IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 24, 2020, accepted May 6, 2020, date of publication May 11, 2020, date of current version May 28, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2993921

Secure Digital Certificate-Based Data Access Control Scheme in Blockchain

BIN LIU[1], LIJUN XIAO[2], JING LONG[3], MINGDONG TANG[4], AND OSAMA HOSAM[5]

[1]Software Engineering Institute, Xiamen University of Technology, Xiamen 361024, China
[2]Big Data Development and Research Center, Guangzhou College of Technology and Business, Guangzhou 528138, China
[3]College of Information Science and Engineering, Hunan Normal University, Changsha 410081, China
[4]School of Information Science and Technology, Guangdong University of Foreign Studies, Guangzhou 510006, China
[5]College of Computer Science and Engineering at Yanbu, Taibah University, Yanbu 46421, Saudi Arabia

Corresponding author: Jing Long (jlong@hunnu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61872138, in part by the Scientific Research Fund of Hunan Provincial Education Department under Grant 19C1157, and in part by the Hunan Provincial Science and Technology Project Foundation under Grant 2018TP1018.

ABSTRACT Previous contract protocols in blockchains ensure their fairness and traceability by utilizing centralized credible nodes. If credible nodes are dishonest or conspire with the signatory, then other nodes are compromised. Meanwhile, the leakage of sensitive information of participant nodes poses a serious threat to the privacy security of data access in blockchains. To address this issue, this study proposes a secure control method of digital certificate-based data access in blockchains. The proposed method combines blockchain and digital certificate technologies and designs a secure authentication protocol for privacy data in blockchains without verifying the encrypted identity signature of the third-party participant. The high-efficiency network forwarding protocol proposed in this work can support the fair contract signing of multiple signers via blockchain. This protocol can protect the privacy of contracts and identities of participants. Experimental results show that the proposed scheme is superior in terms of communication overhead, storage overhead, and detection rate.

INDEX TERMS Blockchain, digital certificate, sensitive information, encrypted signature, authentication protocol.

I. INTRODUCTION

Blockchain technology originated from the peer-to-peer electronic payment system proposed by Nakamoto in 2008 [1]. Establishing a trustable relationship between two strange entities without a third-party center is difficult. The emergence of blockchains can address the trust issue among nodes in decentralized systems by using distributed node verification and consensus mechanism [2]. Private data access in blockchains is the transfer process of digital asset value that can realize a remarkable change in the current network architecture from ''information Internet'' to ''value Internet.'' Blockchains can realize a trustable transaction between two parties without the participation of any intermediary [3], [4]. This technology is a remarkable innovation in traditional trust transactions on the Internet or Internet-of-Things [5], [6].

The associate editor coordinating the review of this manuscript and approving it for publication was Hong-Ning Dai.

Blockchain technology generally utilizes a blockchain data structure to verify and store data. Consensus algorithm of distributed nodes is used to generate and update data. Encryption algorithm ensures the security of data transmission and access [7]–[9]. At present, the condition of all on-chain data transactions depends on the data authenticity provided by authority organizations. For example, authentication organizations determine the dependability of digital certifications in some core data or online banks specify the payment reliability in the capital of data delivery. Furthermore, the provision of security and privacy protection for data transaction only by the third-party reliable entities is not enough. Illegal intrusion, control, or compromise of reliable entities causes the leakage of individual private data and financial loss for companies or individuals. Consequently, the majority of methods of blockchain data authentication are unreliable. This condition requires a blockchain technology with high efficiency and security. The consensus and anonymity can ensure not only fast verification of online data but also protect data

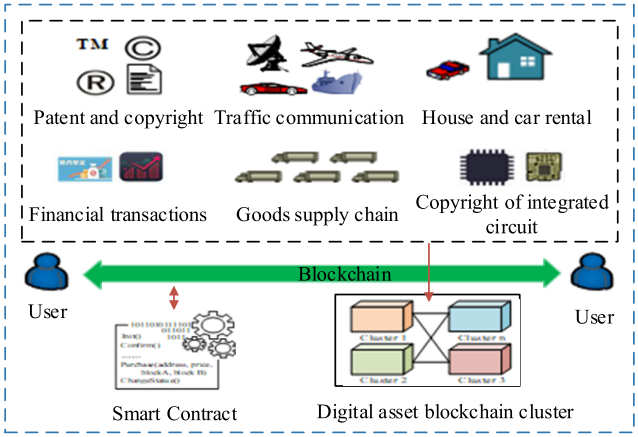


FIGURE 1. Architecture of secure digital access system in blockchain.

privacy [10], [11]. The secure access control of blockchain data (as in Fig.1) has the following features:

1) Low security cost. Traditional technologies of Internet data protection technologies realize secure transactions of digital capital by using the structure of network data security structure. This condition requires large network overheads for operation and maintenance. Data transactions of blockchain-based digital capital can ensure the secure authentications of nodes in the blockchain network by using consensus methods, including POW, POS, and PBFT; improve the protection speed of sensitive data transactions; and reduce the cost of digital capital transactions.

2) High supervision efficiency. Blockchains offer complete transparency in each block node data in the system. Ensuring the consistency of blockchain data via consensus data cooperation technology promotes convenience in distributed supervision. In this case, fraudulent behaviors are reduced and supervision efficiency of decentralization of blockchain is improved.

This work proposed a a secure control method of digital certificate-based data access in blockchain, which is organized as follows. Section II illustrates the related work. Section III and Section IV demonstrate the mathematical model and the proposed scheme. Section V analyzes the security and Section VI evaluates the performance. This work is summarized in Section VII.

II. RELATED WORK

Technology of blockchain-based distributed data transmission becomes an effective solution in addressing the security problem of digital capital transactions. The blockchain network has a tamper-proof environment. The transaction of any digital capital is verified by real participants or miners [12]. The blockchain system can use encryption methods to link the transaction data blocks for retrospective and immutable records. However, some problems exist. The consensus requires the miner to exchange a large reward by using computation ability. Consequently, greedy miners usually

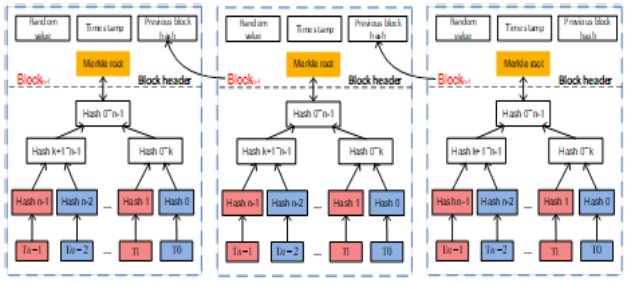


FIGURE 2. Distributed link structure of blockchain.

attempt to increase their mining ability through the system. Therefore, some security flaws target the computation ability in digital capital transactions of blockchains due to the existence of greedy rewards [13].

In recent years, the increase in user information leakage and other security events has raised suspicion on third parties in the security model of abundant private data collection and control. Joshi *et al.* [14] realized a secure blockchain-based protocol that transforms the blockchain into third-party automatic access controllers without trust. This protocol can be used to carry instructions and store, search, and share data. In this case, the dependable computation issue in blockchain networks is addressed. However, consistency analysis theory of algorithm performance and data support is lacking in this specific blockchain. Cachin [15] proposed a private blockchain evaluation scheme for the ethereum and hyperledger structures; this scheme generates a large amount of data-consistency evaluation results by using quantitative analysis of delay time and throughput in the blockchain network. Jeon et al [16] proposed a new IoT server platform by introducing a block chain and store sensor data in a block chain. Mobius selected IoT server platform, Mobius authenticates IoT devices conforming to oneM2M standard, receives real-time sensor data, stores information and data in Mysql server and manages it. Dorri *et al.* [17] proposed a communication approach for data control and audit; this strategy has low overhead, but its security and privacy should be further improved. Mylrea and Gourisetti [18] utilized blockchains and intelligent contracts to improve resilience in smart grids. Liang *et al.* [19] proposed a blockchain-based distributed solution to anchor the hashed data record collected by unmanned aerial vehicles in blockchain networks. Meanwhile, this approach generates specific digital receipts of blockchains in each data record stored in the cloud. Experiments showed that this technology is a reliable system of distributed digital recovery with low cost overhead and good extendability. Cai *et al.* [20] introduced a blockchain clearing system for digital capital transactions. This system disassembles and merges the composite transaction system in clearing procedures. The security of data risk decision and evaluation in blockchain is investigated. The distributed link structure of high-efficiency blockchain is shown in Fig. 2.

To address the issue on secure authentication of blockchain data, Karamitsos *et al.* [21] implemented a prototype of

digital capital transaction by utilizing ethereum blockchains and intelligent contracts. This prototype is audited, transparent, and distributed. However, the transaction security problem is not analyzed. Sharma *et al.* [22] discussed the security evaluation of blockchain data, pointed out the security flaw in blockchain clouds, and proposed corresponding solutions. Liang *et al.* [23] utilized different reward mechanisms in data transactions and simulated block withholding attacks in the blockchain cloud. The experiments showed that if the block withholding attack provides sufficient resource for malicious miners in the blockchain cloud and damages the mining of honest miners, then the attack is successful. The blockchain-based digital capital storage and transaction technology remarkably improves the running efficiency of the entire digital capital blockchain. Meanwhile, the separate changes in account and transaction information allow increased flexibility in digital management. Bahga and Madisetti [24] proposed a distributed computing platform with privacy and extendability features. The secure calculation of multiple parties realized data queries through distributed calculation. Any node can only partially access data. Off-chain storage technologies [25] relate blockchains to distributed hash tables. Blockchains only store the address of stored data. Double chains were utilized in the financial field in [26]. However, transactions are still point-to-point; this transaction type uses the original asymmetrical encryption technology. The use of blockchain in data sharing is practical. Es-Samali *et al.* [27] implemented a framework for automatic authority management by combining intelligent contract and access control. This framework is suitable for distributed digital copyright integration and authority management of different organizations.

In summary, the fairness and traceability of existing contract protocols in blockchains are realized by centralized credible nodes. If credible nodes are dishonest or conspire with the signatory, then other nodes are compromised. Meanwhile, the leakage of sensitive information of participant nodes poses a serious threat to the privacy security of data access in blockchains. In this work, each user in the blockchain should capture the digital certificate before obtaining blockchain data. However, the certificate is applied from the blockchain network provider. The node will offer the corresponding data to the blockchain user after verifying the validity of the certificate. The digital certificate can ensure user privacy, but the anonymity of digital certificate may be utilized by malicious users for unlimited access. Thus, the proposed authentication protocol plays a significant role in secure data detection and authentication.

III. SECURE PROTECTION MODEL OF DIGITAL CERTIFICATE-BASED BLOCKCHAIN

We assume that a blockchain is composed of N nodes to monitor the target environment continuously and provide interesting data to users. Base stations in the chain are unreliable in connecting the inner network to the external network. The collected data are stored at local or other nodes. Users obtain data directly from the nodes.

TABLE 1. Symbol illustration.

Symbols	Illustration
p, q	Two big prime numbers satisfies $q \mid p$-1, namely q and p-1are co-prime.
g	An element with order q *in* Z_p^*, called generation element.
m	Information to be signed.
m_w	Authorized identity, type of signed information, range, valid date, etc.
$x_P, x_A \in Z_p^*$	Private key of blockchain user
$y_P = g^{x_P} (\mathrm{mod}\, p)$, $y_A = g^{x_A} (\mathrm{mod}\, p)$	Public key of network provide and immediate agent institute
$H(\cdot), h(\cdot)$	Hash function with high security
$\|\|$	Connection string

We assume that the nodes in the blockchain can conspire with one other, forge certificates, and even capture some nodes to obtain interesting information. The blockchain user wants to obtain the information of others as much as possible but is not interested in leaking the identity and data access way. Moreover, this illegal behavior is performed when the circumstances are profitable. However, DoS attacks are not performed by the user because it does not benefit the collection of data. Users do not capture many nodes for data by escaping access control due to the large cost and effort involved. Only a few nodes are captured in reusing the certificates.

A. AUTHENTICATION AND PRIVACY PROTECTION OF BLOCKCHAIN USER

B. SCHEME DESCRIPTION

The digital certificate-based access control of nodes in blockchain networks includes three stages, namely, agency, certificate generation, and node authentication. The symbol notation is listed in Table 1.

The following steps are included at this stage:

1) U sends the registration information to P without user's identity.

2) P randomly selects the number and calculates K and s as follows:

$$K = g^k (\mathrm{mod} p), \tag{1}$$

$$s = x_P + k \cdot H(m_w || K) (\mathrm{mod} q). \tag{2}$$

3) P sends (K, s) and mw to the agent via a secure channel.

4) A receives (K, s) and verifies whether it is satisfied. If so, then A accepts the agency task and calculates as the agency key.

C. DIGITAL CERTIFICATE GENERATION

In the process of digital certificate generation, the on-chain anonymous verification institution A implements certificate generation through the following steps:

TABLE 2. Symbol description.

Notation	Description
Rot (x, *ind*)	Left rotate of X by *ind*
RRot(X,*ind*)	Right rotate of X by *ind*
ind	Random position of string
g	Element with order q in Z_p^*
ID_i	ID of user node
x_i,z_i	Secret key of user node
x_j,s_i,g^{x_j}	Secret key of supply chain node
b	Random number of user node
a,c	Random number of supply chain node

1) A randomly selects a number $\lambda \in_R Z_p^*$ and calculates t.

$$t = g^{\lambda + x_A} (\mathrm{mod} p) \tag{3}$$

2) A sends (K, t) to user U who can access the node data with the certificate.

3) U receives (K, t) and randomly selects $a, b \in Z_p^*$ and calculates

$$\begin{aligned} \mu &= t^a (y_P y_A K^{H(m_W||K)})^{ab} (\mathrm{mod} p), \\ e &= h(m||\mu) (\mathrm{mod} q), \\ e' &= a^{-1} e + b (\mathrm{mod} q). \end{aligned} \tag{4}$$

4) A receives e' and calculates $s'' = e's' + \lambda + x_A$ as the message signature. s'' is sent to U.

5) U receives s'' and calculates $\varphi = g^{s'' a} (\mathrm{mod} p)$. The blind agency signature of message m (m, m_w, φ, e, K) is called the digital certificate.

D. NODE VERIFICATION

Each blockchain node has y_P and y_A before deployment. The on-chain information provider can dynamically update y_P and y_A. U can enter the sensor network and access node data once the certificate is obtained. For any node N_i, $e = h(m||\varphi(y_P y_A K^{H(m_W||K)})^{-e} (\mathrm{mod} q)$ will be verified once the certificate (m, m_w, φ, e, K) is received. With $e = h(m||\mu) (\mathrm{mod} q)$, only $\mu = \varphi(y_P y_A K^{H(m_W||K)})^{-e} (\mathrm{mod}\ q)$ should be proven. N_i can detect the certificate in real time. Upon passing the two steps, node N_i provides the data required in the certificate to user U.

E. SECURE AUTHENTICATION PROTOCOL

In this section, a digital authentication protocol of certificate-based secure blockchain identity is proposed. The supply chain nodes in the blockchain authenticate the identity of user nodes in the blockchain. The permitted blockchain blocks store information of legal user node, including each authentication record, time stamp, position, and product. The supply chain node can access the blockchain and search for related information. The user node is verified through the comparison of historical records. The flow of secure authentication protocol is described in Fig.3.

1) PROTOCOL PREPARATION

We assume that G is a cyclic multiplicative group with a large prime and generator g. P is a prime that satisfies $p = 2q + 1$. x_i and x_j are keys in Z_p^*. All the elements in G are considered in Z_p^*. The element style in this section is g^{xi} mod p. ''mod p'' is omitted for easy reading. The symbols in this protocol are listed in Table 2.

2) PROTOCOL EXECUTION

The protocol in this work includes registration and authentication.

1) Registration

The registration center allocates the identity and secret key $(ID_i, x_i, z_i, g^{x_i})$ in each user node. The secret key of each supply chain node is $(x_j, s_i = \frac{1}{z_i}, g^{x_j})$.

2) Authentication

Step 1: The supply chain node generates random a, and indj sends a handshake message $Hello$, a, ind_j, g^{x_j} to the user.

Step 2: After receiving the message, the user node generates a random position message ind_i and a random number b to calculate the request authentication message.

$$\begin{aligned} M_1 &= g^{x_i x_j}, M_2 = g^{x_i x_j + ind_i} \oplus ind_i \\ M_3 &= g^{ind_i \cdot z_i}, M_4 = Rot(b \oplus a \oplus g^{x_i x_j}, ind_i) \\ M_5 &= Rot(z_i \oplus b, ind_i) \oplus b, M_6 = Rot(x_i \oplus z_i, ind_i) \oplus z_i \\ M_7 &= Rot(ID_i \oplus x_i \oplus b \oplus z_i, ind_i) \oplus x_i \end{aligned} \tag{5}$$

The user node sends the request authentication message $g^{x_i}, M_2, M_3, M_4, M_5, M_6, M_7$ to the supply chain node for authentication.

Step 3: The supply chain node receives the message and calculates

$$\begin{aligned} M_8 &= g^{x_i x_j}, \quad M_9 = M_3^{s_i} = g^{ind_i \cdot z_i \cdot \frac{1}{z_i}} = g^{ind_i} \\ M_{10} &= M_8 \cdot M_9 = g^{x_i x_j + ind_i}, \quad ind_i' = M_2 \oplus M_{10} = g^{x_i x_j + ind_i} \\ &\quad \oplus ind_i \oplus g^{x_i x_j + ind_i} \\ b' &= RRot(M_4, ind_i') \oplus a \oplus g^{x_i x_j}, \\ z_i' &= RRot(M_5 \oplus b', ind_i') \oplus b' \\ x_i' &= RRot(M_6 \oplus z_i', ind_i') \oplus z_i', \\ ID_i' &= RRot(M_7 \oplus x_i', ind_i') \oplus x_i' \oplus b' \oplus z_i' \end{aligned} \tag{6}$$

The supply chain node extracts the random number, secret key, and random position information of the user node. Then, $h(x_i' \| x_j' \| z_i' \| ID_i')$ is calculated. The supply chain node searches the block, which stores the value of $h(x_i' \| x_j' \| z_i' \| ID_i')$ in the permitted blockchain. If the block exists, then the supply chain node can trace the historical record of $h(x_i' \| x_j' \| z_i' \| ID_i')$ in each block of the blockchain. The supply chain node successfully authenticates the user node if the historical record of $h(x_i' \| x_j' \| z_i' \| ID_i')$ exists. Then,

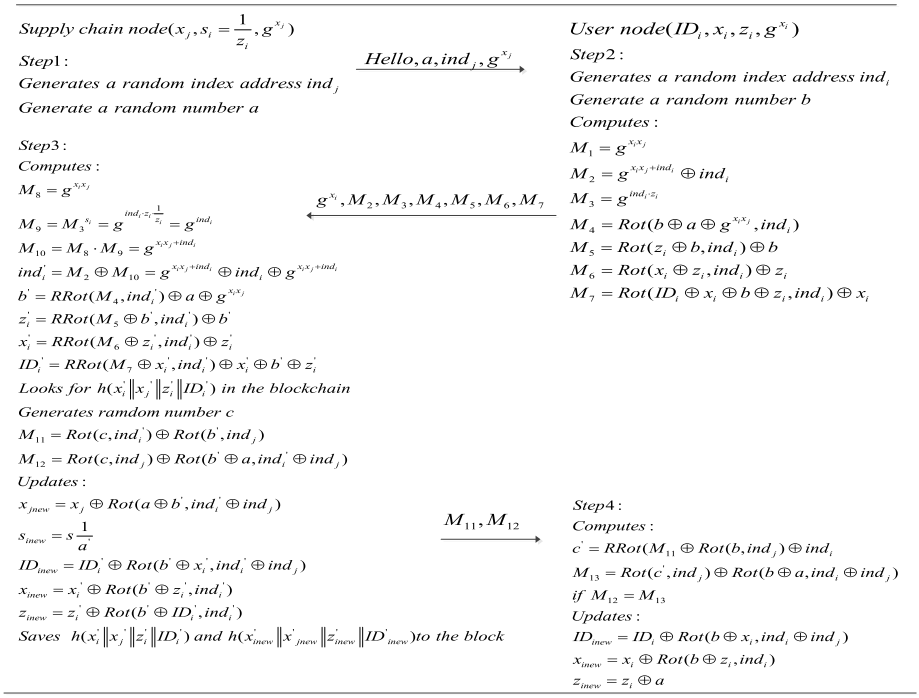


FIGURE 3. Flow of the proposed protocol.

the supply chain calculates (7) and updates data in (9).

$$\begin{aligned} &Generates\ ramdom\ number\ c \\ M_{11} &= Rot(c, ind_i') \oplus Rot(b', ind_j), \\ M_{12} &= Rot(c, ind_j) \oplus Rot(b' \oplus a, ind_i' \oplus ind_j) \end{aligned} \tag{7}$$

$$\begin{aligned} x_{jnew} &= x_j \oplus Rot(a \oplus b', ind_i' \oplus ind_j), \quad s_{inew} = s\frac{1}{a'} \\ ID_{inew} &= ID_i' \oplus Rot(b' \oplus x_i', ind_i' \oplus ind_j), \\ x_{inew} &= x_i' \oplus Rot(b' \oplus z_i', ind_i') \\ z_{inew} &= z_i' \oplus Rot(b' \oplus ID_i', ind_i'). \end{aligned} \tag{8}$$

The supply chain node stores $h(x_i' \| x_j' \| z_i' \| ID_i')$ and $h(x_{inew}' \| x_{jnew}' \| z_{inew}' \| ID_{inew}')$ in random blocks of the blockchain. Both new and old hash values are stored in this block to allow the supply chain to search historical authentication records in the subsequent authentication. Meanwhile, M_{11}, M_{12} are sent to the user node.

Step 4: User node calculates (9) after receiving M_{11}, M_{12}.

$$\begin{aligned} c' &= RRot(M_{11} \oplus Rot(b, ind_j) \oplus ind_i \\ M_{13} &= Rot(c', ind_j) \oplus Rot(b \oplus a, ind_i \oplus ind_j) \end{aligned} \tag{9}$$

If M12 = M13, then the user node successfully authenticates the supply chain node and updates data in (10).

$$\begin{aligned} ID_{inew} &= ID_i \oplus Rot(b \oplus x_i, ind_i \oplus ind_j) \\ x_{inew} &= x_i \oplus Rot(b \oplus z_i, ind_i) \\ z_{inew} &= z_i \oplus a \end{aligned} \tag{10}$$

The user node stores the updated data in itself. In this case, the user and supply chain nodes complete the bidirectional authentication. After each round of authentication, the user and supply chain nodes update the secret key by using a random number in time to ensure that it can resist replay attacks. In the protocol of this work, three random numbers {a,b,c} ensure the security of the secret key. Each message is updated in time. Attackers cannot conduct tracing attacks in the proposed protocol.

F. CERTIFICATE DETECTION

In the procedure of sending the certificate detection request, the nodes within the communication radius of the transmission path can receive the request message. If a node stores m, then a passive message is returned to the original node. With the same signature b, the proposed scheme can remarkably improve the detection range of the digital certificate. For example, node w_i is assumed to be one of the b witness nodes in a-*th* selection of node N_i, which is not selected in a−1 selections. Thus, node w_i will record m. V is a node in the path that acts as the certificate witness and receives the detection request message transmitted from node N_i to w_i.

The jump number between the two nodes is h. For simplification, we assume that each detection chain beginning at node Ni includes h + 1 nodes. The area size in the path of h hops is S_h. b witness block nodes generate the b request messages. Hence, the total area of b paths is expressed as follows:

$$S_b = bS_h - (b-1)S_r. \tag{11}$$

Similarly, we assume that the certificates of c (c < N) block nodes compromise the probability of (a − 1)bc/N. The remaining b (a − 1) (1 − c/N) witness block nodes are safe.

If none of them receive the detection request information, then the a-th certificate detection fails with the probability of (1 − Sb/S)b(a − 1)(1 − c/N). Thus, we have

$$p = 1 - (1 - S_b/S)^{(a-1)b(1-c/N)}. \tag{12}$$

The communication cost of the proposed scheme is $C = (b + \omega)h$, a ≥ 1. If (a − 1) b witness nodes are safe, then the probability of responding to a passive message is Sb/S. Therefore, it has high detection efficiency and low communication and storage cost.

G. SECURE STORAGE

Digital certificates are unique information in blockchains that will be copied and inquired. If the digital certificate of a block node is received, then each node in the blockchain will broadcast a detection request message. If the detection request message is also a witness node that is safe and has the digital certificate record, then the witness node returns a passive message to the original node. Otherwise, the certificate message is regarded as new. The original node selects an arbitrary path. The certificate is copied and stored in all the nodes in the path.

When the block nodes of other chains receive the digital certificate, block node N_i randomly generates a position $H(m, x_1)$. In this study, x_1 is an arbitrary random number. The block node sends an agent query message, including m to the node at the position $H(m, x_1)$. The one that receives the agent query message and is closest to $H(m, x_1)$ is called the agent query node of N_i (node U_1). If node U_1 stores m, then a passive message is returned to node N_i. Otherwise, U_1 sends the query request message containing m in the upward and downward directions. Node N_i refuses the use of a certificate if an abused passive message is received before the timeout of the timer. Otherwise, the certificate cannot be used normally. Block node N_i generates a random number x_2 that is different from x_1 and sends a copied agent request message with m to the nodes around $H(m, x_2)$. The node U_2 closest to $H(m, x_2)$ acts as the copied agent node of N_i after receiving the request message. U_2 subsequently stores m and sends the copied request message along two paths at the horizonal direction. All the nodes at the copied paths store m.

In the detection rate of abnormal nodes in the blockchain, a − 1 times of successful use of digital certificate generates a − 1 copied paths in network. In this study, at least one node exists to receive the query detection request message. We assume that a node receives an anonymous request message in the blockchain network, a − 1 crossing nodes are stored in the blockchain network. Given that the query path of block node N_i is predictable, user U can only randomly capture some nodes for a-th use of certificate $C = C_1 + C_2 = (h + W) + (h + L) = 2h + W + L$. We assume that U captures c nodes. The probability that each crossing node will be compromised is c/N. If a − 1 crossing nodes are compromised, then a-th certificate detection fails with the probability of $(\frac{c}{N})^{a-1}$.

$$p = 1 - (\frac{c}{N})^{a-1} \tag{13}$$

The communication cost of blockchain network exists in two cases, namely, a-th successful detection and unsuccessful detection. The communication cost of the first case includes the following parts: query cost C_1 and copy cost C_2. C_1 is the total cost of transmitting the query agent request and sending two query detection requests. C_2 includes the costs that transmits the copied agent request and sends the two copied requests. We assume the existence of L and W node hops in the horizontal and vertical directions. The average hop number between the two nodes is h. Therefore, we have

$$C = C_1 + C_2 = (h + W) + (h + L) = 2h + W + L. \tag{14}$$

If a-th certificate detection is unsuccessful, then C is the sum of C1 and the cost of sending the passive message. If − 1 crossing nodes are safe, then a − 1 passive messages are returned to node Ni. In this case,. Consequently, we have

$$C = (1 - p)(2h + L) + W + p(ah + W). \tag{15}$$

In addition, the storage cost of the proposed scheme is calculated as follows:

$$M = (a - 1)L + (1 - p)L = (a - p)L. \tag{16}$$

IV. SECURITY EVALUATION

The security performance evaluation of the proposed scheme includes the following aspects:

- Access control. The primary key at each stage is encrypted by a group of features. Attackers will not obtain the encryption key without the primary key due to the one-way feature of the key chain. The encryption with primary key is secure with an assumption. It demonstrates that attackers cannot decrypt the primary key and expect to have the access structure. In this case, the proposed scheme can control the data being accessed by authorized users.
- Limit collusion attack. Collusion users capture the primary key to decrypt the data.
- Limit the effect of node capture attacks. Each sensor node only stores the current key of data encryption. The previously used key is eliminated. Attackers cannot deduce the historical keys with the current key due to the one-way feature of the key. Each node independently stores the encrypted data. Hence, the capture of other nodes by one compromised node is useless.
- Overhead and functionality. Each sensor node is responsible for the following operations at various stages: generating the primary key and encrypting with the proposed scheme, generating the data encryption key based on the primary key, and encrypting the sensor data. These operations are further allocated to various stages. Each node concretely executes scalar multiplication at the elliptic curve, one-way hash, and symmetry data encryption at each stage at most.

Table 3 compares the proposed scheme to other schemes in [27] and [28]. In this table, the scheme in [27] designed

TABLE 3. Performance comparison of various schemes.

Scheme	Extendibility	Descriptiveness	Resistance against 51% attack	User Undo
Scheme [27]	No	Medium	No	No
Scheme [28]	No	Low	No	No
Our scheme	Yes	High	Yes	No

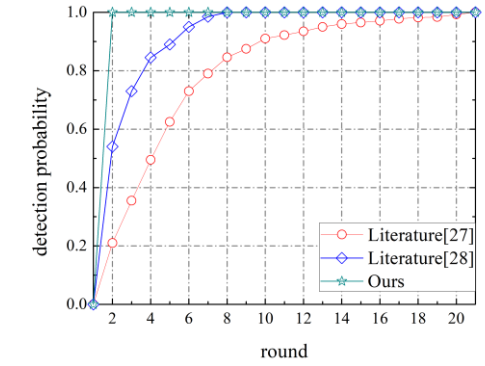
a simple threshold of the encryption key without extendability, resistance against 51% of attacks, and the user's Undo function. The scheme in [28] can resist against collusion attacks. The proposed scheme is extendable and with the Undo function and can resist collusion attacks. By contrast, the functionality of the proposed scheme is more comprehensive.

V. EXPERIMENTS AND ANALYSIS

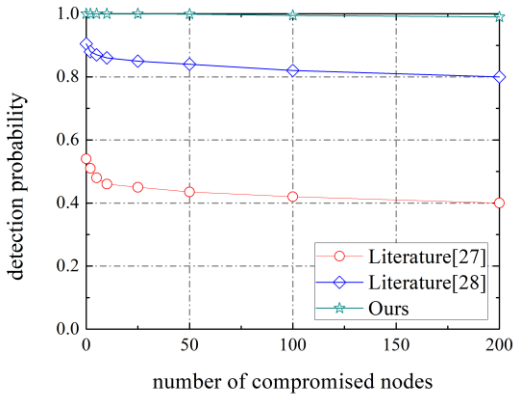
In this work, the hyperledger fabric platform is utilized to implement a universal framework of blockchain experiment. Docker container uses Ubuntu OS. The official real-service data set is adopted for test and verification. The open-source Hyperledger Fabric 1.0.5 can be download from Github at the website https://github.com/hyperledger/fabric. The following experiment steps are presented: (1) Install chain code in a running fabric network completes the end-to-end testing and ensures that all functional components in the blockchain network are normal. This business-related procedure related creates the channel, installs the chain code, instantiates the chain code, simulates the calling chain code, executes the transaction, and inquires the blockchain information. (2) Create channel: the request of installing the chain code is created via Install Proposal Request. The ID and loading position of the chain code is set up. Then, all the peers in the organization are obtained. The installation request is sent to all the peers. The complete data set is used in the experiment. The experiments are conducted in the platform with two servers and more than 20 blockchain terminals, including the evaluation of detection probability and communication cost.

A. DETECTION PROBABILITY

Fig. 4(a) shows the evaluation of the relationship between the detection probability and certificate utilization round and the comparison of the three schemes. In the proposed scheme, the detection probability increases with a. When a reaches 2, it can be detected completely. Given that the proposed scheme selects two cross curves to store and verify the certificate, the detection effect can be ensured if the cross node is safe. The scheme in [27] has the lowest detection probability because it utilizes a completely random method in selecting the target nodes. However, the selected nodes are not optimal and some may be compromised. The scheme in [28] can detect the certificate after eight times and utilizes the nodes in the path for feedback. b paths exist with b nodes. Some of the



(a) detection probability with increase of round



(b) detection probability with increase of compromised nodes

FIGURE 4. Evaluation and comparison of detection probability and the communication cost.

nodes in the path are compromised. The detection accuracy of the scheme in [28] is higher than that in [28] with less required witness nodes.

Fig. 4(b) shows the relationship between the detection probability and the number of compromised nodes c. In the scheme in [27], we set $a = 2$ and $b = 50$. In the scheme in [28], we set $b = 10$. The comparative schemes are affected by the value of c due to the random witness node. For example, when 10% of the nodes are compromised, the scheme in [27] can detect the initially abused certificate with a probability of 98% because the copied path has many nodes that receive the detection request message. This finding demonstrates that the proposed scheme has a good detection effect after suffering a node capture attack.

B. ABILITY AGAINST COLLUSION ATTACKS

Fig. 5 shows the comparison of ability against collusion attacks. For simplification, a node is assumed to generate a value at each period of each stage. The total number of users in the current network is 100. The number of users that simulate collusion attacks changes from 10 to 50. Under this condition, the percentage of data disclosed is compared in the different schemes. A user conducts an attack to capture additional data with the holding key. Compared with the direct

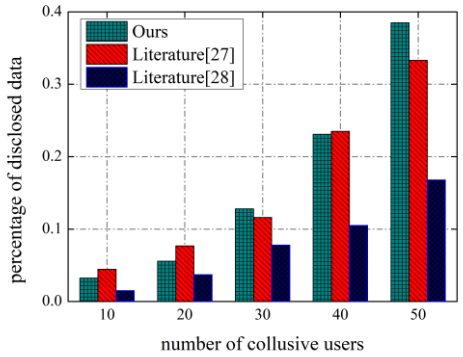


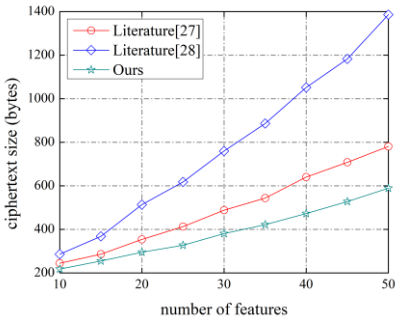
FIGURE 5. Comparison of ability against collusion attacks.

capturing node and eavesdropping attack, the collusion attack saves more cost and is more difficult to detect. In Fig. 5, when the number of collusive users is small, three schemes achieve similar abilities against attacks. However, the increase in collusive users rapidly reduces the data security in the schemes in [27] and [28]. The scheme in [28] decreases slightly. When the number of collusive users exceeds 40, the scheme in [28] shows a sharper decline than that in [27]. Given that the scheme in [27] generates the secret key with a random number, which is actually a pseudorandom number, it can be listed through exhaustion, which becomes easy when the number of collusive users increases. The scheme in [28] eliminates the aforementioned negative factors and updates the primary key in time. If a malicious user is found, then the Undo operation is executed. In this case, the collusive user can only obtain his/her own data rather than the data of other nodes. Therefore, the attack effect is limited within the minimum range.

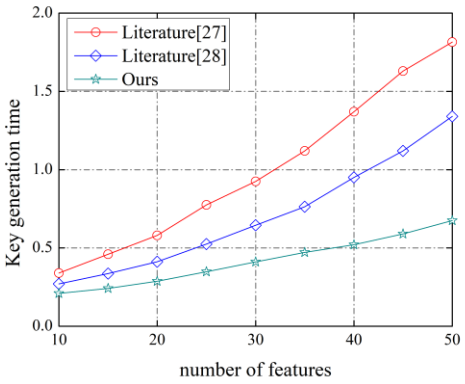
C. OTHER PERFORMANCE METRICS

To verify the validity of compromised nodes, we compare the performance of various metrics to that of [27] and [28], including length of cipher text, key generation time, time cost of encryption, and decryption. Fig. 6 shows the relationship between the performance and feature number. The length of the cipher text includes ID, head, and data block.

As shown in Fig. 6, the performance of the three schemes increases with the growth of features. The result clearly shows that the decryption time is nonlinear because it relates with the number of features and the specific access tree. Different access trees have their own access structures. Overall, the proposed scheme achieves better performance than the comparative schemes because the proposed scheme executes a strict access control strategy with a secure protocol at the encryption stage. In this case, the secret key is reconfigured via polynomial interpolation. Decryption requires many complex matching and exponentiation operations. Although the scheme in [27] used random element instead of secret sharing for strict control at the encryption stage, the size of the cipher text and secret key increase linearly with the growth of feature number. In this case, the efficiency of the scheme in [27]



(a) Ciphertext size



(b) Key generation time

FIGURE 6. Comparison of the performance and feature number.

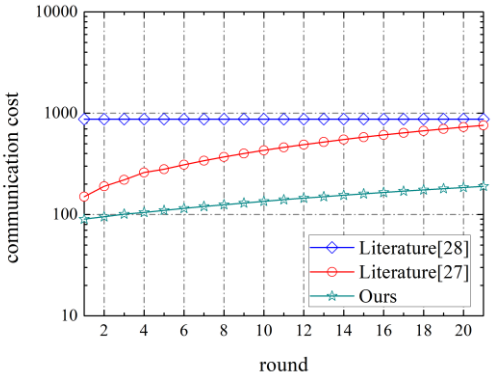


FIGURE 7. Evaluation of communication cost.

is low. The scheme in [28] utilized periodical encryption, and each node is encrypted with a symmetric encryption algorithm. The secret keys of each period form a one-way key chain. One key is used in each period.

Fig. 7 shows the comparison of communication cost among the three schemes. The result is base-10 logarithm. In Fig. 7, the proposed scheme has lower communication cost than those in [27] and [28].

VI. CONCLUSIONS

To address the security issue of data access in the blockchain, this study designs a digital control scheme of

certificate-based data access in blockchains. The main contribution of this work is listed as follows: 1) The proposed scheme divides the blockchain data based on their features and creates the relationship with the secret key. When a user performs a query, the access control strategy related to the secret key is used to assess whether the query is legal and thus realizes data access control. 2) A blockchain-based communication protocol, including a universal slotted protocol, overhead balance strategy, and fault-tolerance strategy, is designed. 3) To protect the data visitor, a public anonymous authentication is realized. Moreover, three methods of distributed certificate detection are designed to avoid the abuse of certificates by malicious anonymous users. The experimental results show that the proposed scheme has good protection ability against collusion and node capture attacks. Furthermore, the user's Undo ability, communication cost, storage cost, and detection efficiency are encouraging. Future studies should improve the existing scheme and apply them in real blockchain nodes. In large-scale transactions, we will further investigate theories, such as homomorphic attribute encryption and storage space balance, consider the requirement of low-delay storage module, and explore the model of hierarchical pluggable storage. Moreover, the parallel extendable distributed storage scheme in the scenario of mass data is investigated to improve the efficiency and extendability of storage modules.

REFERENCES

[1] S. Nakamoto. (2009). *Bitcoin: A Peer-to-Peer Electronic Cash System.* [Online]. Available: https://bitcoin.org/en/bitcoin-paper

[2] W. Liang, M. Tang, J. Long, X. Peng, J. Xu, and K.-C. Li, "A secure FaBric blockchain-based data transmission technique for industrial Internet-of-Things," *IEEE Trans. Ind. Informat.*, vol. 15, no. 6, pp. 3582–3592, Jun. 2019.

[3] H.-N. Dai, Z. Zheng, and Y. Zhang, "Blockchain for Internet of Things: A survey," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8076–8094, Oct. 2019.

[4] Z. Zheng, S. Xie, H.-N. Dai, W. Chen, X. Chen, J. Weng, and M. Imran, "An overview on smart contracts: Challenges, advances and platforms," *Future Gener. Comput. Syst.*, vol. 105, pp. 475–491, Apr. 2020.

[5] W. Liang, W. Huang, J. Long, K. Zhang, K.-C. Li, and D. Zhang, "Deep reinforcement learning for resource protection and real-time detection in IoT environment," *IEEE Internet Things J.*, early access, Feb. 17, 2020, doi: 10.1109/JIOT.2020.2974281.

[6] W. Liang, K.-C. Li, J. Long, X. Kui, and A. Y. Zomaya, "An industrial network intrusion detection algorithm based on multifeature data clustering optimization model," *IEEE Trans. Ind. Informat.*, vol. 16, no. 3, pp. 2063–2071, Mar. 2020, doi: 10.1109/TII.2019.2946791.

[7] X. Li, J. Niu, M. Karuppiah, S. Kumari, and F. Wu, "Secure and efficient two-factor user authentication scheme with user anonymity for network based E-health care applications," *J. Med. Syst.*, vol. 40, no. 12, Dec. 2016.

[8] N. Z. Aitzhan and D. Svetinovic, "Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams," *IEEE Trans. Dependable Secure Comput.*, vol. 15, no. 5, pp. 840–852, Sep. 2018.

[9] X. Li, J. Niu, S. Kumari, S. H. Islam, F. Wu, M. K. Khan, and A. K. Das, "A novel chaotic maps-based user authentication and key agreement protocol for multi-server environments with provable security," *Wireless Pers. Commun.*, vol. 89, no. 2, pp. 569–597, Jul. 2016.

[10] X. Li, A. K. Sangaiah, S. Kumari, F. Wu, J. Shen, and M. K. Khan, "An efficient authentication and key agreement scheme with user anonymity for roaming service in smart city," *Pers. Ubiquitous Comput.*, vol. 21, no. 5, pp. 791–805, Oct. 2017.

[11] T. T. A. Dinh, R. Liu, M. Zhang, G. Chen, B. C. Ooi, and J. Wang, "Untangling blockchain: A data processing view of blockchain systems," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 7, pp. 1366–1385, Jul. 2018.

[12] W. Liang, Y. Fan, K.-C. Li, D. Zhang, and J.-L. Gaudiot, "Secure data storage and recovery in industrial blockchain network environments," *IEEE Trans. Ind. Informat.*, early access, Jan. 13, 2020, doi: 10.1109/TII.2020.2966069.

[13] Y.-L. Gao, X.-B. Chen, Y.-L. Chen, Y. Sun, X.-X. Niu, and Y.-X. Yang, "A secure cryptocurrency scheme based on post-quantum blockchain," *IEEE Access*, vol. 6, pp. 27205–27213, 2018.

[14] A. P. Joshi, M. Han, and Y. Wang, "A survey on security and privacy issues of blockchain technology," *Math. Found. Comput.*, vol. 1, no. 2, pp. 121–147, 2018.

[15] C. Cachin, "Architecture of the hyperledger blockchain fabric," in *Proc. Workshop Distrib. Cryptocurrencies Consensus Ledgers*, vol. 310, 2016, pp. 1–4.

[16] J. H. Jeon, K.-H. Kim, and J.-H. Kim, "Block chain based data security enhanced IoT server platform," in *Proc. Int. Conf. Inf. Netw. (ICOIN)*, Jan. 2018, pp. 941–944.

[17] A. Dorri, S. S. Kanhere, R. Jurdak, and P. Gauravaram, "Blockchain for IoT security and privacy: The case study of a smart home," in *Proc. IEEE Int. Conf. Pervas. Comput. Commun. Workshops (PerCom Workshops)*, Mar. 2017, pp. 618–623.

[18] M. Mylrea and S. N. G. Gourisetti, "Blockchain for smart grid resilience: Exchanging distributed energy at speed, scale and security," in *Proc. Resilience Week (RWS)*, Sep. 2017, pp. 18–23.

[19] X. Liang, S. Shetty, D. Tosh, C. Kamhoua, K. Kwiat, and L. Njilla, "ProvChain: A blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability," in *Proc. 17th IEEE/ACM Int. Symp. Cluster, Cloud Grid Comput. (CCGRID)*. Piscataway, NJ, USA: IEEE Press, May 2017.

[20] W. Tsai *et al.*, "Big data-oriented blockchain for clearing system," *Big Data Res.*, vol. 1, pp. 22–35, 2018.

[21] I. Karamitsos, M. Papadaki, and N. B. A. Barghuthi, "Design of the blockchain smart contract: A use case for real estate," *J. Inf. Secur.*, vol. 9, no. 3, pp. 177–190, 2018.

[22] P. K. Sharma, M.-Y. Chen, and J. H. Park, "A software defined fog node based distributed blockchain cloud architecture for IoT," *IEEE Access*, vol. 6, pp. 115–124, 2018.

[23] X. Liang, J. Zhao, S. Shetty, and D. Li, "Towards data assurance and resilience in IoT using blockchain," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Oct. 2017, pp. 261–266.

[24] A. Bahga and V. K. Madisetti, "Blockchain platform for industrial Internet of Things," *J. Softw. Eng. Appl.*, vol. 09, no. 10, pp. 533–546, 2016.

[25] L. Aniello, R. Baldoni, E. Gaetani, F. Lombardi, A. Margheri, and V. Sassone, "A prototype evaluation of a tamper-resistant high performance blockchain-based transaction log for a distributed database," in *Proc. 13th Eur. Dependable Comput. Conf. (EDCC)*, Sep. 2017, pp. 151–154.

[26] J. Wang, M. Li, Y. He, H. Li, K. Xiao, and C. Wang, "A blockchain based privacy-preserving incentive mechanism in crowdsensing applications," *IEEE Access*, vol. 6, pp. 17545–17556, 2018.

[27] H. Es-Samaali and A. Outchakoucht, "A blockchain-based access control for big data," *Int. J. Comput. Netw. Commun. Secur.*, vol. 5, no. 7, p. 137. 2017.

[28] A. Ouaddah, A. Abou Elkalam, and A. Ait Ouahman, "FairAccess: A new blockchain-based access control framework for the Internet of Things," *Secur. Commun. Netw.*, vol. 9, no. 18, pp. 5943–5964, Dec. 2016.

BIN LIU received the B.S. degree in computer science and technology from Yangtze University, in 2005, and the M.S. degree in computer science and technology from Xiamen University, in 2010. He is currently an Assistant Professor with the School of Software Engineering, Xiamen University of Technology, Xiamen, China. His current research interests include steganography, real-time embedded systems, field-programmable gate arrays, digital integrated circuits intellectual property rights protection, EEG data analysis, and identity security identification.

LIJUN XIAO received the bachelor's degree from the Hunan University of Technology and Business, in 2014, and the master's degree from the Hunan University of Science and Technology, Xiangtan, in 2017. From 2017 to 2018, she was an Assistant Researcher with the Hunan University of Science and Technology. She has published six refereed journals and conference papers. She has been an Assistant Professor with the Guangzhou College of Technology and Business, Guangzhou, since 2009. Her current research interests include the theories of software engineering, IP protection, and software security.

JING LONG received the M.S. degree from the College of Computer Science and Engineering, Hunan University of Science and Technology, China, in 2012, and the Ph.D. degree in computer science and electronics engineering from Hunan University, China, in 2018. She is currently a Lecturer with the College of Information Science and Engineering, Hunan Normal University. Her current research interests include hardware security, IP protection, the Internet of Things, and network security.

MINGDONG TANG received the B.S. degree in electrical engineering from Tianjin University, Tianjin, China, in 2000, the M.S. degree in control engineering from Shanghai University, Shanghai, China, in 2003, and the Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2010. He is currently a Professor with the School of Information Science and Technology, Guangdong University of Foreign Studies, Guangzhou, China. He is also with the Guangdong Provincial Key Laboratory of Computational Intelligence and Cyberspace Information, South China University of Technology, Guangzhou. He has published more than 100 peer-reviewed scientific research articles in various journals and conferences. His research interests include service-oriented computing, data mining, and blockchain.

OSAMA HOSAM received the M.Sc. degree in computer systems and engineering from Azhar University, in 2007, and the Ph.D. degree in computer science and engineering from Hunan University, China, in 2011. He is a Research Associate with SRTA-City, Alexandria, Egypt. He worked with the Nanjing University of Technology. He worked as an Assistant Professor with the College of Computer Science and Engineering at Yanbu, in 2013. In 2017, he was promoted to be an Associate Professor in the fields of computer and information security.