Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

A Blockchain-based Medical Data Sharing and Protection Scheme

XIAOGUANG LIU^{1,2,3}, ZIQING WANG³, CHUNHUA JIN⁴, FAGEN LI³,(Member, IEEE), and GAOPING LI^{1,2}

¹The Key Laboratory for Computer Systems of State Ethnic Affairs Commission, Southwest Minzu University, Chengdu, Sichuan 610041, P. R. China

²School of Computer Science and Technology, Southwest Minzu University, Chengdu, Sichuan 610041, P. R. China

³Center for Cyber Security, School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, P. R. China

⁴The Laboratory for Internet of Things and Mobile Internet Technology of Jiangsu Province, Huaiyin Institute of Technology, Huaian, Jiangsu, 223003, P. R. China

Corresponding author: Xiaoguang Liu (e-mail: 21700128@swun.edu.cn).

This work is supported by the Fundamental Research Funds for the Central Universities of Southwest University for Nationalities (No: 2019NQN26), Key Fund Project of Sichuan Provincial Department of Education (No: 17ZA0414), the Laboratory for Internet of Things and Mobile Internet Technology of Jiangsu Province (No: JSWLW-2017-006), and the National Natural Science Foundation of China (No: 61273311).

ABSTRACT Electronic health record (EHR) has recorded the process of occurrence, development, and treatment of diseases. So it has high medical value. Owing to the private and sensitive nature of medical data for patients, the data sharing and privacy preservation are critical issues in EHR. Blockchain technology may be a promising solution for the problems above since it holds the features of decentralization and tamper resistance. In the paper, we propose a medical data sharing and protection scheme based on the hospital's private blockchain to improve the electronic health system of the hospital. Firstly, the scheme can satisfy various security properties such as decentralization, openness, and tamper resistance. A reliable mechanism is created for the doctors to store medical data or access the historical data of patients while meeting privacy preservation. Furthermore, a symptoms-matching mechanism is given between patients. It allows patients who get the same symptoms to conduct mutual authentication and create a session key for their future communication about the illness. The proposed scheme is implemented by using PBC and OpenSSL libraries. Finally, the security and performance evaluation of the proposed scheme is given.

INDEX TERMS Blockchain, Electronic health record, Medical data, Sharing and protection, Symptoms-matching.

I. INTRODUCTION

ITH the development of computer and communication technology, EHR has become an indispensable tool for medical services [1]. The system utilizes some electronic devices such as the computer to deal with digital medical records, and it has the advantages of easy to use, stronger timeliness, and low cost. EHR not only provides the most useful data for diagnosis and scientific research but also it gives one kind of judgment basis for handling medical disputes. So, it has attracted a wide range of attention including the government, the medical community, cybersecurity department, and so on [2], [3]. Because the medical data is crucial for the diagnosis, and it is personal and sensitive for patients. Thus, data sharing and privacy preservation issues are critical in EHR. The medical data should be stored,

managed, and accessed securely. Notably, the doctor usually needs to know the medical history of the patient when he/she makes the diagnosis or treatment. However, the patient can not professionally describe his/her medical history, which will affect the latest treatment. Thus, in EHR, historical medical data generated by different doctors in different hospitals should be capable of being securely and timely queried by a legitimate doctor with the patient's consent, please see [4]–[7] for more details.

In recent years, the EHR system is markedly developed with the rise of cloud computing. For example, in [8], authors first expounded the security requirements of the EHR system based on cloud computing. Also, some suggestions are suggested to ensure the security of medical data in the cloud. In [9], the attribute-based encryption is utilized to protect

the data in the cloud, and then the proposed EHR system is implemented in an android phone. In [10], Xhafa et al. proposed an attribute-based EHR with privacy awareness in cloud computing. However, as mentioned in [11], [12], these cloud-based schemes have some flaws. For example, they have a dependency on the cloud provider. If some targeted attacks to cloud provider are carried out, then the information leakage is likely to occur. Additionally, the server may suddenly stop if the cloud providers would go bankrupt or be swallowed up by the larger companies. That is, the security of EHR will be threatened. In 2008, the blockchain structure was proposed [13]. It can be viewed as a distributed database and satisfies the features of decentralization, tamper resistance, and asymmetric encryption. This technology can provide a reliable way to manage and store data. So it may be a promising solution for EHR. At present, the blockchainbased researches for EHR have already started attracting attention from medicine. How to design an efficient and secure EHR system by using blockchain is their core task

A. RELATED WORK

[14]–[17].

In 2015, a decentralized personal data management system was presented in [18]. It can ensure the users own and manage their data. In the system, the blockchain is converted into an automatic access control manager in the protocol without a trusted-third-party. In 2016, a decentralized "MedRec" system based on blockchain was proposed to handle EHR [19]. MedRec has contributed to the emergence of data economics. It also provides researchers with big data while allowing patients and providers to choose to publish metadata. In 2017, Xue et al. [20] designed a blockchain-based sharing model for medical data. The scheme solves the problem of checking, saving, and synchronizing medical data among different medical institutions by improving the consensus mechanism. But it has some disadvantages in data storage since the scheme does not possess the ability of machine learning algorithm. Xia et al. [21] designed a blockchainbased data sharing framework. It takes the advantages of blockchain's immutability and the built-in autonomy to address access control challenges related to sensitive data stored in the cloud. At the same year, Xia et al. [22] also proposed a system named MeDShare, which is based on blockchain and has minimal data privacy risks. It is used to solve the problem of medical data sharing among healthcare big data custodians (e.g., cloud service providers) in the untrusted environment. The two schemes have the weaknesses of the cloud since they still need the assistance of the cloud. In 2018, Yang et al. [23] presented a blockchain-based architecture for EHR. It prevents tampering and misuse of EHR by keeping track of all events occurring in the database. Also, the system introduces a new incentive mechanism to create new blocks in the blockchain. In [24], a medical data storage system based on blockchain was proposed. The system not only can guarantee the originality and verifiability of stored medical data but also can preserve the privacy of patients. In [25],

Zhang et al. proposed a medical data sharing scheme based on blockchain to improve the diagnosis level. They utilize the private blockchain possessed by the hospital to store personal health data of patients while the consortium blockchain is used to keep the security indexes. Notably, authors have described the details of the scheme and implemented it on JUICE. Nevertheless, it needs substantial computational and communication cost.

B. MOTIVATION AND CONTRIBUTION

Research on medical sharing schemes based on blockchain is still in its infancy at present. The existing schemes have the following drawbacks: (1) Most schemes only give the framework and don't describe the specific details for implementation [21], [22]. (2) Although the details are given in some schemes, the cost of computation and communication is high [25]. The motivation of this paper is to design a medical data sharing scheme based on blockchain. It is helpful to the storage, management, and sharing of the medical data. The scheme should satisfy the security requirements in medical data sharing schemes. Also, it should have low computational and communication cost. The main contributions of this paper are listed as follows.

- A lightweight medical data sharing and protection model is proposed, which is based on blockchain. Utilizing the proxy re-encryption technology, the model could make data sharing among doctors from different hospitals. The stored medical information is very secure and could not be easily tampered since they are stored in the blockchain.
- 2) An improved consensus mechanism is proposed by improving the traditional delegated proof of stake. It is secure, reliable, and efficient.
- 3) We design a symptoms-matching mechanism for patients who register in different hospitals and have the same disease symptoms. One session key could be set between the patients after they make mutual authentication. The mechanism can help patients to communicate the disease information.

C. ORGANIZATION OF THIS PAPER

The rest of paper is organized as follows. Firstly, some preliminaries are presented in section II. In section III, we give one medical data sharing and protection model based on blockchain. In section IV, we offer the security and performance analysis of the proposed scheme. Finally, the paper is concluded in section V.

II. PRELIMINARIES

A. BLOCKCHAIN

Blockchain mainly solves the trust and security issues of transactions, and it is a kind of distributed database combining data blocks in chronological order. Generally, the blockchain is divided into three classes: private blockchain, consortium blockchain, and public blockchain [15], [26]. As

FIGURE 1: The structure of the blockchain

shown in Figure 1, each blockchain consists of many blocks, and each block contains a block header and a block body. Block header contains multiple meta-information about the current block. For example, timestamp, a hash value for the blockchain body, and a hash value for the previous block. Block body is usually used to record the real data of the current transactions. The main features of the blockchain are as follows [26]:

- Decentralization: there is no central node, and each node is equal. Transaction records are done by multiple nodes that distributed in different places, and each node records and keeps a complete account. All nodes can supervise the transaction and jointly testify for it.
- 2) Tamper resistance: the Hash value of the previous block is contained in the latter block. If one of the blocks is modified, then all the blocks after that will be recalculated. So the modification of the database by a single node is invalid.
- 3) Openness: in addition to the private information of all parties involved in the transaction being encrypted, the data of the blockchain are open to all. Anyone can query block data and develop relevant applications through the public interface.
- 4) Autonomy: the blockchain adopts a consensual protocol (such as an open and transparent algorithm), which enables all nodes in the system to freely and securely exchange data. So, it will not be intervened by a human.
- 5) Anonymity: the exchange between nodes follows a fixed algorithm, so the counter party does not need to make the other party trust it through public identity.

B. GENERAL NETWORK MODEL OF THE BLOCKCHAIN-BASED MEDICAL DATA SHARING AND PROTECTION SCHEME

As shown in Figure 2, the general network model of the blockchain-based medical data sharing and protection scheme is composed of three parties, i.e., system manager, user (patient), and hospital. In the model, the role of system manager (denoted as SM) is usually played by some trusted authorities such as government departments. It is responsible for the management of the whole system. When a user needs to see a doctor, he/she first registers with the hospital. Then,

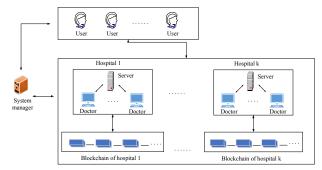


FIGURE 2: General network model of the blockchain-based medical data sharing and protection scheme

the hospital will arrange for a doctor to make a diagnosis for him/her. When the visit is over, the doctor will store them in the blockchain if the medical results have passed the verification by the verifiers. Especially, any doctor with the patient's permission can access the patient's historical data stored in the blockchain when needed.

C. BASIC REQUIREMENTS FOR MEDICAL DATA SHARING AND PROTECTION SCHEME

Ideal medical data sharing and protection scheme should satisfy the following basic requirements, i.e., security and privacy protection, data access, patient control (user engagement), and unified standard [20], [22].

- Security and privacy protection: medical data could not be illegally used by anyone. The scheme should be able to resist malicious attacks, and illegal behavior could be traced.
- Data access: after being authorized, patients can see all their medical records and doctors can access previous medical information under the authorization of patients.
- 3) Patient control: the patient could manage his/her historical medical records, i.e., anyone could not acquire the historical data without the patient's agreement.
- 4) Unified standard: in the model, all participants should use unified data standard and management scheme, which are helpful to implement the data sharing and improve system stability.

D. DELEGATED PROOF-OF-STAKE

Blockchain utilizes the consensus mechanism to ensure that all legitimate nodes maintain the same global ledger. Delegated Proof-of-Stake (DPOS) is an efficient and reliable consensus mechanism [27], [28]. Similar to the board vote, holders of coins select some nodes to vote on behalf of everyone in DPOS. It could improve the efficiency of reaching a consensus. DPOS's process is that everyone who owns the coins to vote and generate 101 delegates firstly. The delegates could be seen as supernodes that have equal rights to each other. Then these supernodes are responsible for generating a new block in turn. If delegates fail to perform their duties

(e.g., when their turn comes, they fail to compute the right value), the network will select new supernodes to replace them, and the old nodes will be punished.

E. PROXY RE-ENCRYPTION

To ensure the security in the data sharing, the proxy reencryption was presented in [29]. In these schemes [29]– [31], one party A entrusts a trusted third party or a semihonest agent to transform the ciphertext encrypted with its public key into ciphertext encrypted with the other party B's public key. Then, B could decrypt the ciphertext with own private key, i.e., the data sharing is realized. During the whole process, the data encrypted is very secure, and A's private key does not have to be disclosed. The specific steps are listed as follows:

- 1) A encrypts the plaintext M with own public key, i.e., $C_A = E_A(M)$, where M is what A wants to give B, and E is an asymmetric encryption algorithm such as classical RSA.
- 2) B sends the request to A, and then A (or the agent) generates one conversion key $PK_{A \leftrightarrow B}$.
- 3) A sends C_A and $PK_{A \leftrightarrow B}$ to the agent.
- 4) The agent converts the ciphertext C_A into C_B using PK_{A⇔B}. Here, C_B is the ciphertext of M encrypted with B's public key. In the step, the agent only provides transformation service, and it cannot obtain the plaintext.
- 5) The agent sends the ciphertext C_B to B.
- B decrypts C_B with own private key to get the plaintext M.

F. BILINEAR MAPS

Let \mathbb{G}_1 and \mathbb{G}_2 denote two cyclic multiplicative groups with same prime order p. The mapping $e: \mathbb{G}_1 \times \mathbb{G}_1 \to \mathbb{G}_2$ is a bilinear map if it satisfies the following conditions [32]:

- 1) Bilinear: $e(U^a, V^b) = e(U, V)^{ab}$ holds for any two points $U, V \in \mathbb{G}_1$ and any two points $a, b \in \mathbb{Z}_p^*$.
- 2) Nondegeneracy: there exists two points $U, \dot{V} \in \mathbb{G}_1$ such that $e(U, V) \neq 1_{\mathbb{G}_2}$, where $1_{\mathbb{G}_2}$ is the identity element of \mathbb{G}_2 .
- 3) Computability: e(U, V) could be calculated efficiently in polynomial time for any two points $U, V \in \mathbb{G}_1$.

III. THE PROPOSED MEDICAL DATA SHARING AND PROTECTION SCHEME

In this section, we will propose a medical data sharing and protection scheme based on the private blockchain of the hospital. The two-way proxy re-encryption technology in [31] is utilized in the scheme. Also, it has provided a symptoms-matching mechanism for patients with the same disease symptoms. The notations used in this paper are given in Table 1.

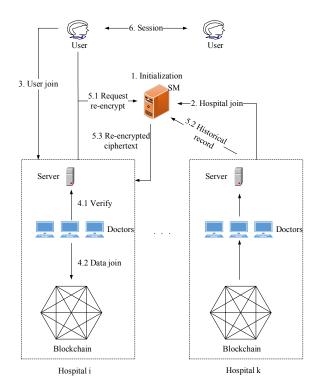


FIGURE 3: Proposed architecture

TABLE 1 Notations

Notations	Description
p, q	Two large prime numbers
SM	The system manager
HO_i	The <i>i</i> th hospital
$US_{i,j}$	The jth user of the ith hospital
$ID_{(.)}$	The identity
$PID_{(.)}$	The pseudo identity
$PK_{(.)}$	The public key
$SK_{(.)}$	The private key
$\mathbb{G}_1,\mathbb{G}_2$	Two multiplicative groups with same order p
k, l, l_1	Three security parameters
g	A generator of the group \mathbb{G}_1
\overline{F}	A random function
E(.)	Encryption
D(.)	Decryption
	String concatenation
$T \subset Z_p^*$	The set of disease symptoms
H_1, H_2^r, H_3	Three secure hash functions
MAC	The message authentication code
K	The session key
e	The bilinear maps

As shown in Figure 3, the system manager SM, the hospital HO_i , and the user $US_{i,j}$ are the three kinds of participants in the network. SM is played by the health management department that is a trusted third party and responsible for generating the master key and system parameters. Hospital HO_i first registers with SM and then generates its private key and public key. If a user $US_{i,j}$ sees a doctor in the hospital HO_i , he/she must register with HO_i and set his/her private key and public key. When the diagnosis has finished, the doctor will broadcast the results in the blockchain. If they have passed the verification by the server, the medical results of $US_{i,j}$ will be stored in the blockchain of HO_i . If a doctor in any hospital wants to query the historical records of the

patient $US_{i,j}$, he/she and the patient should apply to the SM simultaneously. SM will compute the conversion key and generate the ciphertext of the historical records re-encrypted by the doctor's public key. Then SM sends the ciphertext to the doctor. Finally, any two patients $US_{i,j}$ and $US_{i+1,j+1}$ could conduct a mutual authentication and set a session key for their future session. Our scheme includes the following six phases, i.e., the initialization phase, the hospital join phase, the user join phase, the data join blockchain phase, the data search and sharing phase, and patients session phase.

A. INITIALIZATION PHASE

- 1) SM first inputs a security parameter $1^k(k \in N)$, chooses two multiplicative groups \mathbb{G}_1 and \mathbb{G}_2 , and the bilinear map e, where the two groups have the same prime order p, and q is a generator of \mathbb{G}_1 . Then SM picks three secure hash functions $H_1:\{0,1\}^*\to\mathbb{G}_1$, $H_2:\mathbb{G}_2\to\{0,1\}^k$, and $H_3:\mathbb{G}_1\times\{0,1\}^k\times\{0,1\}^l\to\mathbb{Z}_p^*$, and a random function $F:\mathbb{G}_2\times\mathbb{G}_1\times\{0,1\}^k\to\{0,1\}^{l-l_1}\|\{0,1\}^{l_1}$, where l and l_1 both are security parameters. Finally, SM randomly selects $x\in\mathbb{Z}_p^*$ as the system master key, sets the public key $Y=g^x$, chooses random elements $g_1,g_2,u,v,d\in\mathbb{G}_1$, and publishes $\{p,g,g_1,g_2,u,v,d,Y,H_1,H_2,H_3,F,l,l_1,\mathbb{G}_1,\mathbb{G}_2\}$.
- 2) The user $US_{i,j}$ randomly selects $x_j \in \mathbb{Z}_p^*$ as his/her private key and computes the public key $PK_{i,j} = g^{x_j}$.
- 3) The hospital HO_i randomly selects $x_i \in \mathbb{Z}_p^*$ as the private key and its public key is set as $PK_i = g^{x_i}$.
- 4) The doctor S of HO_i randomly selects $x_s \in \mathbb{Z}_p^*$ as the private key and computes its public key $PK_s = g^{x_s}$.

B. HOSPITAL JOIN PHASE

If a new hospital HO_i plans to join into the network, it must execute the following steps combining with SM.

- 1) HO_i sends its identity ID_i to SM.
- 2) If the identity is legal, SM randomly selects $\lambda_i \in \mathbb{Z}_p^*$ and computes $PID_i = E_{SM}(ID_i \oplus \lambda_i || \lambda_i)$ as HO_i 's pseudo identity.
- 3) SM sends PID_i to HO_i through a secure channel.

C. USER JOIN PHASE

If a patient $US_{i,j}$ sees a doctor in the hospital HO_i , he/she needs to do the following steps, where the index j indicates the patient is the jth patient of HO_i .

- 1) $US_{i,j}$ submits the identity $ID_{i,j}$ to the server of HO_i , and then the server assigns a doctor S to give the diagnosis for $US_{i,j}$. Meanwhile, the server randomly selects $\alpha \in \mathbb{Z}_p^*$ as the evidence for the user, sends α to $US_{i,j}$, and stores it for the doctor S.
- 2) When $US_{i,j}$ visits the doctor S, $US_{i,j}$ will show α as the consent to make a diagnosis or access historical records of $US_{i,j}$. S gives the diagnosis result m, extracts a symptom $t_{i,j} \in T$, randomly selects $\lambda_{i,j} \in \mathbb{Z}_p^*$, computes $US_{i,j}$'s pseudo identity $PID_{i,j} = E_s(ID_{i,j} \oplus \lambda_{i,j} || \lambda_{i,j})$. Then,

S inputs $PK_{i,j},\ Y,\ m,$ and $t_{i,j},$ randomly selects $r\in\mathbb{Z}_p^*,$ computes $C_1=g_1^r,\ C_2=PK_{i,j}^r,$ $U=e(g,g_2^{t_{i,j}})^r,\ C_3=H_2(U),\ K=e(g,g)^r,$ $C_4=[F(K,C_1,C_3)]_{l-l_1}\parallel([F(K,C_1,C_3))]_{l_1}\oplus m),$ $h=H_3(C_1,C_3,C_4),$ and $C_5=(u^hvd)^r.$ Thus, the ciphertext of m is $C_{i,j}=(C_1,C_2,C_3,C_4,C_5).$ S sends $(PID_i,PID_{i,j},t_{i,j})$ to $US_{i,j}$ securely. Also, S computes $X_{i,j}=E_i(\alpha\|ID_{i,j}\|PID_i\|ID_s)$ with HO_i 's public key and sends it to the server of the hospital.

D. DATA JOIN BLOCKCHAIN PHASE

In DPOS, legitimate participants need to ballot 101 delegates to record data in the blockchain in turn. In the hospital, doctors from different departments have unique professional knowledge. So, the general DPOS is not very suitable for the blockchain of hospital since how to elect delegates is a thorny problem. Also, the election of delegates will consume computational cost and communication cost. In our scheme, a lightweight and efficient consensus mechanism is proposed, please see Algorithm 1. It could be regarded as an improvement of DPOS. Every doctor is seen as the delegate and responsible for broadcasting and recording data generated by themselves in the blockchain. The server of the hospital is chosen as the only supernode, i.e., the verifier. Especially, we set up one credit score scheme for hospitals and doctors to ensure our mechanism is reliable. SM and server of the hospital have the right to check the effectiveness of every record of the doctor. When there are wrong records, the credit score of the doctor will be reduced. If this score reaches a lower bound, the doctor will be expelled from the hospital. Similarly, hospitals are supervised by SM and doctors. The SM will punish the hospital if it has some illegal behaviors. Doctors have the right to report these behaviors of the hospital to SM.

Algorithm 1 Improved consensus mechanism

- 1: a doctor in hospital HO_i broadcasts the diagnosis result $C_{i,j}$ in the blockchain
- 2: the server of HO_i verifies the data
- 3: if the data passes the verification
- 4: the data is placed in one new block of the blockchain
- 5: else
- 6: return FALSE
- 7: end if

In the hospital, one doctor generates the medical data for $US_{i,j}$, i.e., $C_{i,j} = (C_1, C_2, C_3, C_4, C_5)$ and he/she also responsible for broadcasting the data in the private blockchain of HO_i . The server of HO_i is selected as the only verifier since it is supervised by SM and managed by the core division of the hospital. The process is that the server first decrypts the ciphertext $X_{i,j}$ by HO_i 's private key and checks the identity information of the doctor and patient. If it passes verification, the data is accepted in the blockchain, and all nodes update the records. The structure of the block is shown

FIGURE 4: The structure of a block in the blockchain

in Figure 4. The specific steps are listed as follows.

- 1) Doctor S broadcasts the medical data in the private blockchain of the hospital.
- The server of the hospital verifies the data every minute, and then every ten legitimate records are placed in one new block of the blockchain.
- 3) Other nodes of blockchain update their stored data.

E. DATA SEARCH AND SHARING PHASE

When the patient $US_{i,j}$ interacts with a doctor S in hospital HO_i , S may need to know the historical medical records of the patient in hospital HO_k for more precise diagnosis. Thus, the following steps should be executed by inputting $PK_{i,j} = g^{x_j}$ and $PK_s = g^{x_s}$ if the doctor has obtained the permission of $US_{i,j}$.

- 1) S and $US_{i,j}$ send their private keys and identities to SM respectively, then SM computes the re-encryption key $rk_{j\leftrightarrow s} = x_s/x_j \mod p$.
- 2) SM sends an extraction instruction about $US_{i,j}$'s medical records to the hospital HO_k .
- 3) The server of HO_k sends the encrypted historical records to SM.
- 4) SM first computes $h=H_3(C_1,C_3,C_4)$, if $e(C_1,PK_{i,j}u^hvd)=e(g_1,C_2C_5)$ holds, SM computes $C_2'=C_2^{rk_j\leftrightarrow s}=PK_s^r$, sends the ciphertext $C_s=(C_1,C_2',C_3,C_4,C_5)$ to S through the server of HO_i . Otherwise, SM outputs \bot .
- 5) $US_{i,j}$ computes $U_1 = g^{r'}$, $U_2 = (g_2^{t_{i,j}})^{1/x_j}H_1(PK_i^{r'})$, and sends $U_\alpha = (U_1, U_2)$ to the server of HO_i , where $r' \in \mathbb{Z}_p^*$ is a random number.
- 6) The server of HO_i ensures $e(C_1, PK_{i,j}u^hvd) = e(g_1, C_2C_5)$ holds. If not, the phase is terminated. Otherwise, the server computes $U = U_2/H_1(U_1^{x_i})$ and ensures $C_3 = H_2(e(C_2, U))$ is true. Otherwise, the phase is terminated.
- 7) S computes $K = e(C'_2, g)^{1/x_s}$, if $[F(K, C_1, C_3)]_{l-l_1} = [C_4]_{l-l_1}$, then S recovers $m = [C_4]_{l_1} \oplus [F(K, C_1, C_3)]_{l_1}$. Otherwise, this phase is terminated.

F. PATIENTS SESSION PHASE

As shown in Figure 5, an interaction program will be given for users $US_{i,j}$ and $US_{i+1,j+1}$ in this section. They have the same disease symptoms and hope to make further

communication about their illness. We only consider weak security requirements here since the following two reasons. (a) Since the information can be used to interact with each other, so it has no very strong privacy. (b) It can reduce the computational cost and communication cost. The details are given below.

- 1) $US_{i,j}$ sends PID_i and $PID_{i,j}$ to $US_{i+1,j+1}$, and then $US_{i+1,j+1}$ sends PID_{i+1} and $PID_{i+1,j+1}$ to $US_{i,j}$.
- 2) $US_{i,j}$ selects a secret integer $n_{i,j} \in \mathbb{Z}_p^*$ and a prime number $z \in \mathbb{Z}_p^*$ randomly, computes $w = z^{-1} \mod p$, $P_{i,j} = g^{zn_{i,j}}$, and $Q_{i,j} = g^{wn_{i,j}}$. Then $US_{i,j}$ sends $(z, P_{i,j}, Q_{i,j})$ to $P_{i+1,j+1}$. $P_{i+1,j+1}$ randomly choose $n_{i+1,j+1} \in \mathbb{Z}_p^*$, computes $w = z^{-1} \mod p$, $P_{i+1,j+1} = g^{zn_{i+1,j+1}}$ and $Q_{i+1,j+1} = g^{wn_{i+1,j+1}}$, and sends message $(P_{i+1,j+1}, Q_{i+1,j+1})$ to $US_{i,j}$.
- and sends message $(P_{i+1,j+1},Q_{i+1,j+1})$ to $US_{i,j}$. 3) $US_{i,j}$ computes $k_{i,j} = Q_{i+1,j+1}^{n_{i,j}}$, $MAC_{i,j} = MAC_{k_{i,j}}(Q_{i,j},w,P_{i+1,j+1},PID_{i+1,j+1},t_{i,j})$ and sends $MAC_{i,j}$ to $US_{i+1,j+1}$. Then $US_{i+1,j+1}$ computes $k_{i+1,j+1} = Q_{i,j}^{n_{i+1,j+1}}$ and $MAC_{i+1,j+1} = MAC_{k_{i+1,j+1}}(PID_{i+1,j+1},P_{i+1,j+1},Q_{i,j},\omega,t_{i,j})$. If $MAC_{i+1,j+1} = MAC_{i,j}$ holds, $US_{i+1,j+1}$ computes $MAC_{i+1,j+1}^{\dagger} = MAC_{k_{i+1,j+1}}(PID_{i,j},P_{i,j},Q_{i+1,j+1},k_{i+1,j+1},t_{i+1,j+1})$ and sends it to $US_{i,j}$. Otherwise, the prase is terminated.
- 4) $US_{i,j}$ computes $MAC_{i,j}^{\dagger} = MAC_{k_{i,j}}(PID_{i,j}, P_{i,j}, Q_{i+1,j+1}, k_{i,j}, t_{i,j})$. If $MAC_{i,j}^{\dagger} = MAC_{i+1,j+1}^{\dagger}$ is true, then $US_{i,j}$ computes the session key $K = P_{i+1,j+1}^{n_{i,j}}$, the ciphertext $\widetilde{K} = E_{k_{i,j}}(K)$, and sends \widetilde{K} to $US_{i+1,j+1}$. Otherwise, it is terminated.
- 5) $US_{i+1,j+1}$ decrypts ciphertext K to get the session key K.

If the disease symptoms $t_{i,j}$ and $t_{i+1,j+1}$ are same, the correctness of the protocol is based on the following equation.

$$k_{i+1,j+1} = Q_{i,j}^{n_{i+1,j+1}} = g^{\omega n_{i+1,j+1} n_{i,j}} = Q_{i+1,j+1}^{n_{i,j}} = k_{i,j}.$$

IV. ANALYSIS OF THE MODEL

In this section, we will evaluate the proposed scheme from the following three aspects. (1) Whether the proposed scheme can satisfy the basic requirements described for medical data sharing and protection scheme. (2) According to six factors (no payment, the consensus mechanism, based on the private blockchain, reduce the pressure of the main chain, the demand for calculating power, and symptoms-matching), the comparative analysis method is adopted to compare the proposed scheme with the existing blockchain-based medical data sharing and protection schemes [16], [20], [25]. (3) The comparison of the computational cost and communication cost about the scheme [25] and the proposed scheme will be given (as mentioned in [25], few details are given in the existing blockchain-based medical data sharing and protection schemes including [16], [20], so we only select the scheme in [25] as the comparative scheme for computational cost and communication cost). Additionally, we will implement the proposed scheme by using PBC and OpenSSL libraries.

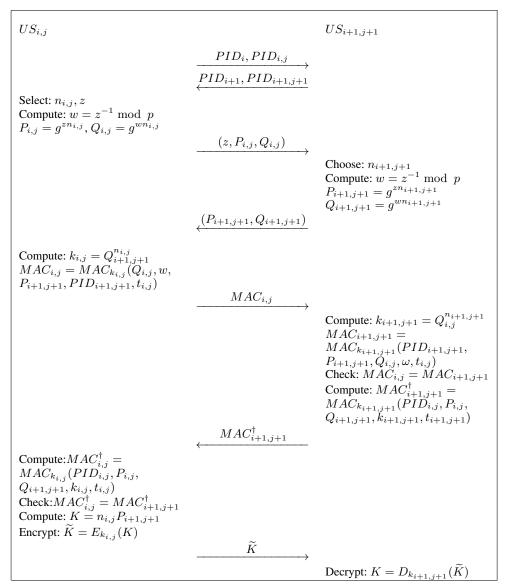


FIGURE 5: The protocol for patients session

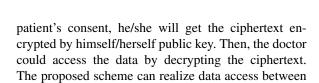
A. THE SOLUTIONS FOR THE BASIC REQUIREMENTS

Our scheme satisfies the five important features described in subsection A of the preliminaries section since it is based on the blockchain. Next, we will analyze the solutions of the proposed model for the basic requirements listed in subsection C of the preliminaries section.

1) Security and privacy. At the time of registration, the hospital or patient will be checked to ensure that all participators of the network are legitimate. After the hospital registers with SM, SM will generate a pseudo identity for the hospital. When one patient sees a doctor, doctor will also compute a pseudo identity for the patient. Thus, user privacy will be protected since the pseudo identity is used instead of true identity in the subsequent processes. All information placed in the blockchain is encrypted by the asymmetric en-

cryption scheme, which could prevent unauthorized nodes from accessing the medical information. When a doctor queries the historical data of a patient, the proxy re-encryption technology is used. It allows the stored information in the blockchain to be transmitted in the ciphertext state. So, the security of the proposed scheme is further improved. If two patients with the same symptoms want to communicate the disease information, and they must make mutual authentication and set a session key to preventing information leakage. In all phases, anyone except the doctor and the patient is unable to obtain the plaintext of medical data. So, the proposed scheme has better security and stability.

2) Data access. The proxy re-encryption is utilized in the proposed scheme. If a doctor has obtained the



different medical institutions. Patients also can query

3) Patient control. The medical records are stored in the blockchain of the hospital. If one legal doctor wants to obtain the stored data in the blockchain, he/she must have the re-encryption key issued by SM. The key is generated by SM utilizing the doctor and patient's private keys. So, patients could control access to data.

their medical records after applying to the hospital.

4) Unified standard. In the proposed model, we use the uniform standard of data such as the keywords of disease symptoms, which is beneficial to data sharing and protection.

B. PERFORMANCE ANALYSIS

In the proposed data sharing and protection scheme, an improved DPOS mechanism is proposed. It does not need nodes to vote and generate delegates, which could reduce the computational cost and communication cost. Every doctor is responsible for broadcasting the message generated by himself/herself in the hospital's private blockchain. The server of the hospital is seen as the only supernode and used to check the information. Then, other nodes in the blockchain will update the stored data if the information has passed the verification. Especially, doctors and hospitals both are supervised by the credit score mechanism.

As shown in Table 2, we will first compare the based-blockchain three medical data sharing schemes [16], [20], [25] with the proposed scheme from the following six factors, i.e., no payment, based on the private blockchain, the consensus mechanism, reduce the pressure for the main chain, the demand for calculating power, and symptoms-matching, they are denoted as F1, F2, F3, F4, F5, and F6 for convenience.

TABLE 2 Comparison of the six factors

	F1	F2	F3	F4	F5	F6
[16]	×	×	POW	×	Big	×
[20]	\checkmark	\checkmark	Improved DPOS	\checkmark	Small	×
[25]			DBFT		Big	×
Ours	\checkmark	\checkmark	Improved DPOS	\checkmark	Small	\checkmark
Support 1/. Not-support ×						

The scheme in [16] uses the POW consensus mechanism and needs to pay for the nodes that participate in the consensus mechanism. It could not satisfy F1, F2, F4, and F6, and requires a big calculating power. In [20], Xue et al. proposed a blockchain-based data sharing and protection scheme. The scheme could satisfy F1, F2, and F4, but it has no symptoms-matching function between patients. Besides, an improved DPOS consensus mechanism is also proposed and used, but only the delegate nodes can record data. It will consume extra communication cost and time. In [25], the scheme adopts DBFT consensus mechanism. It stores personal medical

data in the private blockchain while the security indexes of personal health data are put in the consortium blockchain. The scheme could satisfy many requirements, but it has a high computational cost (please see Table 3) and could not provide the symptoms-matching function. Thus, our scheme has better performance according to the six factors.

Now, we will first compare the computational and communication cost of the scheme in [25] with the proposed scheme. Generally, the server and SM both could be regarded as a cluster head with sufficient computational and communication resource. So, we will only consider the burden of the patient and doctor. Three operators are considered, i.e., the scale multiplication operator in \mathbb{G}_1 (m), the exponentiation operator in the prime finite field (e), and the bilinear pairing operator (b). Then, we will implement the proposed scheme by using PBC and OpenSSL libraries.

TABLE 3 The comparison of the computational cost

Scheme	Patient	Doctor
[25]	7m	(17+n)m+7e+4b
Ours	4e	11e + 5b

In Table 3, we have listed the computational cost of the two schemes. It should be noted that the patients session phase does not necessarily occur, so we ignore it here. We can know that the patient's computational cost in [25] is 7m. In the proposed scheme, the patient's computational cost is 4e. We can note that the cost both is constant for different n and the gap is very small [33]. Here, the parameter n is the size of the disease keyword set, and it is usually a large number such as 1000 set in [25]. On the side of the doctor, the computational cost of our scheme is 11e + 5b. But the computational cost of the scheme [25] increases linearly with the large number n, so it has higher computational cost on this side. Thus our scheme is satisfactory for practical medical data sharing scheme.

The specific experimental environment for implementing the proposed scheme is as follows. The cryptographic primitives are implemented on a computer with Intel(R) Core(TM) i5-5200U CPU @ 2.20Ghz 2.19Ghz, 8 GB RAM, Manjaro Linux 64 bit operating system with KDE desktop, using C++ language. PBC library and OpenSSL library are used for the simulation. The version of PBC library is 0.5.14, and the version of OpenSSL library is 1.1.1c. We deployed five blockchain nodes to receive block information. One of the nodes is deployed on the computer used for the simulation and runs on a different port than the server program. We created four identically configured virtual machines on a computer running Windows 10 operating system to deploy the other four blockchain nodes. The computer running the virtual machine is configured as Intel(R) Core(TM) i5-8400 CPU @ 2.80GHz, 8 GB RAM and running Windows 10 64-bit Home Chinese version operating system. The virtual machine software is Oracle VM VirtualBox 5.2.22. The vir-

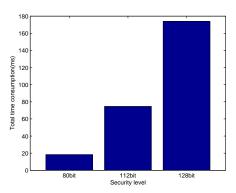


FIGURE 6: The total computational cost of experiments of the proposed scheme with different security levels

tual machine we created uses the Ubuntu 18.04.1 operating system with 1024MB of RAM and one CPU. The virtual machine uses the bridge mode to join the LAN segment where the computer running the server program is located.

TABLE 4 Experimental security level for different p and q

Security level (bit)	Size of p (bit)	Size of q (bit)
80	160	1024
112	224	2048
128	256	3072

The Ate pairing has been widely utilized in the public key cryptography. In this paper, we use a super singular curve $E(F_q)$ with order p over the finite field F_q , where p and q two large prime numbers. We have considered three kinds of AES key size security level, i.e., 80bit, 112bit, and 128bit [34]. Please refer to the corresponding p and q values in Table 4. Additionally, three security parameters k, l, l_1 are set as 256bit, 1024bit, and 512bit, respectively. The experimental results of computational cost are summarized in Table 5. We can know from the results that the computational cost difference is very small in the hospital join and data join two phases. The reason is that they do not involve three types of operators that have high computational costs, i.e., e, m, and b. However, the other three phases are different. Their computational cost gradually increases as the level of safety increases. In Figure 6, the total computational cost of experiments for different security levels is given. All values are the average result of 100 times experiments.

TABLE 5 Experimental results of the proposed scheme (ms)

	Security level		
Phases	80bit	112bit	128bit
Initialization	3.568	13.334	26.189
Hospital join	0.004	0.006	0.006
User join	3.366	11.259	25.150
Data join	1.956	1.892	1.902
Data search and sharing	9.449	48.149	120.742

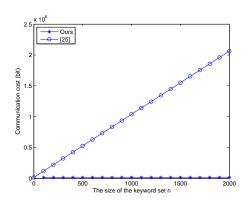


FIGURE 7: Communication cost comparison versus the size of the keyword set

TABLE 6 The comparison of the communication cost

Scheme	Communication cost
[25]	$(n+12) \mathbb{G}_1 + \mathbb{G}_2 + 5 \mathbb{Z}_p^* + 13\lfloor \frac{2}{3}n_p \rfloor +$
	t + x + ID + 2 Hash
Ours	$9 \mathbb{G}_1 + 2 \mathbb{Z}_p^* + 2 x + ID + 2k + 2l$

For the communication cost, we give the comparison results in Table 6. Here, the communication cost for the patient and doctor in the following three main phases is considered, i.e., the data broadcast, the data verification, and the data search and access. In the proposed scheme, the patient $U_{i,j}$ needs to send the trapdoor $U_{\alpha} = (U_1, U_2)$ to the server of the hospital HO_i , where U_1 and U_2 are elements of \mathbb{G}_1 . If the doctor wants to query $U_{i,j}$'s a historical record, $U_{i,j}$ will send the private key x_j to SM, where x_j is an element of \mathbb{Z}_{p}^{*} . For the doctor S, he/she needs to send the private key x_s to SM, and receives the encrypted historical record stored in another hospital, where x_s is the element of \mathbb{Z}_p^* , and the ciphertext of the historical record has the same size with $C_{i,j}$. Also, doctor S is responsible for broadcasting the ciphertext $C_{i,j} = (C_1, C_2, C_3, C_4, C_5)$, block ID, user pseudo identity, doctor's public key, and doctor's signature in the blockchain, where C_1 , C_2 , and C_5 are the elements of \mathbb{G}_1 , C_3 is an element with size k, C_4 is an element with size l, the user pseudo identity is an element of the general ciphertext space (the length of the element is denoted as |x|), the doctor's public key is an element of \mathbb{G}_1 , and the signature could be seen as an element of the general ciphertext space. So the communication cost of our scheme is $9|\mathbb{G}_1| + 2|\mathbb{Z}_p^*| + 2|x| + |ID| + 2k + 2l$.

In [25], a lot of communication cost need to be consumed on the patient side and doctor side. Known from the experimental results in [25], the communication cost of the scheme is $(n+12)|\mathbb{G}_1|+|\mathbb{G}_2|+5|\mathbb{Z}_p^*|+13\lfloor\frac{2}{3}n_p\rfloor+|t|+|x|+|ID|+2|Hash|$ in the three main phases, where n_p is the number of the private blockchain's verifiers, t is a timestamp. Apparently, the communication cost of [25] is higher than the proposed scheme since n is a large number. So the proposed scheme has better performance. Finally, in order to clearly

show the cost of communication, we set p and q two large prime numbers are 160 bits and 1024 bits respectively. The lengths of elements in \mathbb{G}_1 and \mathbb{G}_2 are 1024 bits and 512 bits separately. We assume that the lengths of the identity and the timestamp both are 32 bits, the point in the ciphertext space is 160 bits, the $n_p=3$, and the hash value is 256 bits. The comparison diagram of communication cost is given in Figure 7. It is easy to find that the communication cost of the proposed scheme is constant. However, as the size n of the keyword set increases, the communication cost of the scheme in [25] increases linearly. The communication cost of our scheme is significantly low.

V. CONCLUSION

The features of blockchain technology such as the decentralization and tamper resistance make it very suitable for the protection and sharing of medical data. In this paper, a lightweight medical data sharing scheme based on blockchain is proposed and implemented. Proxy reencryption technology is used to help the doctors to access historical records of patients. It can ensure the security of the proposed scheme since the inquired information is transmitted in the ciphertext form. Besides, an improved DPOS mechanism is proposed to act as the consensus mechanism that is lightweight and reliable. Finally, our scheme provided the symptoms-matching mechanism that allows two patients with the same symptoms can make communication about their illness. The analysis results show that the proposed scheme satisfies many requirements and has a low computational and communication cost.

A. REFERENCES

REFERENCES

- A K Jha, D Doolan, D Grandt, T Scott, and D W Bates, "The use of health information technology in seven nations," Int. J Med. Inform., vol.77, no.12, 2008, pp. 848–854.
- [2] Y Guo and C Liang, "Blockchain application and outlook in the banking industry," Financ. Innov., vol.2, no.24, 2016, pp. 1–12.
- [3] Y Yuan and F Y Wang, "Blockchain: the state of the art and future trends," Acta Automat. Sinica, vol.42, no.4, 2016, pp. 481–494.
- [4] B Shickel, P J Tighe, A Bihorac, and P Rashidi, "Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis," IEEE J Biomed Health, vol.22, no.5, 2018, pp. 1589– 1604.
- [5] G S Birkhead, M Klompas, and N. R. Shah, "Uses of electronic health records for public health surveillance to advance public health," Annu. Rev. Publ. Health, vol.36, no.1, 2015, pp. 345–359.
- [6] F G Li, Y N Han, and C H Jin, "Cost-effective and anonymous access control for wireless body area networks," IEEE Syst. J., vol.12, no.1, 2018, pp. 747–758.
- [7] M M Hassan, K Lin, X Yue, and J Wan, "A multimedia healthcare data sharing approach through cloud-based body area network," Future Gener. Comp. Sy., vol.66, 2017, pp. 48–58.
- [8] J J P C Rodrigues, I D L Torre, and G Fernández, and M L Coronado, "Analysis of the security and privacy requirements of cloud-based electronic health records systems," J Med. Internet Res., vol.15, no.8, 2013, pp. 418–426.
- [9] M Preethi and R Balakrishnan, "Cloud enabled patient-centric EHR management system," 2014 IEEE International Conference on Advanced Communications, Control and Computing Technologies, Ramanathapuram, India, 2014, pp.1678-1680.

- [10] F Xhafa, J L Feng, Y H Zhang, X F Chen, and J Li, "Privacy-aware attribute-based PHR sharing with user accountability in cloud computing," J Supercomput., vol.71, no.5, 2015, pp. 1607–1619.
- [11] N Leavitt, "Is cloud computing really ready for prime time?," Computer, vol.42, no.1, 2009, pp. 15–20.
- [12] N Sultan, "Making use of cloud computing for healthcare provision: opportunities and challenges," International Journal of Information Management, vol.34, no.2, 2014, pp. 177–184.
- [13] S Nakamoto. "Bitcoin: a peer-to-peer electronic cash system," 2008. [Online] Available: http://bitcoin.org/bitcoin.pdf.
- [14] C Lin, D B He, X Y Huang, K R Choo, and A V Vasilakos, "BSeIn: A blockchain-based secure mutual authentication with fine-grained access control system for industry 4.0," J Netw. Comput. Appl., vol.116, no.15, 2018, pp. 42–52.
- [15] M Mettler, "Blockchain technology in healthcare: the revolution starts here," 2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom), Munich, Germany, 2016, pp.1-3.
- [16] A Azaria, A Ekblaw, T Vieira, and A Lippman, "MedRec: using blockchain for medical data access and permission management," 2016 2nd International Conference on Open and Big Data (OBD), Vienna, Austria, 2016, pp.25-30.
- [17] X Yue, H J Wang, D W Jin, M Q Li, and W Jiang, "Healthcare data gateways: found healthcare intelligence on blockchain with novel privacy risk control," J Med. Syst., vol.40, no.218, 2016, pp. 1–8.
- [18] G Zyskind, O Nathan, and A Pentland, "Decentralizing privacy: using blockchain to protect personal data," 2015 IEEE Security and Privacy Workshops, San Jose, USA, 2015, pp.180-184.
- [19] A Ekblaw, A Azaria, J D Halamka, and A Lippman, "A case study for blockchain in Healthcare: "MedRec" prototype for electronic health records and medical research data," the 2016 IEEE of International Conference on Open and Big Data, Iscataway, USA, 2016, pp.25-30.
- [20] T F Xue, Q C Fu, C Wang, and X Y Wang, "A medical data sharing model via blockchain," Acta Automat. Sinica, vol.43, no.9, 2017, pp.1555–1562.
- [21] Q Xia, E B Sifah, K O Asamoah, J Gao, X J Du, and M Guizani, "MeDShare: trust-less medical data sharing among cloud service providers via blockchain," IEEE Access, vol.5, 2017, pp.14757-14767.
- [22] Q Xia, E B Sifah, A Smahi, S Amofa, and X S Zhang, "BBDS:Blockchain-based data sharing for electronic medical records in cloud environments," Information, vol.8, no.44, 2017, pp.1-16.
- [23] G Yang and C L Li, "A design of blockchain-based architecture for the security of electronic health record (EHR) systems," 2018 IEEE International Conference on Cloud Computing Technology and Science, Nicosia, Cyprus, 2018, pp.261-265.
- [24] H Y Li, L H Zhu, M Shen, F Gao, X L Tao, and S Liu, "Blockchain-based data preservation system for medical data," J Med. Syst., vol.42, 2018, pp.141.
- [25] A Q Zhang and X D Lin, "Towards secure and privacy-preserving data sharing in e-health systems via consortium blockchain," J Med. Syst., vol.42, 2018, pp.140.
- [26] Z B Zheng, S A Xie, H N Dai, X P Chen, and H M Wang, "An overview of blockchain technology: architecture, consensus, and future trends," 2017 IEEE International Congress on Big Data, Honolulu, USA, 2017, pp.557-564
- [27] I Bentov, C Lee, A Mizrahi, and M Rosenfeld, "Proof of activity: extending bitcoin's proof of work via proof of stake [extended abstract]y," Acm Sigmetrics Performance Evaluation Review, vol.42, no.3, 2014, pp.34-37.
- [28] Y Yuan and F Y Wang, "Blockchain: the state of the art and future trends," Acta Automat. Sinica, vol.42, no.4, 2016, pp.481-494.
- [29] M Blaze, G Bleumer, and M Strauss, "Divertible protocols and atomic proxy cryptography," International Conference on the Theory and Applications of Cryptographic Techniques, Espoo, Finland, 1998, pp.127-144.
- [30] R Canetti and S Hohemberger, "Chosen-ciphertext secure proxy reencryption," 14th ACM conference on Computer and communications security, Alexandria, USA, 2007, pp.185-194.
- [31] L F Guo and B Lu, "Efficient proxy re-encryption with keyword search scheme," Journal of Computer Research and Development, vol.51, no.6, 2014, pp.1221-1228.
- [32] T Okamoto, "Cryptography based on bilinear maps," International Symposium on Applied Algebra, Algebraic Algorithms, and Error-Correcting Codes, Las Vegas, USA, 2006, pp.35-50.
- [33] H Xiong and Z G Qin, "Revocable and scalable certificateless remote authentication protocal with anonymity for wireless body area networks," IEEE T. Inf. Foren. Sec., vol.10, no.7, 2015, pp.1442-1455.

[34] J Daemen and V Rijmen, "The design of rijndael: AES-the advanced encryption standard," Springer, Berlin.



XIAOGUANG LIU is now a lecturer in the School of Computer Science and Technology, Southwest Minzu University, Chengdu, P.R. China. He received his Ph.D. degree in mathematics from Shaanxi normal university, Xi'an, P.R. China in 2014. From 2017 to now, he is a postdoctoral fellow in the School of Computer Science and Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu, P.R. China. His recent research interests include cryp-

tography, network security, and optimization computation.



ZIQING WANG is now a master student in the School of Computer Science and Technology of the University of Electronic Science and Technology of China (UESTC), Chengdu, P.R. China. He received his B.S. degree in Information Security from UESTC in 2019. His current research interests include blockchain and identity authentication.



CHUNHUA JIN is now a lecturer in the Faculty of Computer and Software Engineering, Huaiyin Institute of Technology, Huai'an, P.R. China. She received her Ph.D. degree in Information Security from University of Electronic Science and Technology of China (UESTC), Chengdu, P.R. China in 2016. Her recent research interests include cryptography and network security.



FAGEN LI is now a professor in the School of Computer Science and Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu, P.R. China. He received his Ph.D. degree in Cryptography from Xidian University, Xi'an, P.R. China in 2007. From 2008 to 2009, he was a postdoctoral fellow in Future University-Hakodate, Hokkaido, Japan, which is supported by the Japan Society for the Promotion of Science (JSPS). He worked as a research fellow

in the Institute of Mathematics for Industry, Kyushu University, Fukuoka, Japan from 2010 to 2012. His recent research interests include cryptography and network security. He has published more than 80 papers in international journals and conferences. He is a member of the IEEE.



GAOPING LI is now a professor in the School of Computer Science and Technology, Southwest Minzu University, Chengdu, P.R. China. He received his M.S. degree in application mathematics from Chongqing University, Chongqing, P.R. China in 2005. His recent research interests include fractal theory and its applications in image processing, and Compressed sensing.