with traditional eclipse attacks, BHE allows attackers to occupy the network connection of the victim node for a relatively short time

4.3. Attack Process

4.3.1. Attack Preparation. During the attack preparation phase, the attacker needs to collect malicious IPs for the selected victim nodes. The malicious IP is determined by the topological relationship between the attacker and the

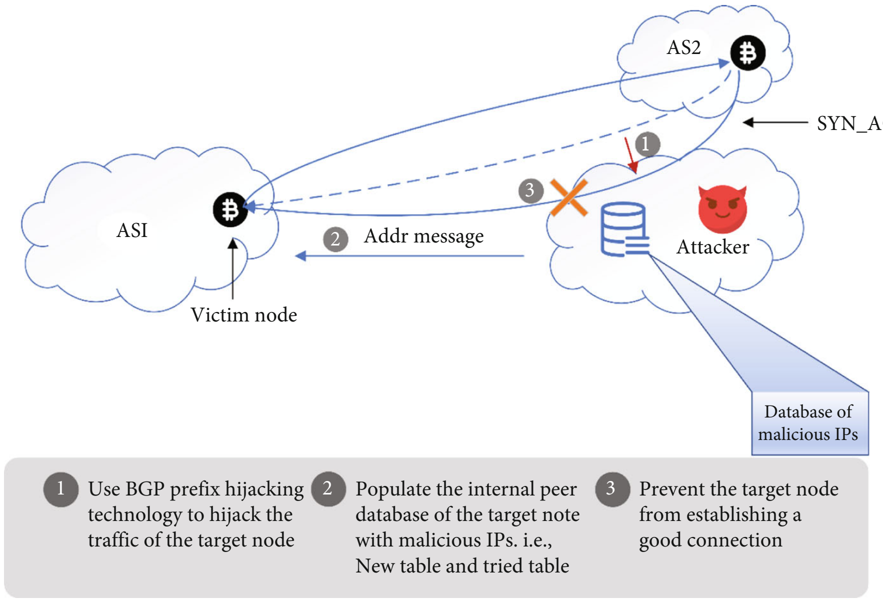

AS2
SYN_ACK
1
ASI
3
2 Addr message
Attacker
Victim node
Database of malicious IPs
1 Use BGP prefix hijacking technology to hijack the traffic of the target node
2 Populate the internal peer database of the target note with malicious IPs. i.e., New table and tried table
3 Prevent the target node from establishing a good connection

Figure 2: Describe the execution flow of a BHE attack, which consists of 3 steps and is able to control the peering decision of the victim node.

victim node, and the corresponding malicious IP can be enumerated by evaluating the interdomain routing state. Research has proven that most AS can easily enumerate a large number of malicious IPs (several million or more) [15]. Please note that the malicious IPs enumerated by the attacker are all valid IPs. Validity here is not meant to represent a real Bitcoin node, but it means that these IPs are legitimate (the network segment is correct).

4.3.2. Attack Execution. In the attack execution phase, the attacker will control the peering decision of the victim node based on traffic hijacking, which is embodied by controlling the internal routing table (new table and tried table) of the victim node and preventing the victim node from establishing a good connection. Eventually, the attacker will occupy all network connections of the victim node. Below, we describe our attack strategy in detail.

(1) Hijack Traffic. Similar to the Bitcoin hijacking attack, the attacker first transfers the network traffic of the victim node through the BGP hijacking attack, then deletes the leak point in the victim node (if the target is a set) [4], and finally hijacks all the network traffic of the victim node. Note that, unlike the Bitcoin hijacking attack, this step only hijacks the network packets of the victim node, and the manipulation of these packets will take place in the next two steps.

(2) Dominate the Internal Peer Database of Victim Node. (2)1. How to Dominate New Table. The running node mainly obtains the IP addresses through the addr messages. The node stores the learned IP addresses in the new table. Due to the characteristics of bitcoin's plaintext transmission, after hijacking the traffic of the victim node, the attacker can capture the addr message sent to the victim node and modify the IP entry in it. The above operations do not modify the structure of the data packet and therefore do not attract the attention of both communicating parties.

When a Bitcoin node inserts an IP address into its new table, it hashes the IP prefix group (IP group) (i.e., the/16 of IPv4 addresses or/32 IPv6 addresses) and the prefix group of the peer relayed that IP (peer group) to determine the bucket for the IP among 1,024 buckets in total; i.e.

$$\begin{aligned} h_1 &= \mathrm{H}(\mathrm{SK}, \mathrm{ip\_group}, \mathrm{peer\_group}), \\ h_2 &= H(\mathrm{SK}, \mathrm{peer\_group}, h_1\%64), \\ \mathrm{bucket\_new} &= h_2\%1024, \end{aligned} \tag{1}$$

where H(·) is the SHA-256 hash function and SK is a secret key of the node. The exact slot for IP in the bucket (which contains 64 slots) is determined by hashing the bucket index and the entire IP address; i.e.,

$$\mathrm{solt\_new} = \mathrm{H}\left(\mathrm{SK}, {}'\mathrm{N}', \mathrm{bucket\_new}, \mathrm{ip}\right). \tag{2}$$

If the slot is already occupied, a validity check is performed on the existing IP (for example, if the existing IP is more than 30 days old, or if the connection fails several times). If the existing IP has expired, it will be replaced by the new IP being inserted; otherwise, the IP being inserted is ignored. Note that the IP address is stored with the timestamp. If the IP is already in the new table, its timestamp will be updated.

The eclipse attack populates a new table of victim nodes by sending malicious IPs quickly. In an eclipse attack, the attacker will continuously establish connections to the victim node and send addr messages containing a large number of malicious IPs to the victim node through these connections and then slowly wait for the victim node's new table to be filled with malicious IPs [11]. The disadvantage of this method is that it cannot prevent the victim node from receiving good IPs; that is, good IPs compete with malicious IPs.

We propose a new strategy to fill up the new table of the victim node based on BGP hijacking. The strategy makes the victim node unable to learn good IPs, which makes up for the defect that traditional eclipse attack cannot prevent good IP insertion when filling new table [11, 15].

The strategy consists of the following two parts:

(1) The first part is similar to the eclipse attack. The attacker simulates different malicious IPs to repeatedly establish incoming connections to the victim node and then sends addr messages containing 1000 malicious IPs to the victim node

(2) The attacker prevents the victim node from learning a good IP through BGP hijacking. The specific process is described in Algorithm 1; the attacker first filters out the addr messages sent by the ordinary node to the victim node by identifying the destination IP, source IP, and message type of the data packet, then replaces the IP addresses in messages with the malicious IPs in sequence, and finally sends it to the victim node

(2)2. How to Dominate Tried Table. In the new version of bitcoin, the IP address in the tried table can only come from the new table, and the tried table cannot be accessed directly from the outside, which is to prevent attackers from inserting malicious IP addresses directly into the tried table. Therefore, ordinary eclipse attack methods are difficult to impact the tried table directly.

The bitcoin node will migrate the IP address in the new table to the tried table in two cases. First, when the node successfully establishes an outgoing connection to an IP in the new table, the IP address will be transferred to the tried table. Second, the bitcoin node will randomly select an IP from the new table every two minutes to establish an outgoing connection called a probe connection. If the probe connection establishes successfully, the selected IP will be transferred to the tried table.

When a Bitcoin node inserts an IP address into the tried table, it needs to perform a series of hash operations on the IP to obtain the index of its bucket and slot.

$$\begin{aligned} h_1 &= H(\mathrm{SK}, \mathrm{ip}), \\ h_2 &= H(\mathrm{SK}, \mathrm{ip\_group}, \mathrm{h}_1\%8), \\ \mathrm{bucket\_tried} &= h_2\%256, \\ \mathrm{solt\_tried} &= H\left(\mathrm{SK}, {}'\mathrm{K}', \mathrm{bucket\_tried}, \mathrm{ip}\right)\%64. \end{aligned} \tag{3}$$

The strategy of BHE filling the tried table is similar to trickle-down attack; that is, by filling the new table, it slowly affects the tried table [15]. In addition, we have further optimized the attack process based on BGP hijacking, so that the victim node cannot establish a connection to a good IP, which accelerates the rate of IP transfer. This method is similar to preventing the victim node from establishing a good connection and is described in detail in the next subsection.

4.3.3. Prevent the Victim Node from Establishing a Good Connection. When a Bitcoin node establishes an outgoing connection to an IP, it needs to perform a TCP three-way handshake. Because the attacker hijacks all network traffic of the victim node, the data packets (SYN or SYN_ACK) used by the victim node to establish an outgoing connection can be captured (e.g., the data packet sent by AS2 to AS1 in Figure 2). If the peer of the victim node is not a malicious IP, the attacker will delete the corresponding packet to prevent the connection from being established.

When the victim node establishes a new outgoing connection due to disconnection or restart, the attacker can prevent it from establishing an outgoing connection to a good IP through the following steps (as shown in Algorithm 2):

(1) According to the source IP, port number (Bitcoin default port number is 8333), and data packet type (SYN data packet) of the data packet, determine whether the data packet is a TCP handshaking packet used by the victim node to establish a new connection, if so, enter the second step. If not, go to step 3 for judgment

(2) If the destination address of the handshaking packet is a malicious IP, then the purpose of the attack is achieved. The attacker will pretend to be that IP to communicate with the victim node. Otherwise, the data packet will be discarded, and the attack will enter the next cycle

(3) The attacker may fail to intercept the SYN packet sent by the victim node, but the attacker may intercept the response packet sent by the legitimate node to the victim node. The identification method is similar to that of 1. If the packet is identified as a response packet (SYN-ACK), drop the packet. Otherwise, forward the packet along the original path to avoid forming a black hole and enter the next cycle

During the attack, the victim node will establish an outgoing connection to malicious IP (such as A to D and A to F in Figure 1), and the attacker will intercept the connection and communicate with the victim node by disguising the malicious IP through source IP spoofing. Ultimately, the attacker will occupy all network connections of the victim node in this way and then partition it. Because the route from the victim node to the malicious IP contains the attacker, even if the attacker cancels the BGP hijacking, he can still intercept the network packet of the victim node and occupy its network connection. In other words, the attacker can occupy the network connection of the victim

```
Input: S=[pkt1…]: hijacked network packets. M: the set of malicious IPs. dp: the ip of victim node
for pkt ∈ S do
    if pkt.ipDst=dp and pkt.ipSrc not in M and pkt.payload=Addr then
        for index in len(pkt.payload) do
        pkt.payload[index].ip = M[index]
        end for
    end if
    send(pkt)
end for
```

Algorithm 1: Modify the IPs in the addr message to malicious IPs to prevent the victim node from learning a good IP.

```
Input: S = [pkt1…]: hijacked network packets. P: the set of malicious IP prefixes. dp: the ip of victim node
for pkt ∈ S do
    if pkt.ipSrc=dp and pkt.dport=8333 and pkt.payload=SYC then
        ipStr ←‘ ’
        ipStr ← Prefix(pkt.ipDst)
        if ipStr in P then
        success(pkt)
    else
        drop(pkt)
        end if
      else if pkt.ipDst=dp and pkt.sport=8333 and pkt.payload=SYC_ACK then
      drop(pkt)
  else
      send(pkt)
  end if
end for
```

Algorithm 2: Selectively deleting TCP handshaking packets forces the victim node to establish an outgoing connection to the victim node.

node for a long time, and even if the BGP hijacking is canceled, the formed partition will not disappear.

Please note that it is relatively hidden to prohibit the victim node from establishing a new connection with a legal IP by discarding the TCP handshaking packet. Because the connection has not formed, the victim node will think that the remote node does not exist and then randomly select an IP from the routing table to connect.

4.3.4. Occupy Incoming Connections. Since Bitcoin nodes receive unsolicited incoming connections, and in most cases, incoming connections are very short-lived (e.g., a couple of minutes) [15], the attacker can easily occupy all incoming connections of the victim node; e.g., the attacker can simulate malicious IPs to repeatedly establish a connection to the victim node.

4.4. Implications of BHE Attack. BHE is the strengthened eclipse attacks, so the damage for the bitcoin with the eclipse attacks is similar. It destroys the information consistency of bitcoin network, which allow an attacker to control the network view of the victim node and easier to launch the traditional attacks to the victim node, such as the double-spending attack [21–23] and selfish mining [24–27]. In particular, compared to the partition attack [4], BHE supports the n-confirmation double spend; in the scenario of a n-confirmation double spend, the victim node has to wait for the confirmation of n blocks before accepting the current transaction; and since BHE can forge the blockchain view of the victim node by sending fake blocks, an attacker with sufficient computing power can achieve the n-confirmation double spend attack.

5. Experiment

We simulate the attack scenario of BHE in the laboratory environment and evaluate the effectiveness of BHE. The experiment configures 3 Ubuntu 18 machines with public IP to simulate the attack environment. Two machines equipped with Bitcoin core (v0.19.1) act as the attack node and the victim node, and the remaining one acts as the proxy of the attacked node.

5.1. Attack Preparation. Since BGP hijacking in the public network environment requires a lot of network resources, it is difficult to implement. So we simulate traffic hijacking by the proxy host. And in the experimental environment, we can evaluate and compare different combinations of attack strategies.

Our attack script is configured in the proxy host and implements the ability to flood the victim node routing table with malicious IPs and prevents the victim from establishing a good connection (the attack process in chapter 4). In addition, we implement some additional functions in the script to simulate the execution of the attack more realistically: (1) The traffic sent to the malicious IP is directed to the attacking host through the address translation function (NAT) of the firewall (iptables) to ensure that the malicious IP is always reachable [28, 29], and it is convenient to record the attack result. (2) When the time of flooding the malicious IP address reaches the set attack duration, restart the victim node and prevent it from establishing a good connection. When the attacking host occupies all the victim's outgoing connections, the victim node state is rolled back to before restarting and repeated the above restart operation to record multiple sets of experimental data.

5.2. Attack Setup. We evaluate the attack effect of BHE under different attack configurations: (1) The number of malicious IPs, which refers to how many malicious IP prefixes does the attacker collected to fill the routing table of the victim node. (2) The age of the victim node, which refers to the number of running days since the victim first ran.

The validity period of IP addresses in the internal peer database of the bitcoin node is 30 days, and we selected the 4 most representative nodes for experiments (running for 0 days, 10 days, 20 days, and 30 days). We found that by running the node on the public network many times when the node runs for some time (more than a week), its internal database will save IP addresses with various prefixes; that is, for the latter three nodes, it can also make the attack successful without flooding the victim with malicious IPs (experimentally proved this). Research has proved that one hundred IP prefixes can control the victim's internal routing table very well [15]. Therefore, in the experiment, the number of malicious IP prefixes ranges from 0 to 100, and finally, four representative groups of data were selected for analysis.

5.3. Experimental Results and Analysis

5.3.1. The Impact of Node Age on Attack Efficiency. Figure 3 shows some experimental data under 100 malicious IP prefix configurations, which describes the attack efficiency of the attack against nodes of different ages under different attack duration. The attack duration is the duration of filling the malicious IP to the victim node before the victim node restarts. The attack efficiency refers to the time required to monopolize all outgoing connections of the victim node after the victim node restarts. The data in Figure 3 shows that with the increase of attack duration, the attack efficiency on all nodes is increasing, but under all attack duration, the attack effect on nodes running for 0 days is the best (except when the attack duration is 0, because in this case the node does not store the IP and cannot complete the connection), followed by nodes running 30 days (our script intercepts and deletes the IPs returned by the DNS seed so that the node can only receive malicious IPs in the attack). This is because more malicious IPs will be inserted into the peer database of these two nodes under the same attack duration. Furthermore, since the peer databases of nodes running 0 days only store malicious IPs (our experiments show that malicious IPs account for nearly 100%), the attack works best on nodes running 0 days with little variation over time (except the case where the attack duration is 0).

We experimentally find that for a node running for 0 days, an attacker can always monopolize all its connections with any number of malicious IP prefixes (less than 0.5 minutes) in a very short period of time. This is because, relative to other nodes, the database of 0-day nodes lacks legitimate IPs to compete with malicious IPs. In fact, when the node runs for about 10 days, the IP (legal IP) in the database will tend to a stable value, which we call a strong node. Our subsequent experimental analysis is mainly performed on strong nodes (10, 20, and 30 days nodes), because nonstrong nodes (0-day nodes) can always be easily attacked (not dependent on the number of malicious IPs), so it is difficult to show the experimental law under different attack conditions (for example, the more malicious IP prefixes, the higher the attack efficiency).

To better understand the impact of node age on attack efficiency, this article demonstrates the process of filling the routing table of each node with malicious IP in Figure 4. In general, the number of malicious IPs in the routing table of each node is increasing. However, the insertion rate of malicious IP is very slow for nodes running for 10 and 20 days, while nodes running for 30 days are easy to be filled by malicious IP. This is because many IP addresses in the routing table of nodes running for about 30 days are marked as terrible, making it easier for malicious IP to insert into the routing table.

5.3.2. The Impact of the Number of Malicious IP Prefixes on Attack Efficiency. Table 1 describes the relationship between the optimal attack efficiency and the number of malicious IP prefixes. The optimal attack efficiency refers to the time threshold at which the time required to monopolize all connections no longer decreases significantly as the attack duration increases. Taking the data in Figure 3 as an example, when the attack duration reaches about 25 minutes, the attack reaches the optimal efficiency, and the optimal attack efficiency of the attack on a node that has been running for 30 days is within 1 minute. The data in Table 1 shows that with the reduction of the number of malicious IPs, the attack efficiency on all types of nodes will reduce accordingly. This is because the smaller the number of malicious IP prefixes, the fewer malicious IPs will be inserted eventually. Among them, nodes running for 30 days are least affected by this rule, because this type of node is more likely to be populated by malicious IPs, so the nodes can also be well-populated while the malicious IP prefix decreases.

5.3.3. Best Attack Strategy. We call the attack duration that reaches the optimal attack efficiency for the first time for the optimal attack duration. After reaching the optimal attack duration, the attack efficiency will not change significantly as the attack duration increases. It is not difficult to see from the data in Figure 3 that 25 minutes is the optimal attack duration for that configuration.

The optimal attack strategy refers to the optimal attack duration and attack efficiency combination. When the attack

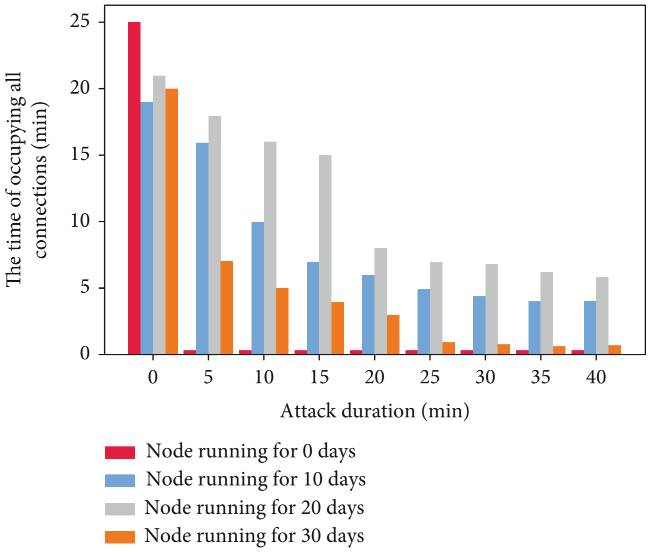


Figure 3: The relationship between attack duration and attack efficiency. Under the same attack time, the attack has the best effect on nodes that have been running for about 30 days.

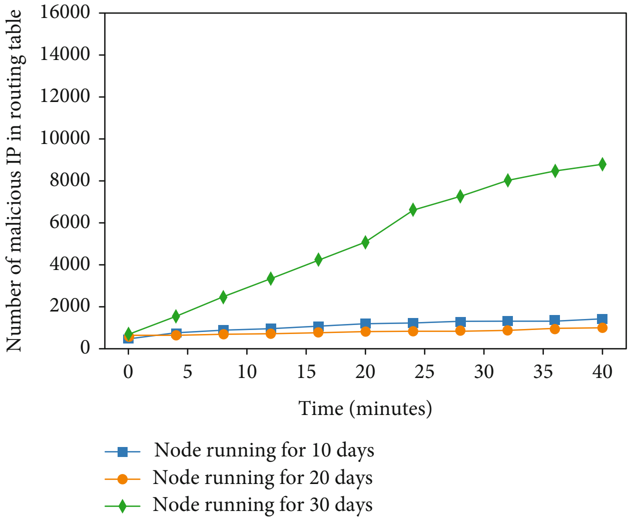


Figure 4: The number of malicious IP addresses in the routing table of nodes running for 30 days increased faster than those running for 10 days and 20 days.

Table 1: For the same type of node, the more malicious IP prefixes, the higher the optimal attack efficiency (the shorter the required attack time). The nodes running for 30 days are the least affected by this rule.

Node type	Number of malicious IP prefixes	Optimal attack efficiency(min)
Node running for 30 days	25	<6 min
	50	<5.5 min
	75	<2.5 min
	100	<1 min
Node running for 20 days	25	<17 min
	50	<14 min
	75	<7 min
	100	<6 min
Node running for 10 days	25	<13 min
	50	<9.5 min
	75	<5.5 min
	100	<5 min

achieves the optimal attack efficiency, it can monopolize all the network connections of the victim node in the shortest time. At this time, it is most difficult to be found by the victim node. The optimal attack efficiency will be obtained when the attack duration reaches the optimal attack duration. Even if the attack duration increases, the attack efficiency will not significantly improve. Therefore, the optimal attack strategy = optimal attack duration + optimal attack efficiency, when filling the victim node with malicious IP reaches the optimal attack duration of the current configuration, restart the node and occupy its outgoing connection. Take the data in Figure 3 as an example. When the attack duration reaches 25 minutes, stop filling malicious IP, launch a denial-of-service attack to restart the victim node, and occupy all its connections.

6. Countermeasure

In this section, we introduce some countermeasures against the BHE attack.

6.1. Add Connection within AS. The attack is based on the fact that BGP hijacking can intercept all network connections of the victim node, but BGP hijacking cannot intercept the traffic inside the AS to which the victim node belongs. Therefore, Bitcoin nodes can add an additional connection, which actively connects to the IP in the target AS and can give the node a view of the blockchain first. In this way, even if the attacker controls all other connections, it is not easy to control the network view of the victim node.

6.2. Diverse Connections. The nodes inside the mining pool use private protocols to connect, and it is difficult for attackers to capture this traffic. Therefore, when Bitcoin nodes join the bitcoin network, they can join some organizations (such as mining pools) and use private protocols for information exchange, which can effectively resist BHE attacks.

6.3. Replace the Port. The attacker can accurately identify the bitcoin traffic of the victim node in a large amount of traffic because most of the bitcoin traffic selects port 8333 by default. Therefore, the bitcoin node can set up several more ports for other nodes to connect, significantly increasing the difficulty for attackers to filter traffic.

6.4. Record the Time Interval between the Arrival of Adjacent Blocks. The purpose of the BHE attack is the same as the eclipse attack, which is to prevent the victim node from receiving the latest block or receiving the wrong block. It takes longer than normal for the victim to receive the next block [30]. Research has proved that under normal circumstances, the node will receive the next block within 40 minutes, but in the case of an attack, it is significantly higher than this value [30]. Therefore, Bitcoin can detect the attack by recording the time interval between the arrival of adjacent blocks.

6.5. Observe the Restart Time of Node. When a node is restarted after being attacked by BHE, it takes longer to establish 8 outgoing connections than a normal node because the node will experience many failures. Our experimental data shows that under normal circumstances, the node establishes all its outgoing connections within 5 minutes, but in the case of being attacked, most of time it exceeds this value (such as the nodes running for 10 days and 20 days in Figure 3). Therefore, we can judge whether the node is attacked by observing the time required for the node to establish eight outgoing connections after a restart. If the restart time significantly exceeds the average restart time of the node, the node is considered to be under attack. However, if the attacker collects enough malicious IPs and adopts the optimal attack strategy, this method will not apply.

7. Conclusion

This paper presents a new attack on the bitcoin network—BHE. This attack can control all network connections of the victim node in a short time. An attacker can use this attack to increase his mining advantage or launch a traditional blockchain attack on the victim node or target mining pool, which destroys the original intention of bitcoin design. The paper implemented the attack model in our experimental environment and proved its powerful destructive power. At the end of this paper, the paper gives some practical measures to mitigate the harm of attack.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

The authors would like to thank the anonymous reviewers for their elaborate reviews and feedback. This work was supported by the National Natural Science Foundation of China (No. 61906099) and the Open Fund of Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Natural Resources (No. KF-2019-04-065).

References

[1] C. G. Akcora, Y. R. Gel, and M. Kantarcioglu, “Blockchain networks: data structures of Bitcoin, Monero, Zcash, Ethereum, Ripple, and Iota,” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 12, no. 1, p. e1436, 2022.

[2] S. Nakamoto, “Bitcoin: a peer-to-peer electronic cash system,” *Decentralized Business Review*, no. article 21260, 2008.

[3] X. Han, Y. Yuan, and F. Y. Wang, “Security problems on blockchain: the state of the art and future trends,” *Acta Automatica Sinica*, vol. 45, no. 1, pp. 206–225, 2019.

[4] M. Apostolaki, A. Zohar, and L. Vanbever, “Hijacking bitcoin: routing attacks on cryptocurrencies,” in *2017 IEEE Symposium on Security and Privacy (SP)*, pp. 375–392, San Jose, CA, USA, 2017.

[5] F. Franzoni, X. Salleras, and V. Daza, “AToM: active topology monitoring for the bitcoin peer-to-peer network,” *Peer-to-Peer Networking and Applications*, vol. 15, no. 1, pp. 408–425, 2022.

[6] L. Zhang, T. Wang, and S. C. Liew, “Speeding up block propagation in Bitcoin network: uncoded and coded designs,” *Computer Networks*, vol. 206, article 108791, 2022.

[7] W. Yue and L. Junxiang, “The evolution process of blockchain P2P network protocol,” *Computer Application Research*, vol. 36, no. 10, pp. 2881–2886, 2019.

[8] S. Secci, J. L. Rougier, A. Pattavina, F. Patrone, and G. Maier, “Peering equilibrium multipath routing: a game theory framework for internet peering settlements,” *IEEE/ACM Transactions on Networking*, vol. 19, no. 2, pp. 419–432, 2011.

[9] H. S. Alotaibi, M. A. Gregory, and S. Li, “Multidomain SDN-based gateways and border gateway protocol,” *Journal of Computer Networks and Communications*, vol. 2022, 23 pages, 2022.

[10] H. Ballani, P. Francis, and X. Zhang, “A study of prefix hijacking and interception in the internet,” *ACM SIGCOMM Computer Communication Review*, vol. 37, no. 4, pp. 265–276, 2007.

[11] E. Heilman, A. Kendler, A. Zohar, and S. Goldberg, “Eclipse attacks on bitcoin’s peer-to-peer network,” in *24th {USENIX} Security Symposium ({USENIX} Security 15)*, pp. 129–144, Washington,D.C., August 2015.

[12] A. Gervais, H. Ritzdorf, G. O. Karame, and S. Capkun, “Tampering with the delivery of blocks and transactions in bitcoin,”

in *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, pp. 692–705, Denver, Colorado, USA, October 2015.

[13] M. Walck, K. Wang, and H. S. Kim, "Tendril staller: block delay attack in Bitcoin," in *2019 IEEE international conference on Blockchain (Blockchain)*, pp. 1–9, Atlanta, GA, USA, 2019.

[14] A. E. Yves-Christian, B. Hammi, A. Serhrouchni, and H. Labiod, "Total eclipse: how to completely isolate a bitcoin peer," in *2018 Third International Conference on Security of Smart Cities, Industrial Control System and Communications (SSIC)*, pp. 1–7, Shanghai, China, 2018.

[15] M. Tran, I. Choi, G. J. Moon, A. V. Vu, and M. S. Kang, "A stealthier partitioning attack against bitcoin peer-to-peer network," in *2020 IEEE symposium on security and privacy (SP)*, pp. 894–909, San Francisco, CA, USA, 2020.

[16] J. Marçal, L. Rodrigues, and M. Matos, "Adaptive information dissemination in the bitcoin network," in *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*, pp. 276–283, St. Raphael Resort, Limassol, Cyprus, April 2019.

[17] G. Naumenko, G. Maxwell, P. Wuille, A. Fedorova, and I. Beschastnikh, "Erlay: efficient transaction relay for bitcoin," in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, pp. 817–831, London,-United Kingdom, November 2019.

[18] K. Otsuki, Y. Aoki, R. Banno, and K. Shudo, "Effects of a simple relay network on the bitcoin network," in *Proceedings of the Asian Internet Engineering Conference*, pp. 41–46, Andaman Cannacia Resort and Spa Hotel, Kata Beach, Phuket, Thailand., August 2019.

[19] Q. Dai, B. Zhang, and S. Dong, "Eclipse attack detection for blockchain network layer based on deep feature extraction," *Wireless Communications and Mobile Computing*, vol. 2022, 19 pages, 2022.

[20] Q. Dai, B. Zhang, and S. Dong, "A DDoS-attack detection method oriented to the blockchain network layer," *Security and Communication Networks*, vol. 2022, 18 pages, 2022.

[21] C. Pinzón and C. Rocha, "Double-spend attack models with time advantange for bitcoin," *Electronic Notes in Theoretical Computer Science*, vol. 329, pp. 79–103, 2016.

[22] G. O. Karame, E. Androulaki, and S. Capkun, "Double-spending fast payments in bitcoin," in *Proceedings of the 2012 ACM conference on Computer and communications security*, pp. 906–917, Raleigh, NC, USA, October 2012.

[23] G. O. Karame, E. Androulaki, M. Roeschlin, A. Gervais, and S. Čapkun, "Misbehavior in bitcoin," *ACM Transactions on Information and System Security (TISSEC)*, vol. 18, no. 1, pp. 1–32, 2015.

[24] Q. Bai, X. Zhou, X. Wang, Y. Xu, X. Wang, and Q. Kong, "A deep dive into blockchain selfish mining ICC," in *2019-2019 IEEE International Conference on Communications (ICC)*, pp. 1–6, Shanghai, China, 2019.

[25] A. Sapirshtein, Y. Sompolinsky, and A. Zohar, "Optimal selfish mining strategies in bitcoin," in *International Conference on Financial Cryptography and Data Security*, pp. 515–532, Berlin, Heidelberg, 2017.

[26] Z. Wang, Q. Lv, Z. Lu, Y. Wang, and S. Yue, "ForkDec: accurate detection for selfish mining attacks," *Security and Communication Networks*, vol. 2021, 8 pages, 2021.

[27] S. Solat and M. Potop-Butucaru, "Zeroblock: preventing selfish mining in bitcoin," 2016, https://arxiv.org/abs/1605.02435.

[28] M. S. Rahman, M. Y. Uddin, T. Hasan, M. S. Rahman, and M. Kaykobad, "Using adaptive heartbeat rate on long-lived TCP connections," *IEEE/ACM Transactions on Networking*, vol. 26, no. 1, pp. 203–216, 2018.

[29] Y. Xiang and D. Loker, "Trans-causalizing NAT-modeled Bayesian networks," *IEEE Transactions on Cybernetics*, vol. 52, no. 5, 2020.

[30] B. Alangot, D. Reijsbergen, S. Venugopalan, and P. Szalachowski, "Decentralized lightweight detection of eclipse attacks on bitcoin clients," in *2020 IEEE international conference on Blockchain (Blockchain)*, pp. 337–342, Rhodes, Greece, 2020.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 5866570, 14 pages
https://doi.org/10.1155/2022/5866570

Research Article

Cloud Platform Credibility Assessment System Based on D-S Theory and Blockchain Technology

Ming Yang, Li Jia, Tilei Gao, Yuanyuan He, Bin Gui, and Tao Zhang

School of Information, Yunnan University of Finance and Economics, Kunming 650221, China

Correspondence should be addressed to Tilei Gao; gtllei@ynufe.edu.cn

Received 17 May 2022; Revised 1 July 2022; Accepted 7 July 2022; Published 1 August 2022

Academic Editor: Yin Zhang

Even well-known cloud platforms will have sudden credibility problems in the long-term application process. Effectively evaluating the credibility of the cloud platform and providing users with scientific evaluation results can help users reasonably choose a trusted cloud platform. However, there are often conflicting opinions or malicious assessments in the process of assessment. In addition, the personal privacy information of the users participating in the assessment is at risk of being leaked, and the data that the users have evaluated is also easy to be modified. In order to solve the above problems, this paper defines the credibility category and confidence interval of cloud platform, puts forward a quantitative assessment method combined with fuzzy theory, and realizes the fusion of different users' assessment results based on D-S theory. On this basis, this paper further proposes an effective cloud platform credibility assessment system combined with blockchain technology. Finally, through experimental analysis, this paper shows that the credibility assessment system proposed in this paper is feasible and illustrates the characteristics of the system through method comparison. The system solves the problem of conflicting information in the assessment process, can effectively assess the credibility of the cloud platform, and effectively protects user privacy and the security of assessment data with blockchain technology.

1. Introduction

According to "the first quarter 2020 global data center infrastructure revenue data" released by synergy research group, benefiting from the significant growth of cloud computing demand during the epidemic, the revenue of the global cloud computing market in the first quarter increased by 37% year-on-year. According to "2020 cloud status report" published by Flexera [1], 59% of enterprises expect cloud usage to exceed previous plans. The above report shows that the demand for cloud services in the global market is gradually increasing. However, due to the cloud platform characteristics such as improper management, complex network transmission, huge data storage demand, large number of tenants, and diverse services of cloud platform, there are large credibility problems in the cloud platform. According to the report of Amazon which is the largest cloud computing provider, its company's cloud platform and services had 22 sudden failures during 2010-2019. The report shows that even well-known platforms will have credibility problems. Therefore, when choosing a cloud platform, users need to understand the credibility of the platform. The most effective way is to refer to the comments of users who have used it. However, in the absence of effective evaluation methods and tools, the value of the assessment results given by users who do not have professional knowledge will be greatly reduced. In addition, when users participate in the assessment process, there is bound to be the problem of privacy information being leaked, and the users' assessment results will also have the possibility of being tampered with or deleted. Therefore, in order to effectively assess the cloud platform credibility and give scientific assessment results, it is necessary to establish special assessment systems, methods, and tools.

Shen [2] pointed out that credibility includes reliability and safety. Yang [3] pointed out that credibility includes ability credibility, integrity credibility, predictability, correctness, privacy, and loss cost. As a kind of credibility evaluation, the credibility evaluation process of cloud platform is bound to be affected by human subjective factors [4]. In the process of assessment, due to the influence of human subjective factors, conflict information is bound to appear. In addition, in the process of assessment, it is difficult to give an accurate credibility assessment result due to the influence of users or experts' own complex psychology.

Therefore, how to ensure the objectivity of the assessment, solve the problem of conflicting information in the assessment process, and reduce the scoring difficulty of users are the problems that need to be solved to realize the credibility assessment of cloud platform. In order to ensure the objectivity, relevant studies at home and abroad include the assessment method based on AHP (Analytic Hierarchy Process) [5–11] and the uncertainty assessment method based on information entropy [12–16]. These relevant studies establish an effective credibility assessment system, and realize the quantitative assessment of multi-index system through pairwise comparison, which effectively reduces the impact of human subjective factors on the assessment results. In order to solve the conflict information in the assessment process, scholars at home and abroad have carried out many studies based on D-S evidence theory [17–22]. These related studies point out that using D-S fusion method can effectively solve the problem of conflicting information in the assessment process. In order to ensure the accuracy of the assessment results and reduce the scoring difficulty of users in the assessment process, Wang et al. [23, 24] proposed effective solutions based on fuzzy theory. It can be seen that the comprehensive use of the above methods will effectively solve the problems existing in the cloud platform credibility assessment.

However, in addition to solving the above problems, the cloud platform credibility assessment also faces the problems of privacy security and how to ensure data integrity. It is known that when participating in the assessment, users will leave relevant transaction information and personal information. This leads to the risk that the user's privacy will be stolen or leaked. In order to protect the privacy information of users during evaluation, Shi [25] and Yang [26] both proposed an assessment mechanism based on blockchain, which effectively protects the privacy of users participating in assessment through blockchain technology and can trace the responsibility of malicious users through blockchain traceability technology [27]. In addition, the tamper-proof characteristics of blockchain can also ensure the integrity of assessment data and provide users with continuous and real assessment results in time.

Therefore, in order to realize the effective cloud platform credibility assessment, this paper comprehensively uses the above-mentioned methods to carry out the analysis. Firstly, combined with fuzzy theory, this paper defines the credibility category of cloud platform and its corresponding confidence interval, puts forward the assessment method of cloud platform credibility, and realizes the fusion of different users' assessment results based on D-S theory. On this basis, in order to ensure the privacy of users participating in the evaluation and ensure that the generated evaluation results cannot be tampered with, this paper combines the blockchain technology with the proposed credibility assessment method, proposes an effective assessment block generation method, designs the corresponding consensus mechanism, smart contract and incentive mechanism, and finally proposes a credibility evaluation system based on blockchain technology and D-S theory. The system integrates the characteristics of blockchain technology and D-S theory and provides an effective scheme for the cloud platform credibility assessment.

This paper can be divided into the following parts: in Section 1, this paper introduces the research background and content; in Section 2, the credibility category and confidence interval of cloud platform are defined based on fuzzy theory, and a credibility assessment result fusion method based on D-S theory is proposed to realize the effective evaluation of cloud platform credibility; in Section 3, based on the proposed cloud platform credibility assessment method, this paper further proposes a cloud platform credibility assessment system combined with blockchain technology; in Section 4, in order to verify the effectiveness of the proposed credibility evaluation system, this paper carries out relevant experimental analysis and compares the proposed assessment method with other methods in many aspects; in Section 5, the authors summarize the research work of this paper and point out the future research direction.

2. Cloud Platform Credibility Assessment Method Based on D-S Evidence Theory

Cloud platform generally refers to cloud service platform, which provides users with computing, network, and storage capabilities through distributed processing technology. Because the cloud platform has the characteristics such as large number of tenants, huge data storage demand, complex network transmission, and diverse service functions, its credibility will be affected by many factors in the actual application process, as shown in Figure 1.

These factors include infrastructure credibility, service function credibility, network credibility, service provider management credibility, and platform internal environment credibility. Their meanings and examples are shown in Table 1.

Therefore, to assess the credibility of cloud platform, we need to focus on the credibility category β_i described in Table 1 and carry out comprehensive assessment from multiple aspects.

For example, the infrastructure credibility in Table 1 can be judged by users through the infrastructure information published by the platform. Common information includes number of global acceleration nodes, number of servers, number of data centers, and coverage areas. Users can make basic judgments and give scores through this information. In addition, with the operation of the cloud platform, users who have participated in the assessment can also add scores according to relevant reports or infrastructure failure problems during use. If the platform does not publish the relevant infrastructure information and the user cannot obtain the infrastructure information of the platform, the user can consider the platform's infrastructure credibility as untrusted.

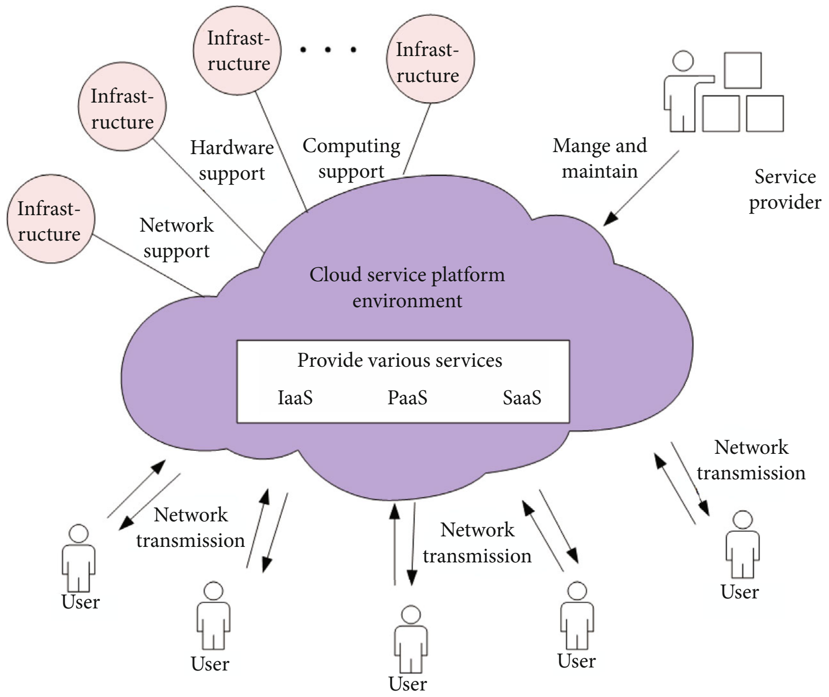


Figure 1: Multiple factors affecting the credibility of cloud platform.

Table 1: Cloud platform trustworthiness category.

β_i	Credibility category	Meaning	Example of credibility problems
β_1	Infrastructure credibility	It refers to the credibility of the platform physical infrastructure.	Such as dilapidated infrastructure, damaged infrastructure, and data center disaster
β_2	Service function credibility	It refers to the credibility of the platform service in terms of function.	Such as lack of function, poor usability, and difficulty in function expansion
β_3	Network credibility	It refers to the credibility of the platform in terms of network transmission.	Such as unstable network transmission, lack of effective network defense support, and vulnerable to DDoS attack or CC attack
β_4	Service provider management credibility	It refers to the credibility of the platform provider in the management of platform environment, services, infrastructure, network, etc.	Such as untimely maintenance and updating and no clear responsibility attribution agreement
β_5	Platform internal environment credibility	It refers to the credibility of the platform internal environment.	Such as the mandatory function of the platform and the attack of other cloud tenants in the same platform

Next, this paper will focus on these 5 credibility categories β_i and put forward effective assessment methods from the perspective of users.

2.1. Confidence Interval of Cloud Platform and Its Assessment Method. It is known that for users who do not have professional knowledge, it is difficult to give an accurate assessment when evaluating the credibility of cloud platform. They can only give a general assessment according to their own use experience, that is, ordinary users can only give a vague assessment. Therefore, this paper will assess the credibility of cloud platform based on fuzzy theory.

Fuzzy theory [28] is based on Fuzzy Set, and its research goal is to deal with uncertain things with fuzzy concepts. Fuzzy Set refers to the set with uncertain boundaries. Since the cloud platform credibility is also a fuzzy concept that is difficult to describe, its credibility can be described by Fuzzy Set.

Firstly, according to the fuzzy theory, this paper sets 5 fuzzy confidence intervals of cloud platform, which are defined as shown in Table 2.

According to the division of Table 2, users can give a fuzzy assessment result according to their use experience. There are 5 possible arbitrary sets of the result, namely, $\{1, 2, 3\}$, $\{3, 4, 5\}$, $\{4, 5, 6, 7\}$, $\{6, 7, 8\}$, $\{8, 9, 10\}$. Among them, $A_3 = \{4, 5, 6, 7\}$ indicates that the credibility level of the cloud platform is between 4 and 7. The greater the credibility level, the more credible the cloud platform is.

As mentioned above, when judging the credibility of the cloud platform, users do not need to give an accurate value,

Table 2: The 5 fuzzy confidence intervals of cloud platform.

Confidence interval	Meaning	The arbitrary sets of credibility level
A_1 completely credible	The platform has few credibility problems and can be fully trusted.	$A_1 = \{8910\}$
A_2 more credible	During the use of the platform service, the credibility problem occasionally occurs.	$A_2 = \{678\}$
A_3 basically credible	The platform has potential credibility problems and belongs to a general trusted platform.	$A_3 = \{4567\}$
A_4 basically untrusted	The platform has obvious credibility problems and is basically untrusted.	$A_4 = \{345\}$
A_5 completely untrusted	The platform is completely untrusted.	$A_5 = \{123\}$

but only need to select one of the confidence intervals. Using the above methods can reduce the difficulty of scoring and obtain the effective assessment results given by users.

2.2. Assessment Result Fusion Method Based on D-S Theory. Next, after collecting the assessment results given by multiple users, these assessment results can be fused with the fusion rules of D-S theory, so as to obtain a more accurate credibility assessment result.

D-S evidence theory [21] is an uncertain reasoning method, which is often used for multi-information fusion. It can effectively deal with the problem of conflict information in the fusion process and fuse the relevant information through calculation.

Suppose that the assessment results given by user 1 and user 2 for a certain platform are shown in Table 3, and the fusion process is as follows.

Step 1: the trust degree $m(A_j)$ of cloud platform confidence interval.

Assessment 1 and Assessment 2 in Table 3 represent the assessment results of the two users, respectively. $m_i(A_j)$ represents the trust degree of A_j given by user i. The greater the value of $m(A_j)$, the greater the possibility that the credibility level of the cloud platform belongs to A_j. The calculation formula of $m(A_j)$ is as follows.

$$m(A_j) = \sum_{\forall \mathrm{lev}(\beta_i) \in A_j} W(\beta_i). \tag{1}$$

In formula (1), $\mathrm{lev}(\beta_i)$ indicates the confidence interval of β_j, and $\mathrm{lev}(\beta_j) \in A_i$ indicates that the user assesses the confidence interval of β_j as A_i.

$W(\beta_j)$ represents the assessment weight of the credibility category β_j, $\sum_{j=1}^{5} W(\beta_j) = 1$. The greater the value of $W(\beta_j)$, the greater the influence weight of credibility category β_j on the credibility of the whole cloud platform. In order to ensure the effectiveness of the assessment, this paper proposes to use the entropy weight method to update the weight value of each credibility category in real time. According to the entropy weight method [29], for a credibility category β_j, the greater the difference between the assessment results, the higher the value of $W(\beta_j)$. Conversely, the lower the value of $W(\beta_j)$. With the increase of user assessment data, the weight value $W(\beta_j)$ of each credibility category will gradually change.

Table 3: Assessment results given by two users.

The arbitrary set A_i	Assessment 1	Assessment 2
$A_5 = \{8910\}$	$m_1(A_5)$	$m_2(A_5)$
$A_4 = \{678\}$	$m_1(A_4)$	$m_2(A_4)$
$A_3 = \{4567\}$	$m_1(A_3)$	$m_2(A_3)$
$A_2 = \{345\}$	$m_1(A_2)$	$m_2(A_2)$
$A_1 = \{123\}$	$m_1(A_1)$	$m_2(A_1)$

Table 4: The 5 credibility category assessments given by user 1 and user 2.

Credibility category β_j	User 1's assessment of $\mathrm{lev}\left(\beta_j\right)$	User 2's assessment of $\mathrm{lev}\left(\beta_j\right)$
β_1	A_5	A_4
β_2	A_4	A_4
β_3	A_4	A_4
β_4	A_4	A_3
β_5	A_3	A_3

As shown in the following example, assume that the assessment weights of the 5 credibility categories of a cloud platform are equal, $W(\beta_j) = 0.200, j = 1, 2, \cdots, 5$. The 5 credibility category assessments given by user 1 and user 2 are shown in Table 4.

By substituting the user assessment data of Table 4 into formula (1) for calculation, the trust degree $m_i(A_j)$ of the cloud platform's confidence interval can be obtained. The results are as follows.

$$m_1(A_5) = W(\beta_1) = 0.200, m_2(A_5) = 0.000, \tag{2}$$

$$m_1(A_4) = \sum_{j=2}^{4} W\left(\beta_j\right) = 0.600, m_2(A_4) = \sum_{j=1}^{3} W\left(\beta_j\right) = 0.600, \tag{3}$$

$$m_1(A_3) = W(\beta_5) = 0.200, m_2(A_3) = \sum_{j=4}^{5} W\left(\beta_j\right) = 0.400, \tag{4}$$

$$m_1(A_2) = 0.000, m_2(A_2) = 0.000, \tag{5}$$

$$m_1(A_1) = 0.000, m_2(A_1) = 0.000. \tag{6}$$

The above results represent the trust degree of the cloud platform's confidence interval, $m_1(A_j)$ represents the assessment result given by user 1, $m_2(A_j)$ represents the assessment result given by user 2, and the two users give different assessment results, respectively. Next, in order to integrate the views of the two users, this paper will fuse the assessment results of the two users combined with the fusion rules of D-S theory.

Step 2: fuse different assessment results based on D-S fusion rules.

Taking the data of Table 3 as an example, in order to reduce the complexity of calculation before fusion, set A_j and its trust $m_i(A_j)$ can be simplified according to Bayes approximation method [30], and the calculation method is shown in

$$m_i(\underline{A}) = \frac{\sum_{\underline{A}\subseteq A} m_i(A)}{\sum_{A\subseteq\Theta} m_i(A) * N}. \tag{7}$$

In formula (7), $\underline{A}$ is the simplified set of A, Θ Represents the complete set, and N is the total number of factors contained in A. As described above, the data in Table 3 can be substituted into formula (2) for calculation, and its calculation process is as follows.

$$\begin{aligned}
\sum_{A\subseteq\Theta} m_i(A) * N &= 3m_i(A_1) + 3m_i(A_2) + 4m_i(A_3) \\
&\quad + 3m_i(A_4) + 3m_i(A_5), \\
m_i(\underline{1}) = m_i(\underline{2}) &= \frac{m_i(A_1)}{\sum_{A\subseteq\Theta} m_i(A) * N}, \\
m_i(\underline{3}) &= \frac{m_i(A_1) + m_i(A_2)}{\sum_{A\subseteq\Theta} m_i(A) * N}, \\
m_i(\underline{4}) = m_i(\underline{5}) &= \frac{m_i(A_2) + m_i(A_3)}{\sum_{A\subseteq\Theta} m_i(A) * N}, \\
m_i(\underline{6}) = m_i(\underline{7}) &= \frac{m_i(A_3) + m_i(A_4)}{\sum_{A\subseteq\Theta} m_i(A) * N}, \\
m_i(\underline{8}) &= \frac{m_i(A_4) + m_i(A_5)}{\sum_{A\subseteq\Theta} m_i(A) * N}, \\
m_i(\underline{9}) = m_i(\underline{10}) &= \frac{m_i(A_5)}{\sum_{A\subseteq\Theta} m_i(A) * N}.
\end{aligned} \tag{8}$$

Through the above calculation method, the simplified $\underline{A}$ and its trust degree $m_i(\underline{A})$ can be obtained. Among them, $\underline{A}$ is the simplified set of A. $\underline{A}$ is different from set A, it contains only one element. $\underline{A} = \{\underline{A}_1, \underline{A}_2, \cdots, \underline{A}_{10}\} = \{\underline{1}, \underline{2}, \cdots, \underline{10}\}$.

As mentioned above, after simplifying the assessment results of two users in Table 3, $m_1(\underline{A}_j)$ and $m_2(\underline{A}_j)$ can be obtained, $j = 1, 2, 10$. Next, by substituting them into the fusion formula of D-S theory shown in formula (9), the final fusion result can be obtained.

$$m(\underline{A}) = (m_1 \oplus m_2)(\underline{A}) = \frac{1}{k} \sum_{A_i \cap A_j = A} m_1(\underline{A}_i) m_2\left(\underline{A}_j\right). \tag{9}$$

In formula (9), K is the normalization factor, and its calculation method is shown in

$$k = \sum_{A_i \cap A_j \neq \emptyset} m_1(\underline{A}_i) m_2\left(\underline{A}_j\right). \tag{10}$$

The final fusion results are shown in Table 5.

In Table 5, $m(\underline{6}) = m(\underline{7}) = 0.377$, and the values of $m(6)$ and $m(7)$ are the largest, indicating that the cloud platform credibility level is most likely to be 6 and 7.

2.3. Method Improvement. Through the above method, the two results can be fused to update the current credibility assessment results of the cloud platform. However, this method of pairwise integration has defects in the actual assessment process. As shown in Table 5, in the process of fusion, if the value of $m_j(\underline{A}_1)$ given by user j is equal to 0, the value of the fusion result $m(\underline{A}_1)$ will always be equal to 0 in all subsequent fusion processes. This situation is not consistent with the actual assessment. Therefore, this paper will improve the above method. The improved method is as follows.

Step 1: add the complete set U on the basis of Table 2.

U is the complete set of cloud platform credibility level, $U = \{1, 2, \cdots, 10\}$. It contains all possible values of cloud platform credibility level. In order to solve the problem mentioned at the beginning of this section, this paper sets the value of $m_j(U)$to the average value, that is, $m_j(U) = 0.1$.

Step 2: recalculate $m_j(\underline{A}_i)$ according to formula (7).

When the complete set U is added, according to the method of D-S theory, the value of $m_j(\underline{A}_i)$ needs to be recalculated before fusion. Taking the data of Table 2 as an example, when the complete set U is introduced, $m_j(\underline{A}_i)$ is recalculated according to formula (7). The results are shown in Table 6.

Step 3: refuse users' assessment results according to formula (9).

After obtaining the assessment result $m_1(\underline{A}_i)$ and $m_2(\underline{A}_i)$ of the two users according to step 2, fuse them according to formula (9), and the obtained results are shown in Table 6.

As shown in Table 6, the results of the improved method are consistent with those before the improvement. $m(\underline{6})$ and $m(\underline{7})$ are still the largest, indicating that the cloud platform credibility level is most likely to be 6 and 7. On the premise of ensuring the correctness of the results, this method retains all possibilities of the cloud platform credibility level, so as to

Table 5: The fusion results of two assessment results in Table 3.

$\underline{A}$	$m_1(\underline{A}_i)$	$m_1(\underline{A}_i)$	$m(\underline{A})$
{10}	0.063	0.000	0.000
{9}	0.063	0.000	0.000
{8}	0.250	0.176	0.214
{7}	0.250	0.294	0.357
{6}	0.250	0.294	0.357
{5}	0.063	0.118	0.036
{4}	0.063	0.118	0.036
{3}	0.000	0.000	0.000
{2}	0.000	0.000	0.000
{1}	0.000	0.000	0.000

Table 6: The fusion results calculated by the improved method.

$\underline{A}$	$m_1(\underline{A}_i)$	$m_1(\underline{A}_i)$	$m(\underline{A})$
{10}	0.071	0.023	0.010
{9}	0.071	0.023	0.010
{8}	0.214	0.159	0.210
{7}	0.214	0.250	0.330
{6}	0.214	0.250	0.330
{5}	0.071	0.114	0.050
{4}	0.071	0.114	0.050
{3}	0.024	0.023	0.003
{2}	0.024	0.023	0.003
{1}	0.024	0.023	0.003

effectively solve the problems mentioned at the beginning of Section 2.3.

So far, based on D-S theory, this paper puts forward an effective assessment method for the credibility of cloud platform. This method has low requirements for users' professionalism and can effectively integrate the assessment results of different users, so as to solve the conflict information between different users in the assessment process. Although this method reduces the scoring difficulty of users and solves the problem of conflicting information in the evaluation process, this method still has many defects, such as the risk of user's privacy information exposure, the risk of the assessment results be changed, malicious users, and the low user assessment enthusiasm. In order to solve these problems, this paper will make further study combined with blockchain technology and integrate blockchain technology and D-S theory to improve the assessment method.

3. Design of Cloud Platform Credibility Assessment System Based on Blockchain Technology

In order to realize the system, based on the architecture of Ethereum, this paper will combine the blockchain technology with the assessment method proposed in Section 2 to establish a cloud platform credibility assessment system.

It is known that there are 6 layers in Ethereum structure. The 6 layers from bottom to top are data layer, network layer, consensus layer, actuator layer, contract layer, and application layer, as shown in Figure 2.

Among them, the application layer refers to the application scenario of blockchain technology. In this paper, it refers to the assessment of the cloud platform credibility. Like most blockchains, the network layer of the system to be established in this paper adopts a typical P2P network, including data dissemination and verification. Therefore, in order to integrate blockchain technology into the research process of this paper, in addition to the above application layer and network layer, it is also necessary to clarify the meaning of the other 4 layers and put forward effective construction schemes for these 4 layers.

3.1. Data Layer. The data layer refers to the data structure in the blockchain, that is, the "block + chain" structure. In order to ensure the privacy security of users participating in the assessment and ensure the assessment results cannot be modified, this paper intends to encrypt the corresponding assessment data with the encryption technology of blockchain, so as to generate the corresponding block, as shown in Figure 3.

The block header includes the hash value "PreHash" of the previous block, the Merkle root generated by the assessment data contained in this block after layer-by-layer encryption, timestamp, and the random parameter "Nonce" of workload proof. The block body stores the detailed data of this block. In this study, it refers to the assessment data of the cloud platform credibility. The assessment data is composed of the unique address of the user, the trust degree $m(\underline{A})$ of the cloud platform credibility level, and the assessment weight $W(\beta_i)$ of the 5 credibility categories of the cloud platform.

Example: a user's address is 0x6c19a33efc41a1beddc91133a8422e89f041b7, the assessment weight $W(\beta_i) = \{0.200,0.200,0.200,0.200,0.200\}$, and the trust degree of the cloud platform credibility level obtained by the assessment method proposed in Section 2 is $m(\underline{A}) = \{0.000,0.000,0.176,0.294,0.294,0.118,0.000,0.000,0.000,0.000\}$. According to the encryption method of blockchain, this series of values can be spliced together into a string, which is recorded as Assessment Data1. Next, encrypt the Assessment Data1, we can get the encrypted value, namely, Hash1. Similarly, we can get another encrypted value Hash2 generated by another user's assessment data. Then, the Merkle root can be obtained by encrypting Hash1 and Hash2. The Merkle root can be used to verify the data contained in the block and ensure that the block data cannot be modified.

According to the privacy protection technology of Ethereum blockchain, the nonpublic data of the block can only be viewed by the data owner. Therefore, compared with the scoring method proposed in Section 2, the combination of blockchain technology can further effectively protect the user's hidden information.

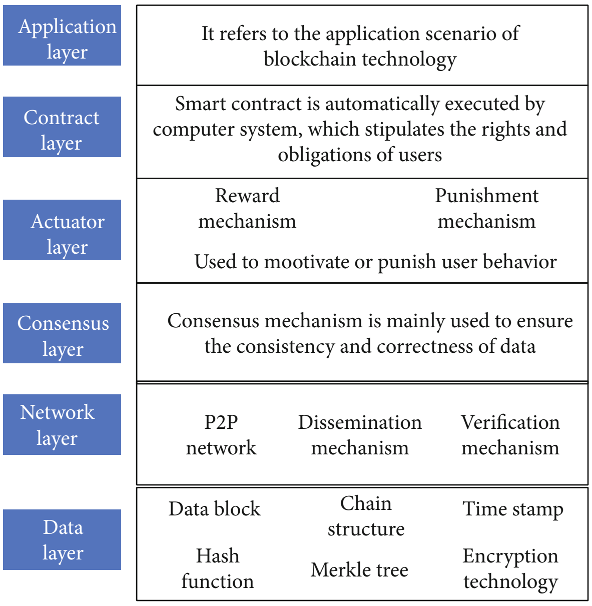

Application layer
It refers to the application scenario of blockchain technology
Contract layer
Smart contract is automatically executed by computer system, which stipulates the rights and obligations of users
Actuator layer
Reward mechanism
Punishment mechanism
Used to mootivate or punish user behavior
Consensus layer
Consensus mechanism is mainly used to ensure the consistency and correctness of data
Network layer
P2P network
Dissemination mechanism
Verification mechanism
Data layer
Data block
Chain structure
Time stamp
Hash function
Merkle tree
Encryption technology

Figure 2: The 6 layers of blockchain.

3.2. Consensus Layer. Consensus mechanism is mainly to solve the problem of data consistency and correctness in unreliable networks. If there is no consensus mechanism, whoever calculates the block can be regarded as an effective block, and the consistency of data cannot be guaranteed. Therefore, only the blocks that meet the requirements can be regarded as effective blocks, and then be added to the chain as new blocks.

At present, the consensus mechanism of Ethereum is a workload proof algorithm based on Ethash (consensus engine) [25]. The algorithm will set the target hash value in the block header. Only when the hash value of the new block is less than or equal to the target hash value, the block can be regarded as a valid block and added to the blockchain. The schematic diagram is shown in Figure 4.

If the hash value of the new block does not meet the conditions, the random number Nonce needs to be changed continuously until the hash value of the new block meets the conditions. For example, when calculating the hash value of a new block, we can return the value of Nonce to zero. When the calculated hash value does not meet the conditions, the value of Nonce can be incremented by 1 until the new block's hash value is less than or equal to the target hash value. The greater the value of Nonce, the greater the difficulty of calculation. Therefore, the workload can be proved according to the value of Nonce.

The cloud platform credibility evaluation system proposed in this paper is based on Ethereum architecture. According to the current Ethereum workload proof algorithm diff = 2^{256}/difficulty, the block is valid only when the hash value of the new block is less than diff. The greater the number of difficulty, the more difficult the calculation is and the slower the block output speed is. On the contrary, the smaller the value of difficulty, the lower the calculation difficulty and the faster the block output speed. Therefore, in order to ensure the calculation speed of the whole assessment system, the value of difficulty needs to be set to a smaller value in Genesis block.

3.3. Actuator Layer. In order to improve the enthusiasm of users and punish malicious users, the system will set up a special mechanism which includes reward mechanism and punishment mechanism. As described below, in order to support this mechanism, the system will set a certain number of initial reputation points for authenticated users, which can be used for scoring and query.

(i) Reward mechanism: when a new score is generated, all users in the system can participate in the calculation of new blocks. The system will reward the first user who successfully calculates the effective block and give the user a certain reputation point

(ii) Punishment mechanism: on the contrary, if the user has malicious behavior and is identified as a malicious user, the system will find the user and deduct a certain reputation point of the user according to the traceability method of the blockchain. When users' reputation point is insufficient, they will no longer be able to participate in the assessment.

The above incentive mechanism and punishment mechanism will be written into the smart contract and automatically executed by the system.

3.4. Contract Layer. Combined with the assessment method proposed in Section 2, this paper will set the smart contract of the system according to the reward mechanism and punishment mechanism. The execution process of the whole system is shown in Figure 5.

Several important smart contract functions are involved in the system, as shown below.

(1) Initial.sol: this function is mainly used to grant users the initial reputation point. When the user becomes a contract user and obtains the account address, the system will automatically perform the contract and remit a certain initial reputation point to the account address

(2) ScorePayment.sol: this function is mainly used to deduct a certain number of users' reputation points before scoring, and the deducted reputation points will be used as collateral. When the system judges that the user's reputation score is insufficient, according to the contract, the user will not be able to participate in the scoring. When the new effective block is be calculated, the user's mortgaged reputation points will be returned

(3) PayAndGetInfo.sol: this function is mainly used to deduct the user's query fee and return the queried block information to the user

(4) Reward.sol: this function is mainly used to reward users who successfully calculate new blocks. When

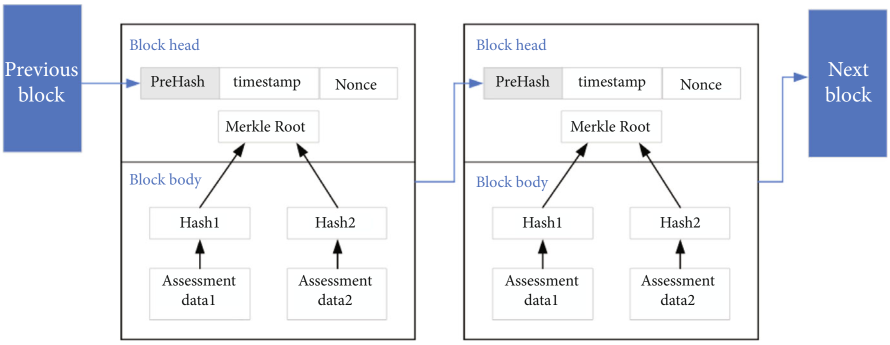


Figure 3: The blocks of the system proposed in this paper.

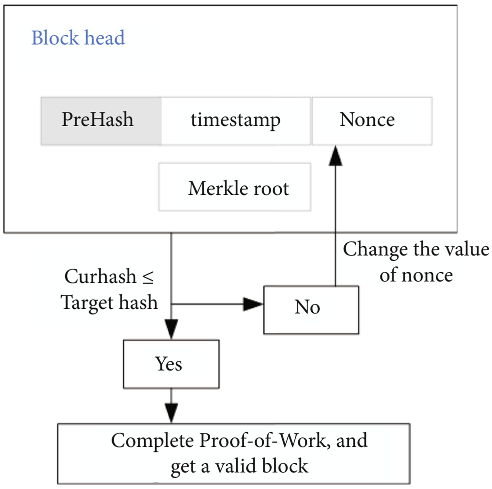

Block head
PreHash
timestamp
Nonce
Merkle root
Change the value of nonce
Curhash ≤ Target hash
No
Yes
Complete Proof-of-Work, and get a valid block

Figure 4: The schematic diagram of blockchain workload proof algorithm.

the user calculates a new effective block, the system will give the user a certain reputation point as a reward according to the reward mechanism

(5) Punish.sol: this function is mainly used to punish malicious users. When a user is identified as a malicious user, the system will trace back to the user through the traceability method of blockchain and deduct the user's reputation points according to the punishment mechanism

As mentioned above, this paper proposes a cloud platform credibility assessment system based on blockchain technology and D-S theory. The blockchain node of the system performs scoring operation through the smart contract and uses the privacy protection technology in blockchain technology to realize the anonymity of scoring process and protect the user's personal privacy.

At the same time, in order to prevent users from scoring maliciously on the platform, the system will deduct a certain number of users' reputation points as collateral before scoring according to the contract. When a new effective block is generated, the system will automatically return the users' mortgaged reputation points. In addition, with the help of blockchain traceability technology, the system can also find users or organizations with malicious behavior and punish them accordingly.

The consensus algorithm in the assessment system ensures the reliability of the blockchain system. After the blockchain nodes reach a consensus, the system will fuse the user's assessment results according to the credibility assessment method based on D-S theory proposed in Section 2, so as to update and record the credibility assessment results of the cloud platform. The result will be uploaded to the blockchain for users to access and query. The results include the current block address, the previous block address, the address of the user who participating in the assessment, the assessment date, the trust degree $m(\underline{A})$of the cloud platform credibility level, the assessment weight $W(\beta_i)$ of the 5 credibility categories of the cloud platform, and the random parameter Nonce of workload proof, as shown in Table 7.

4. Experimental Design and Analysis

4.1. Experimental Analysis of D-S Fusion Method in This Paper. Before the experimental analysis of the credibility assessment system, this paper first verifies the effectiveness of the proposed fusion method. Suppose that for a cloud platform, the assessments given by 3 different users is shown in Table 8.

As shown in Table 8, there is a big conflict between Assessment 2 and other assessments. In this case, the results obtained by traditional D-S fusion method and the results obtained by the improved D-S fusion method proposed in this paper are shown in Table 9.

It can be seen from the results in Table 9, when there are occasional conflicts or malicious assessments in the assessment process, the assessment results obtained by the traditional D-S fusion method will be greatly affected. However, the fusion results obtained by this paper method will not be greatly affected and can still reflect the views of most effective assessments. The above experiments show that the proposed fusion method is effective and feasible.

4.2. Experimental Analysis of the Credibility Assessment System. After verifying the effectiveness of the proposed fusion method, this paper will verify the effectiveness of the proposed assessment system.

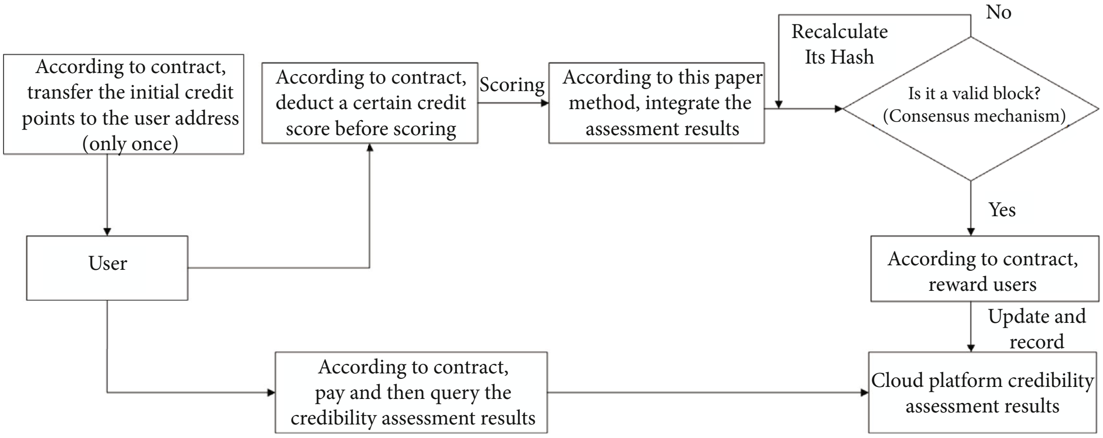

According to contract, transfer the initial credit points to the user address (only once)
According to contract, deduct a certain credit score before scoring
Scoring
According to this paper method, integrate the assessment results
Recalculate Its Hash
No
Is it a valid block? (Consensus mechanism)
Yes
User
According to contract, reward users
Update and record
According to contract, pay and then query the credibility assessment results
Cloud platform credibility assessment results

Figure 5: The process of cloud platform credibility assessment system proposed in this paper.

Table 7: Data uploaded to blockchain.

Data	Example
User address	0x6c19a33EF2cc41a1bedDC91133a8422e89f041B7
$m(\underline{A})$	0.001,0.001,0.162,0.389,0.389,0.027,0.027,0.0002,0.00006,0.00006
$W(\beta_i)$	0.180,0.223,0.107,0.242,0, 248
BlockNumber	101
The previous block address	0xa13782ab4bcb6e9670d315fb341ebbc95d45a2bdb0ea5034ef432b74f30b1b4f
The current block address	0x78dacc2af60900d2e4cae90b71e27446e6e883df36c53f21cbc9e071f7a586f4
Assessment date	20220407
Nonce	4

Table 8: The assessments of 3 different users.

	Assessment 1	Assessment 2	Assessment 3
$A_5 = \{8910\}$	0.500	0.000	0.600
$A_4 = \{678\}$	0.300	0.000	0.200
$A_3 = \{4567\}$	0.200	0.000	0.200
$A_2 = \{345\}$	0.000	0.400	0.000
$A_1 = \{123\}$	0.000	0.600	0.000

4.2.1. Experimental Design. The consensus mechanism of this experiment adopts the workload proof algorithm based on Ethash. The test framework is Remix provided by Ethereum, the experimental server is configured with CPU 5.0ghz and Ram 32g. After setting up the environment required for the experiment, the initial weight of the 5 credibility categories of the platform is set to 0.200, namely, $W(\beta_j) = 0.200$, $j = 1, 2, \cdots, 5$. Then, this experiment convened 10 experts to score a cloud platform and generated the initial block data according to the method proposed in this paper. Next, this paper has visited and consulted users who have used the platform and asked them to assess the platform in the form of questionnaire. Finally, this experiment substitutes the assessment data of all users into the system and obtains the data of more than 300 blocks.

Table 9: The assessments of 3 different users.

	Results obtained by traditional D-S fusion method	Results obtained by the improved D-S fusion method proposed in this paper
{10}	0.00000	0.12353
{9}	0.00000	0.12353
{8}	0.00000	0.23824
{7}	0.00000	0.08824
{6}	0.00000	0.08824
{5}	0.50000	0.13235
{4}	0.50000	0.13235
{3}	0.00000	0.03235
{2}	0.00000	0.02059
{1}	0.00000	0.02059

Through the system, the user can obtain the block information returned by the system after paying according to the contract. According to the address of the block, the user will be able to further query the weight $W(\beta_i)$ of the 5 credibility categories of the cloud platform and can also query the trust degree $m(\underline{A})$ of the cloud platform credibility level.

4.2.2. Experimental Result Analysis. Using the expert account to query, the following results can be obtained.

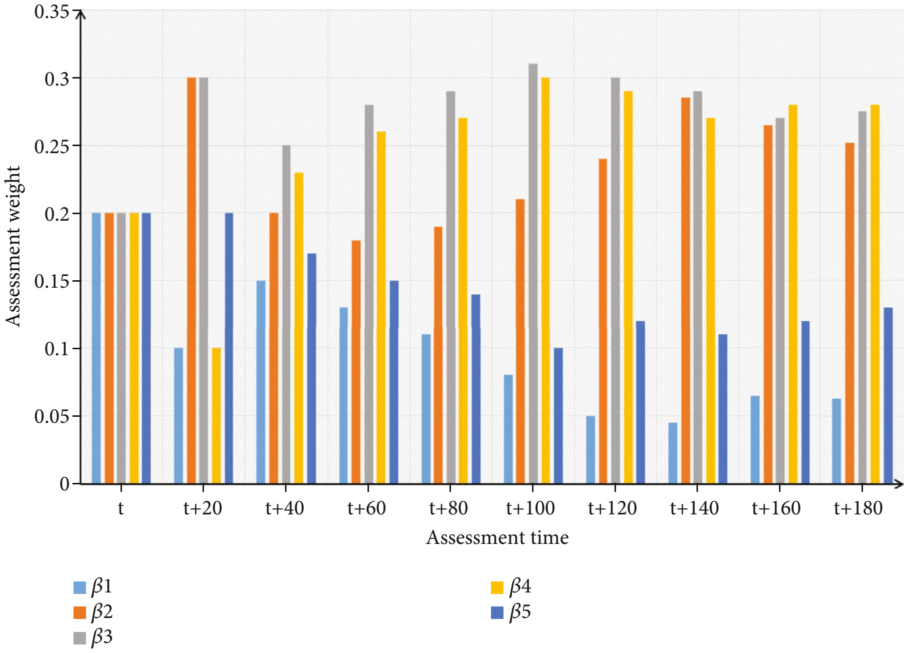


Figure 6: Changes in the assessment weight $W(\beta_i)$.

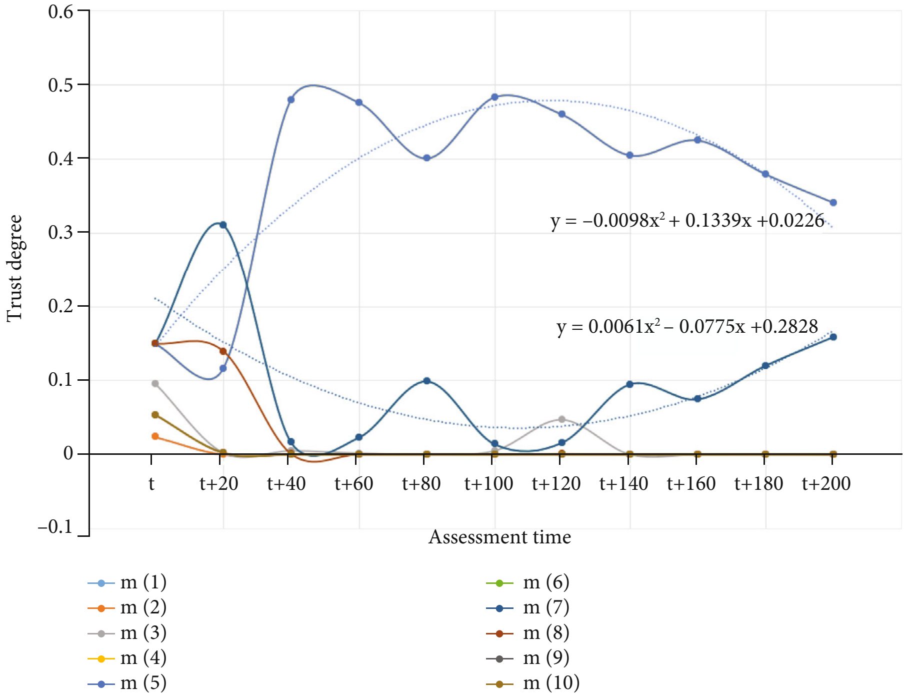


Figure 7: The change of $m(\underline{A})$.

(1) Changes in the Assessment Weight $W(\beta_i)$ of the 5 Credibility Categories. In Figure 6, *t* represents the generation time of the first block, that is, the time of the first assessment. As can be seen from Figure 6, in the initial stage, the assessment weight $W(\beta_i)$ of each credibility category changes greatly. However, with the increase of the number of user assessments, the assessment weight $W(\beta_i)$ of the 5 credibility categories will gradually stabilize. Finally, the assessment weight sorting result is $W(\beta_4) > W(\beta_3) > W(\beta_2) > W(\beta_5) > W(\beta_1)$. According to the entropy weight method, the sorting results show that β_4 has the greatest impact on the credibility of the whole platform, and users have the greatest difference in the assessment of "service provider management credibility β_4"; on the contrary, the value of $W(\beta_1)$ is the lowest, indicating that β_1 has the lowest impact on the credibility of the whole platform, and users have the

Table 10: Cost comparison and comprehensive comparison.

	Cost	Comprehensiveness
This paper method	The cost required includes the following: (1) Assessment weight $W(\beta_i)$of each credibility category given by users (2) Confidence interval $\text{lev}(\beta_i)$ of each credibility category given by users (3) D-S fusion method is required in the assessment process, and its average time complexity is $O(n^2)$ (4) In addition, this method also needs to build a blockchain system	The output assessment results include the following: (1) The change of $W(\beta_i)$ (2) Cloud platform credibility level and its trust degree $m(\underline{A})$ (3) The change of $m(\underline{A})$
Method based on AHP	The cost required includes the following: (1) Weight judgment matrix of credibility categories (2) Asymptotic normalization coefficient (ANC) is required in the assessment process, and its average time complexity is $O(n^2)$	The output assessment results include the following: (1) Assessment weight $W(\beta_i)$ of each credibility category calculated by ANC
Method based on entropy	The required input data include the following: (1) Risk frequency of each credible category (2) Risk loss severity of each credible category (3) Entropy weight method is required in the assessment process, and its average time complexity is $O(n^2)$	The output assessment results include the following: (1) Entropy weight of each credibility categories, namely, the assessment weight $W(\beta_i)$ (2) The uncertainty degree of cloud platform risk, namely, cloud platform credibility level
Method based on D-S theory	The required input data include the following: (1) The confidence interval and trust degree of each credibility category (2) D-S fusion method is required in the assessment process, and its average time complexity is $O(n^2)$	The output assessment results include the following: (1) Cloud platform credibility level and its trust degree $m(\underline{A})$
Method based on fuzzy theory	The required input data include the following: (1) The confidence interval and trust degree of each credibility category (2) Directly assess the cloud platform credibility level and its trust degree according to Fuzzy Sets. The average time complexity is $O(1)$	The output assessment results include the following: (1) Cloud platform credibility level and its trust degree $m(\underline{A})$

smallest difference in the assessment of "infrastructure credibility β_1".

(2) Changes in the Trust Degree $m(\underline{A})$ of the Cloud Platform Credibility Level. Through query, the change of $m(\underline{A})$ is shown in Figure 7.

Starting from the initial assessment records, the query is conducted every 20 blocks. A total of 11 assessment records are queried in this experiment.

In Figure 7, t represents the generation time of the first block, that is, the time of the first assessment. As can be seen from Figure 7, the platform credibility level is 6 and 7, followed by 4 and 5. However, from the change trend, the values of $m(\underline{6})$ and $m(\underline{7})$ show a downward trend, while the values of $m(\underline{4})$ and $m(\underline{5})$ show an upward trend, indicating that the credibility level of the platform shows a downward trend with the increase of the number of user assessments.

In addition, on the whole, the possibility of the platform credibility level belonging to other levels is low, indicating that the platform is relatively stable. The values of $m(\underline{10})$ and $m(\underline{1})$ are close to 0, indicating that the platform is neither a highly trusted platform nor a low trusted platform and is always in a generally trusted state.

4.3. Method Comparison

The above experiments show that the assessment method proposed in this paper is effective and feasible. Next, this paper compares the proposed method with other similar methods.

The methods proposed in this paper are mainly aimed at assessing the credibility of cloud platforms. In order to illustrate the advantages of the methods proposed in this paper, it is necessary to compare it with other similar assessment methods, such as method based on AHP, method based on entropy, method based on D-S theory, and method based on fuzzy theory.

Suppose a cloud platform contains n credibility categories, and the above methods are used to assess the platform. The comparative analysis of each method is shown in Tables 10 and 11.

Summarizing the above comparison, the results are shown in Table 12.

In Tables 10–12, cost represents the cost required for assessment when using this method; comprehensiveness

Table 11: Comparison of privacy security, data stability, and objectivity.

	Privacy security	Data stability	Objectivity
This paper method	Adopt blockchain technology for privacy protection	The assessment result cannot be tampered with	D-S fusion method can effectively solve the conflict information in the assessment process and ensure the objectivity of the assessment results.
Method based on AHP	No privacy protection	The assessment result is easy to be tampered with	In weight assessment, the method of pairwise comparison can effectively reduce the impact of human subjective factors.
Method based on entropy	No privacy protection	The assessment result is easy to be tampered with	Describing the credibility level by the risk uncertainty can effectively reduce the impact of human subjective factors on the assessment results.
Method based on D-S theory	No privacy protection	Because the assessment result is the fusion of different users' assessment results, the result is not easy to be tampered with	D-S fusion method can effectively solve the conflict information in the assessment process and ensure the objectivity of the assessment results.
Method based on fuzzy theory	No privacy protection	The assessment result is easy to be tampered with	The assessment results are obtained by human subjective assessment. There is no effective method to improve the objectivity of the assessment results in the assessment process. Compared with other methods, its objectivity is low.

Table 12: Comparison between this paper method and other methods.

	Cost	Comprehensiveness	Privacy security	Data stability	Objectivity
This paper method	High	High	High	High	Medium
Method based on AHP [5–11]	Medium	Low	Low	Low	Medium
Method based on entropy [12–16]	Medium	Medium	Low	Low	Medium
Method based on D-S theory [17–22]	Medium	Medium	Low	Medium	Medium
Method based on fuzzy theory [31–33]	Low	Medium	Low	Low	Low

means the comprehensiveness of the evaluation results. The more assessment results this method can provide to users, the more comprehensiveness this method; privacy security refers to the security degree of the method in user privacy protection; data stability indicates the stability of the evaluation results. The higher the stability, the less likely the assessment results will be modified by malicious users; objectivity means the objectivity of the assessment results. The higher the objectivity, the lower the impact of human subjective factors on the assessment results.

5. Conclusion

This paper integrates blockchain technology and D-S theory and carries out a series of research on the credibility assessment of cloud platforms. Firstly, based on D-S theory and fuzzy theory, this paper proposes an effective cloud platform credibility assessment method, which solves the conflict problem in the assessment process by integrating the user's assessments and reduces the difficulty of user scoring. On this basis, combined with blockchain technology, this paper regards the fused assessment results as effective blocks on the blockchain, proposes an effective block generation method, and designs the corresponding consensus mechanism, smart contract, and incentive mechanism. As mentioned above, combined with D-S theory and blockchain technology, this paper designs and proposes an effective cloud platform credibility assessment system. Through the encryption technology and traceability technology of blockchain, the system makes up for the defects of the assessment method based on D-S theory, effectively protects the privacy of users participating in the assessment process, ensures the assessment results cannot be tampered with, and improves the assessment enthusiasm of users. Finally, the experimental analysis results show that the assessment system proposed in this paper is effective and feasible.

However, as an assessment system, the assessment results that the system can provide to users are not comprehensive enough. In the follow-up research, we also need to sort out the specific impact indicators based on the cloud platform credibility categories divided in this paper and carry out the assessment combined with the specific credibility evidence, so as to improve the objectivity of the assessment results.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (no. 11861071), the Yunnan Fundamental Research Projects (nos. 202101AT070211, 202201AT070142, and 202201AT070166), and the Talent Introduction Projects of Yunnan University of Finance and Economics (no. 2021D16).

References

[1] Flexera, *Flexera 2020 State of the Cloud Report*, Flexera, America, 2020.

[2] S. Chang-xiang, "Scientific concept of network security and trusted computing 3.0," in *China Software Industry Annual Conference*, Beijing, 2018.

[3] Y. Xi, L. Ping, and G. Jabeen, "The concept model of software trustworthiness based on trust-theory of sociology," *Acta Electronica Sinica*, vol. 47, no. 11, pp. 2344–2353, 2019.

[4] T. Zhang, K. Zhao, M. Yang, T. Gao, and W. Xie, "Research on privacy security risk assessment method of mobile commerce based on information entropy and Markov," *Wireless Communications and Mobile Computing*, vol. 2020, Article ID 8888296, 11 pages, 2020.

[5] K. A. Alam, R. Ahmed, F. S. Butt, S. G. Kim, and K. M. Ko, "An uncertainty-aware integrated fuzzy AHP-WASPAS model to evaluate public cloud computing services," *Procedia Computer Science*, vol. 130, pp. 504–509, 2018.

[6] C. Li, S. Wang, L. Kang, L. Guo, and Y. Cao, "Trust evaluation model of cloud manufacturing service platform," *International Journal of Advanced Manufacturing Technology*, vol. 75, no. 1-4, pp. 489–501, 2014.

[7] P. Lou, L. Yuan, J. Hu, J. Yan, and J. Fu, "A comprehensive assessment approach to evaluate the trustworthiness of manufacturing services in cloud manufacturing environment," *IEEE Access*, vol. 6, pp. 30819–30828, 2018.

[8] R. Fattahi and M. Khalilzadeh, "Risk evaluation using a novel hybrid method based on FMEA, extended MULTIMOORA, and AHP methods under fuzzy environment," *Safety Science*, vol. 102, pp. 290–300, 2018.

[9] M. Fagundes, T. C. Keler, E. O. Teles, S. Melo, and F. Freires, "Multicriteria decision-making system for supplier selection considering risk: a computational fuzzy AHP-based approach," *IEEE Latin America Transactions*, vol. 19, no. 9, pp. 1564–1572, 2021.

[10] Z. Li and R. Jie, "Cloud service trust evaluation algorithm optimization based on multi-level structure model," *Journal of Nanjing University of Science And Technology*, vol. 44, no. 1, p. 6, 2020.

[11] C. Ze-Qian, S. Xiao-Tong, Z. Na-Jing, and Y. Shuo, "Construction and application of evaluation index of public cultural cloud service," *Library and Information Knowledge*, vol. 2020, no. 6, pp. 54–66, 2020.

[12] T. Tilei, L. Tong, Y. Ming, and J. Rong, "Research on a trustworthiness measurement method of cloud service construction processes based on information entropy," *Entropy*, vol. 21, no. 5, p. 462, 2019.

[13] T. Gao, T. Li, R. Jiang, M. Yang, and R. Zhu, "Research on cloud service security measurement based on information entropy," *International Journal of Network Security*, vol. 21, no. 6, pp. 1003–1013, 2019.

[14] H. Guesmi, A. Kalghoum, C. Ghazel, and L. A. Saidane, "FFED: a novel strategy based on fast entropy to detect attacks against trust computing in cloud," *Cluster Computing*, vol. 24, no. 3, pp. 1945–1954, 2021.

[15] A. Sharma, P. Munjal, and H. Banati, "Entropy-based classification of trust factors for cloud computing," *International Journal of Grid and Utility Computing*, vol. 11, no. 6, pp. 747–754, 2020.

[16] S. Nie, "A novel trust model of dynamic optimization based on entropy method in wireless sensor networks," *Cluster Computing*, vol. 22, no. S5, pp. 11153–11162, 2019.

[17] L. Wei, Z. Lu-Kun, B. A. Yuan-Jie, L. I. Guang-Li, and Z. Zhi-Gang, "A relevance aware cloud service trust model based on convex evidence theory," *Computer Engineering & Science*, vol. 41, no. 1, pp. 47–55, 2019.

[18] L. Zuan-shi and G. Xiu-li, "Trusted cloud service evaluation method research based on D-S theory," *Computer Engineering and Applications*, vol. 53, no. 17, pp. 70–76, 2017.

[19] D. X. Wang and Q. Wang, "Trustworthiness evidence supporting evaluation of software process trustworthiness," *Journal of Software*, vol. 29, no. 11, pp. 3412–3434, 2018.

[20] W. Xu, W. Yang, and Y. Yao, "Multi-dimensional trust evaluation method based on D-S evidence theory," *Computer and Digital Engineering*, vol. 47, no. 2, p. 7, 2019.

[21] M. Yang, T. Gao, R. Jiang, L. Jia, and D. Yang, "Comprehensive assessment of mobile service privacy security based on FAHP and D-S theory," *Wireless Communications and Mobile Computing*, vol. 2, 20 pages, 2022.

[22] M. Yang, T. Gao, W. Xie, L. Jia, and T. Zhang, "The assessment of cloud service trustworthiness state based on D-S theory and Markov chain," *IEEE Access*, vol. 10, pp. 68618–68632, 2022.

[23] W. Tiedan, Z. Yang, and P. Dinghong, "Research on cloud service safety evaluation based on improved IVHF-TODIM method," *Computer Engineering and Applications*, vol. 54, no. 4, pp. 84–89, 2018.

[24] W. Tie-dan, T. Miao, and P. Ding-hong, "Hesitant fuzzy Taguchi multi-attribute decision making method for cloud service quality evaluation," *Fuzzy Systems and Mathematics*, vol. 33, no. 3, p. 16, 2019.

[25] S. Huiyang, L. Peng, and W. He, "Threat intelligence evaluation based on blockchain and a neural network," *Journal of Tianjin University(Science and Technology)*, vol. 55, pp. 527–534, 2022.

[26] Y. Ming, H. Xuexian, Z. Qihui, W. Jianghong, and L. Wenfen, "Federated learning scheme for mobile network based on reputation evaluation mechanism and blockchain," *Chinese Journal of Network and Information Security*, vol. 7, no. 6, pp. 99–112, 2021.

[27] L. Haiou, H. Xutao, L. Kai, and G. Yue, "A literature review of blockchain traceability mechanism," *Journal of Intelligence*, vol. 2022, no. 4, pp. 1–7, 2022.

[28] T. Aiguo and H. Chunhua, “Application of fuzzy theory in software project risk assessment,” *Journal of Central South University (Science and Technology)*, vol. 48, no. 2, pp. 411–417, 2017.

[29] D. Li, J. Chen, and M. Qiu, “The evaluation and analysis of the entropy weight method and the fractional grey model study on the development level of modern agriculture in Huizhou,” *Mathematical Problems in Engineering*, vol. 2021, 8 pages, 2021.

[30] F. Voorbraak, “A computationally efficient approximation of Dempster-Shafer theory,” *International Journal of Man-Machine Studies*, vol. 30, no. 5, pp. 525–536, 1989.

[31] X. Hu, R. Jiang, M. Shi, and J. Shang, “A privacy protection model for health care big data based on trust evaluation access control in cloud service environment,” *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 3, pp. 3167–3178, 2020.

[32] A. Mohsenzadeh, H. Motameni, and J. E. Meng, “Retraction note to: a new trust evaluation algorithm between cloud entities based on fuzzy mathematics,” *International Journal of Fuzzy Systems*, vol. 21, no. 6, p. 1988, 2019.

[33] R. Pei-Zhi, L. Wei, B. Ran, and M. Ping, “A simulation credibility assessment method based on improved fuzzy comprehensive evaluation,” *Journal of System Simulation*, vol. 32, no. 12, pp. 185–190, 2020.