BDSS: Blockchain-based Data Sharing Scheme With Fine-grained Access Control And Permission Revocation In Medical Environment

Lejun Zhang[1,2,3*], Yanfei Zou[1,4], Muhammad Hassam. Yousuf[1], Weizheng Wang[5], Zilong Jin[6], Yansen Su[7], Kim Seokhoon[8]

[1] College of Information Engineering, Yangzhou University
Yangzhou, 225127, China
[e-mail: zhanglejun@yzu.edu.cn, MZ120190702@yzu.edu.cn, hassam.yousaf02@gmail.com]
[2] Research and Development Center for E-Learning , Ministry of Education
Beijing, 100039, China
[e-mail: zhanglejun@yzu.edu.cn]
[3] Cyberspace Institute Advanced Technology, Guangzhou University
Guangzhou, 510006, China
[e-mail: zhanglejun@yzu.edu.cn]
[4] College of Mechatronics and Information, Wuxi Open University
Wuxi, 214001, China
[e-mail: MZ120190702@yzu.edu.cn]
[5] Computer Science Department, City University of Hong Kong
Hong Kong
[e-mail: m5232117@u-aizu.ac.jp]
[6] School of Computer and Software, Nanjing University of Information Science and Technology
Nanjing, 21004, China
[e-mail: zljin@nuist.edu.cn]
[7] Key Laboratory of Intelligent Computing and Signal Processing of Ministry of Education, School of Computer Science and Technology, Anhui University
Hefei, 230601, China
[e-mail: suyansen@ahu.edu.cn]
[8] Dept. of Computer Software Engineering, Soonchunhyang University
Asan, Korea
[e-mail: seokhoon@sch.ac.kr]
*Corresponding author: Lejun Zhang

Received November 29, 2021; revised April 25, 2022; accepted May 12, 2022; published May 31, 2022

Abstract

Due to the increasing need for data sharing in the age of big data, how to achieve data access control and implement user permission revocation in the blockchain environment becomes an urgent problem. To solve the above problems, we propose a novel blockchain-based data sharing scheme (BDSS) with fine-grained access control and permission revocation in this paper, which regards the medical environment as the application scenario. In this scheme, we separate the public part and private part of the electronic medical record (EMR). Then, we use symmetric searchable encryption (SSE) technology to encrypt these two parts separately, and

http://doi.org/10.3837/tiis.2022.05.012 **ISSN : 1976-7277**

use attribute-based encryption (ABE) technology to encrypt symmetric keys which used in SSE technology separately. This guarantees better fine-grained access control and makes patients to share data at ease. In addition, we design a mechanism for EMR permission grant and revocation so that hospital can verify attribute set to determine whether to grant and revoke access permission through blockchain, so it is no longer necessary for ciphertext re-encryption and key update. Finally, security analysis, security proof and performance evaluation demonstrate that the proposed scheme is safe and effective in practical applications.

Keywords: data sharing; blockchain; access control; permission revocation; medical environment.

1. Introduction

With the development of the Internet, the fourth industrial revolution comes. People have entered the era of using information technology to promote industrial transformation [1-2]. But how to store huge amounts of data effectively becomes an urgent problem to be solved. The emergence of cloud servers solves the storage problem of local devices while reducing costs and improving stability [3-4]. Therefore, cloud servers have become one of the indispensable applications for enterprises and individuals in work or study. In medical fields, with the continuous increase of medical data, traditional paper medical records can no longer meet people's needs. Therefore, EMR, an emerging storage model, attracts people's attention gradually. It has data integrity and low interaction costs so that it can provide supports in the fields of telemedicine, disease treatment, and research of new drugs in the medical field. EMR contains the personal data of patients. Once illegally leaked, it will bring huge losses to the spirit and reputation of the patient. To protect the safety of EMR, patients should encrypt and upload it to the cloud server, but encryption also means that the keyword search technology based on plaintext cannot be used.

The proposal of SSE technology allows people to search encrypted data without revealing document content [5-7]. The development of SSE technology can be traced back to 2000. Song et al. [8] first proposed a symmetric searchable encryption (SSE) scheme based on the symmetric encryption algorithm, but it takes a lot of expenses to restrict access to certain information.

ABE technology proposed by Sahai and Waters [9] is an emerging encryption technology based on identity-based encryption (IBE) technology, which can realize one-to-many encrypted communication and fine-grained access control to data. Therefore, researchers are devoted to combining the cloud server, SSE technology, and ABE technology to achieve searchable encryption and access control at the same time, but this method cannot control the access to the public and private parts of the same data, and it is difficult to revoke access permission.

Subsequently, blockchain proposed by Satoshi Nakamoto [10] is known by people gradually and applied in many fields later. As the underlying technology of Bitcoin, blockchain uses chained data structure to store data, consensus algorithm to upload data, cryptographic principles to ensure the security of data storage, and smart contract to program [11]. Therefore, it is decentralized, secure, and tamper-proof. The characteristics of blockchain can solve the problem of data tampering and collusion attack. However, the tamper-proof feature of the blockchain guarantees the integrity of the data, but it also leads to the permission revocation difficulty.

Therefore, a blockchain-based data sharing system applied in the medical environment is proposed in this paper. The main contributions of this paper are summarized as follows:

(1) This paper proposes a BDSS system model, which combines the traditional cloud server with the distributed blockchain. It not only solves security problems in the cloud server but also relieves storage pressure on the blockchain.

(2) Based on SSE and ABE technology, our scheme encrypts the public part and the private part of EMR with different access policies respectively, thereby enabling patients to control the sharing of different parts of EMR more precisely.

(3) By updating the latest attribute sets in the private blockchain, this paper verifies the identity of the patient and decide whether to grant or revoke his permission. This makes it impossible for adversary to use false attribute set to cheat hospital and obtain the data he wants.

The rest of this paper is organized as follows. Section 2 discusses the related works in our scheme. Section 3 shows the preliminaries of critical technologies in this paper. Section 4 introduces the system design of our scheme. Section 5 describes system implementation. Section 6 performs security analysis, security proof and performance analysis. Section 7 is the conclusion of this scheme.

2. Related Works

Great achievements have been made in the application of blockchain in the medical field. Azaria et al. [12] proposed a blockchain-based electronic medical record system that accesses medical information across providers and treatment sites. Fan et al. [13] designed an efficient and safe medical record sharing system on the blockchain. The system allows effective medical record access and retrieval and uses the ring signature algorithm and zero-knowledge proof technology to enhance data anonymity. Ji et al. [14] proposed a blockchain-based telemedicine information system that realizes multi-level privacy protection location sharing and ensures the retrievability of the complete location.

Since blockchain is decentralized and tamper-proof, the combination of blockchain and cloud becomes a hot topic. Many scholars successively put forward a large number of provable schemes with special properties in this field [15-16]. Tang et al. [15] designed a middleware system that secures cloud storage services using a minimally trusted blockchain. It hardens the cloud-storage security against forking attacks. Xia et al. [16] designed a data-sharing model between cloud service providers using the blockchain. The design employs the use of smart contracts and an access control mechanism to effectively trace the behavior of the data as well as revoke access to violated rules and permissions on data. However, these schemes do not take the problem of ciphertext query in the cloud-blockchain environment into account.

To solve this kind of problem, researchers discover the advantages of searchable encryption technology and add it into the cloud-blockchain environment. This technology implements keyword search on encrypted data and obtains the interested target data [17]. Under the premise of ensuring data security, it can make people search ciphertext more easily. Liu et al.

[18] design an innovative decentralized public key searchable encryption scheme based on a three-layer blockchain network that uncovers illegal and criminal transactions and achieves crime traceability. Cai et al. [19] utilize searchable encryption techniques and smart contract in blockchain to preserves encrypted search capability and enforce ecosystem healthiness. Yang et al. [20] propose a multi-keyword searchable encryption scheme based on blockchain which locates encrypted files precisely and returns the desired files. It also ensures that users can receive accurate search results without any third-party verification. Chen et al. [21] propose a blockchain-based searchable scheme for electronic health records (EHRs). This scheme constructs an index for EHRs through complex logic expressions and stored in the blockchain so that the data user can utilize the expressions to search the index. The above schemes mainly focus on query optimization, but it is also important to ensure the data owner's data access control rights while facilitating the query of the data requester.

Therefore, researchers propose searchable attribute encryption which combines searchable encryption and attribute-based encryption (ABE) to ensure data security while performing fine-grained access control. Feng et al. [22] propose a blockchain data privacy protection control scheme based on searchable attribute encryption, which solves the privacy exposure problem in traditional blockchain transactions. The attribute encryption combined with linear secret sharing performs fine-grained access control on transaction ciphertext in the blockchain. However, most existing attribute-based searchable encryption schemes are inefficient and not suitable for Internet of things devices because of the large amounts of attributes and keys. To solve the critical problems, Niu et al. [23] propose a key aggregation searchable encryption scheme based on blockchain with auxiliary input, which achieves secure data sharing on encrypted data.

Till now, the searchable attribute encryption scheme does not consider that the public and private parts of the data can be separately encrypted and controlled. In addition, the permission revocation problems in the blockchain still exist. Therefore, this paper is dedicated to building a blockchain-based data sharing scheme with fine-grained access control and permission revocation.

3. Preliminaries

3.1 Symmetric Searchable Encryption

Traditional SSE algorithm can be described as a quintuple [24]:

1) $SSE\_KeyGen(\lambda) \rightarrow (K)$: Executed by the hospital. Input security parameters λ, output symmetric key K;

2) $SSE\_Encrypt(K, A) \rightarrow (B)$: Executed by the hospital. Input symmetric key K and plaintext set $A = \{a_1, a_2, \cdots, a_n\}$, output ciphertext set $B = (b_1, b_2, \cdots, b_n)$.

3) $SearchToken(K, w) \rightarrow T_w$: Executed by the hospital. Input symmetric key K and keyword w, output search token T_w;

4) $Search(I, T_w, B) \rightarrow B(w)$: Executed by the cloud server. Input index I, trapdoor T_w and encrypt file set B, output specific encrypted file set $B(w)$;

5) $SSE\_Decrypt(K, B_i) \rightarrow A_i$: Executed by the hospital. Input symmetric key K and encrypted file B_i, output corresponding plaintext A_i.

3.2 Attribute-Based Encryption

Traditional ABE algorithm can be described as a quaternion [25]:

1) $ABE\_Setup(\alpha, \beta, g) \rightarrow (MSK, PK)$: Executed by the hospital. Input random exponents

$\alpha, \beta \in Z_p$ and the generator of the bilinear group g, output system master key MSK and system public key PK;

2) $ABE\_KeyGen(MSK, PK, \omega) \rightarrow (SK_\omega)$: Executed by the hospital. Input system master key MSK, system public key PK and attribute set ω, output attribute private key SK_ω;

3) $ABE\_Encrypt(M, AT, MSK) \rightarrow (C_M)$: Executed by the hospital. Input message M, tree access structure AT and system master key MSK, output encrypted message C_M;

4) $ABE\_Decrypt(C_M, \omega, AT, PK, SK_\omega) \rightarrow (M)$: Executed by the hospital. Input encrypted message C_M, attribute set ω, tree access structure AT, system public key PK, and attribute private key SK_ω, output message M.

4. BDSS System Design: Take the Medical Environment as the Application Scenario

4.1 Design Goals

In order to facilitate users to store and share data, we design a BDSS system based on cloud server and blockchain. We take the medical environment as the application scenario which makes this scheme more practical.

First of all, since the patient wants to guarantee the security and searchability of EMR at the same time when they share data with others, our scheme generates I and T_w with SSE technology. Thus, patient can share data more safely, and user can search data more efficiently than before.

Next, EMR includes D_{pu} and D_{pr}. D_{pu} refers to the data that the patient is willing to share with people who may research the treatment. D_{pr} refers to the data that the data owner only wants to share with specific people who may be doctors or relatives. For example, D_{pu} may include data such as diagnosis results and prescription. D_{pr} may include data such as personal information, insurance information, and medical history. In **Fig. 1**, we summarize the D_{pu} and D_{pr} of EMR in most cases. The green parts represent the D_{pu} of EMR, and the red parts represent the D_{pr} of EMR. In order to solve this problem, our scheme uses ABE technology to encrypt K_1 and K_2 which encrypt D_{pu} and D_{pr} with AT_1 and AT_2 respectively, so that patients can guarantee the fine-grained sharing of their EMR.

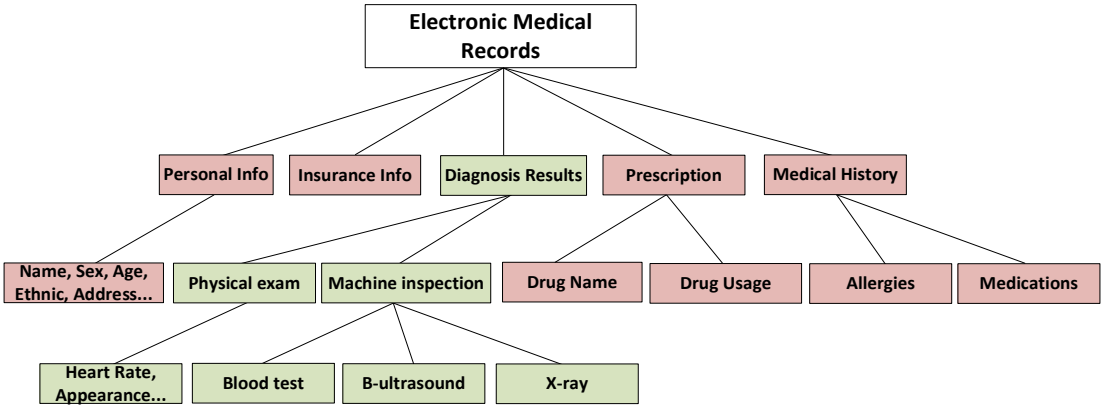

Fig. 1. The public and private parts of Electronic Medical Records

Besides, since the storage of blockchain is limited, hospital stores large-scale ciphertext of EMR in the cloud server. In order to prevent malicious users from tampering with crucial data, hospital records access event in the blockchain. Not only solves security problems in the cloud storage server but also solves the problem of limited storage space on the blockchain.

Finally, blockchain is a double-edged sword. It prevents data tampering and makes it difficult for patients to revoke the access permission. Our scheme uses the smart contract to verify the attribute set ω, thus even if the user's ω has changed after obtaining T_w, the smart contract can check the changes of ω and revoke the access permission.

4.2 BDSS System Model

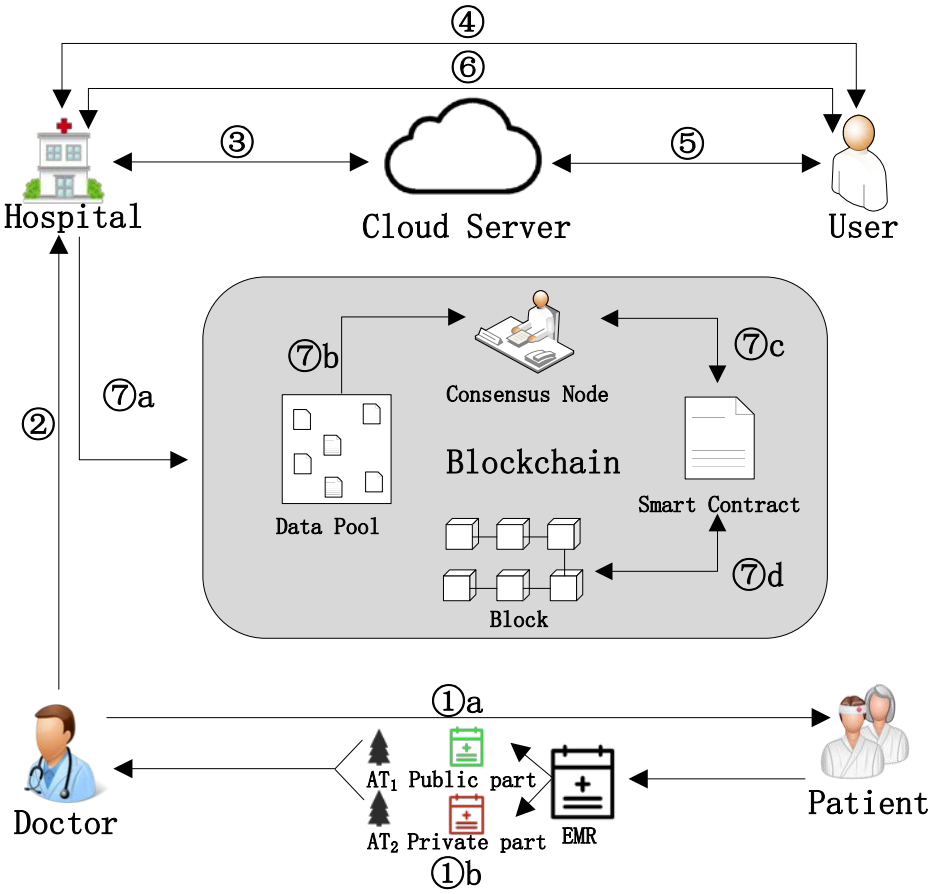

Fig. 2. System model

In our scheme, the hospital is the trusted authorization center, which is responsible for the key generation and encryption of SSE and ABE. Medical record is the historical file of patient treatment, while electronic medical record is digital medical record. Medical record is the historical file of patient treatment, while EMR is digital medical record. As EHR records the patient's treatment history, patients and their families have the right to view the EHR and to determine the level of granular sharing of EMR. The main process of system model is shown in **Fig. 2**. Next, we describe these procedures in detail.

① After treating the patient, doctor generates EMR. Patient divides his own EMR into pubic part D_{pu} and private part D_{pr}. At the same time, patient sends AT_1 and AT_2 that can control the permissions of D_{pu} and D_{pr} of EMR to the doctor;

② Doctor collects and uploads AT_1 and AT_2 to the hospital;

③ Hospital encrypts D_{pu} and D_{pr} of EMR with SSE technology separately, and uploads C_{pu}, C_{pr} and I to cloud server. Then, hospital encrypts K_1 and K_2 that encrypt D_{pu} and D_{pr} with ABE technology, and generates C_{K_1} and C_{K_2} so that the access of EMR can be controlled.

④ When user wants to search the data, the request with w for searching EMR is sent to the hospital. Hospital generates the corresponding T_w, and returns it to the user.

⑤ User submits the T_w to the cloud server. The cloud server matches the corresponding index I according to T_w and returns the C_{pu}, C_{pr} and ID_C to the user.

⑥ User submits ω, C_{pu}, C_{pr} to hospital. Hospital generates SK_ω by ω. If the user's attribute satisfies AT_1, then decrypt successfully generate K_1; if the user property meets AT_2, then decrypt successfully generates K_2. Finally, the hospital decrypts C_{pu} and C_{pr} through K_1 and K_2, and returnsD_{pu} and $D_{pr}$to the user. In this way, our scheme implements fine-grained access control while protecting the symmetric key.

⑦ The hospital puts the ciphertext hash value $H(C_{pu})$ and $H(C_{pr})$, the public key of the hospital, the user's public key, and the signature of hospital to the data pool. The consensus node validates the data in the data pool and records the access event via the smart contract. Users can verify whether the ciphertext in the cloud server is complete and correct through the blockchain, and patients can also know the access of their own data through the blockchain.

At this point, the EMR sharing process is complete.

4.3 Notations

Many different notations are used in our scheme. For the convenience of reading, these notations and their descriptions are summarized in **Table 1**.

Table 1. Notations

Notations	Description
D_{pu}, D_{pr}	The public part and private part of EMR set
C_{pu}, C_{pr}	The ciphertext of C_{pu} and C_{pr}
w	Keyword
I	Index
K_1	The symmetric key that is used to encrypt the public part of D_{pu}
K_2	The symmetric key that is used to encrypt the private part of D_{pr}
K_3	The symmetric key that is used to encrypt search token and index
C_{K_1}, C_{K_2}	The ciphertext of K_1 and K_2
MSK	System master key
PK	System public key
AT_1, AT_2	Tree access structures for encrypting K_1 and K_2
ω	Attribute set
SK_ω	The attribute private key of the user with attribute set ω
T_w	Search token generated by the keyword w
F	HMAC-SHA256

5. BDSS System Implementation

The BDSS system is divided into two stages: EMR storage stage and EMR sharing stage. The detailed implementations of each stage are as follows.

5.1 EMR Storage Stage

At this stage, hospital runs SSE_KeyGen and ABE_Setup to initialize key parameters of SSE technology and ABE technology.

Hospital runs $SSE\_KeyGen(\lambda)$ to generate K_1, K_2 and K_3. K_1 is used to encrypt D_{pu}. K_2 is used to encrypt D_{pr}. K_3 is used to generate I and T_w.

Hospital runs $ABE\_Setup(\alpha, \beta, g)$ to generate MSK and PK. MSK is used to encrypt C_{K_1} and C_{K_2}. PK is used to decrypt C_{K_1} and C_{K_2}.

5.1.1 EMR Encryption and Index Generation

In real life, EMR includes public part and public part. Patient wants to share them with different people. However, traditional schemes cannot meet this need. In order to solve this problem, SSE technology is adopted in our scheme to encrypt and generate indexes for patients' private data and public data respectively, as shown in **Algorithm 1**.

Algorithm 1 EMR Encryption

Input: D_{pu}, D_{pr}
Output: C_{pu}, C_{pr}, I

Extract keyword set W
Set $I = \emptyset$
for each w in W **do**
  $E_w = F(K_3, w)$
  Set the public part of EMR which contain keyword w as D_{pu_w}
  Set the private part of EMR which contain keyword w as D_{pr_w}
  Run $SSE\_Encrypt(K_1, D_{pu_w})$ to generate C_{pu_w}
  Run $SSE\_Encrypt(K_2, D_{pr_w})$ to generate C_{pr_w}
  $ID_{C_w} = \{ID_{C_{pu_w}}, ID_{C_{pr_w}}\}$
  $I_w = (E_w, ID_{C_w})$
end for
return C_{pu}, C_{pr}, I

In this algorithm, the hospital extracts the keywords in the plaintext to form the keyword set W. For each keyword w in W, the hospital collated the public part set D_{pu_w} and the private part set D_{pr_w} related to the keyword, and generated ciphertext C_{pu_w} and C_{pr_w} using SSE technology.

5.1.2 Symmetric Keys Encryption

Note that the most important step to realizing ABE technology is the access structure. An access structure defines a combination of attributes with decryption authority. Only ω that meets the access structure can recover the correct SK_ω to decrypt the ciphertext. There are three access structures. They are threshold access structure [26], tree access structure (AT) [27], and linear secret sharing matrix structure [28] respectively. Because tree access structure represents a more flexible access control strategy which is more suitable for the cloud environment, we choose it as the access structure of our scheme. AT consists of leaf nodes and

non-leaf nodes. Each leaf node is described by an attribute value. Each non-leaf node represents a threshold gate, such as 'AND' gate or 'OR' gate. When the root node in AT can meet its threshold, it means the ω satisfies the AT.

In this scheme, symmetric keys of public and private parts are encrypted with different access control trees so that patients can control the fine-grained sharing of public and private parts. **Fig. 3** shows the example of AT_1. **Fig. 4** shows the example of AT_2.

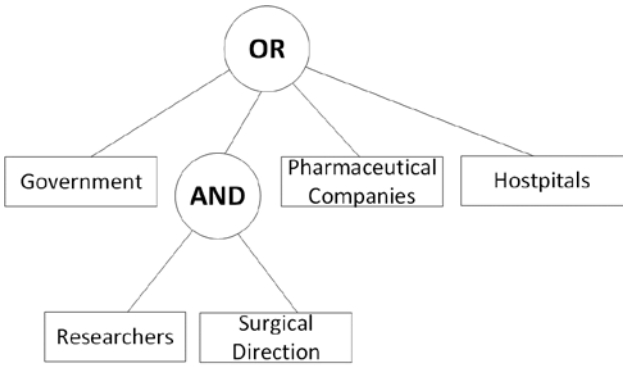

Fig. 3. The example of AT_1

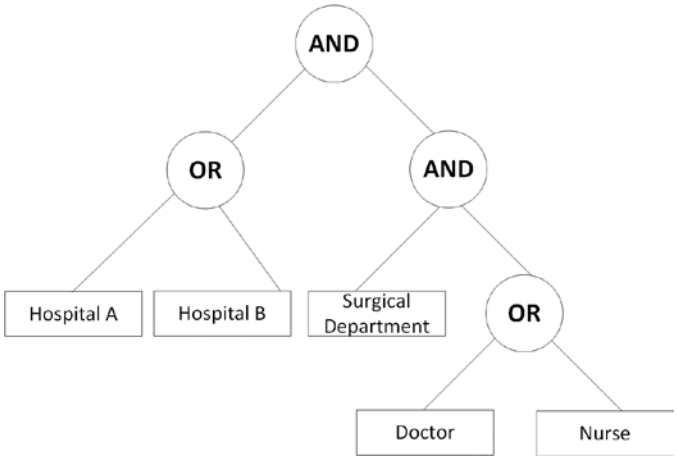

Fig. 4. The example of AT_2

In the above figures, we can find that AT_1 has broader requirements, while AT_2 has stricter requirements relatively. Through them, patients can control the access D_{pu} and D_{pr} more precisely.

The symmetric keys are very important for the security of EMR. Even if the system uses a highly secure encryption scheme, the adversary can attack it successfully when keys are leaked. Therefore, hospital executes $ABE\_Encrypt(K_1, AT_1, MSK)$ and $ABE\_Encrypt(K_2, AT_2, MSK)$ to generate C_{K_1} and C_{K_2}, and stores them in their own database.

5.2 EMR Sharing Stage

5.2.1 Search Token Generation

To access the patient's EMR, user needs to send an access request containing w to the hospital. After verification, patient runs $SSE\_Trapdoor(K_3, w)$ to generate T_w. The generation of T_w

is shown in (1) and (2).

$$E_w = F(w, K_3) \quad (1)$$

$$T_w = (E_w) \quad (2)$$

Then, hospital sends T_w to the authenticated user through the secure channel.

5.2.2 EMR Investigation

User sends T_w to cloud server. Cloud server verifies whether the E_w in T_w is consistent with E_w in I. If they are consistent, cloud server runs $SSE\_Search(I, T_w)$ to generates a file identifier collection $D(w)$, then finds ciphertext result C_{pu_w} and C_{pr_w} according to $D(w)$, and sends it to the user.

5.2.3 EMR Permission grant and revocation

The tamper-proof feature of the blockchain guarantees the integrity of the data, but it also leads to the permission revocation difficulty. The attributes of users will change over time, so their access permissions should be changed accordingly. For example, when user was a doctor in a certain hospital, he achieved T_w from the patients. But he resigns and becomes a personal physician now. He could access the medical data of that hospital before, but now he cannot access these data by T_w. Therefore, when ω of user has changed and not met AT any more, the system must revoke his access permission. Otherwise, private data may be leaked. Therefore, it is necessary to design a mechanism for EMR permission revocation.

Our solution uses the method in **Fig. 5** to improve the incompatibility between the tamper-proof feature of the blockchain and the permission revocation of ABE technology.

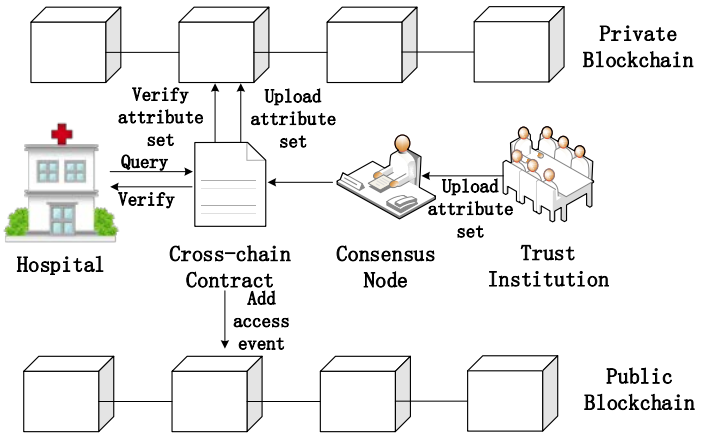

Fig. 5. EMR Permission grant and revocation mechanism

This mechanism combines private blockchain and public blockchain to achieve EMR permission revocation.

Step 1: Trust authority submits user's latest ω and public key to the consensus node in the private blockchain. After consensus, the consensus nodes upload them to the private blockchain through cross-chain smart contracts. Trust authority is regulated by the government who tracks user and updates ω in time. The private blockchain protects the user's attribute privacy from being seen by malicious external users.

Step 2: When the user requests data, he submits ω, C_{pu} and C_{pr} to the hospital. Hospital

uses cross-chain smart contract to verify that whether the ω sent by users is consistent with the latest ω' in the private blockchain. If consistent, the ω submitted by the user is the latest, the hospital runs $ABE\_KeyGen(MSK, PK, \omega)$ to generate SK_ω for the user. If it is inconsistent, it means that the ω sent by the user is fake, then the hospital will refuse to serve it. Then, the hospital will continue to verify whether ω meets the access control tree. If ω meets AT_1, hospital runs $ABE\_Decrypt(C_{K_1}, \omega, AT_1, PK, SK_\omega)$ to achieve K_1, and runs $SSE\_Decrypt(K_1, C_{pu_w})$ to achieve the plaintext of public part of EMR D_{pu_w}. If ω meets AT_2, hospital runs $ABE\_Decrypt(C_{K_2}, \omega, AT_2, PK, SK_\omega)$ to achieve K_2, and runs $SSE\_Decrypt(K_2, C_{pr_w})$ to achieve the plaintext of private part of EMR D_{pr_w}. EMR Access Permissions Grant and Undo Algorithm is shown in **Algorithm 2**.

Algorithm 2 EMR access permissions grant and revocation

Input: user attribute information IS, attribute submitted by users ω, the ciphertext of public and private part of EMR C_{pu_w} and C_{pr_w}, use's public key K_U

Output: the plaintext of public and private part of EMR D_{pu_w} and D_{pr_w}

Cross chain smart contract records user's new attribute information of IS (K_U, ω', ct) in the blockchain //ct represents the current update time
User submits $\omega, C_{pu_w}, C_{pr_w}$ and K_U to the hospital
Hospital searches the latest ω' according to K_U and ct
if $\omega = \omega'$ **then**
    if ω meet AT_1 **then** //AT_1 represents tree access structure of the public part of EMR
        Hospital decrypts C_{pu_w} for the user
    end if
    if ω meet AT_2 **then** //AT_2 represents tree access structure of the private part of EMR
        Hospital decrypts C_{pr_w} for the user
    else
        Hospital refuses to provide the service
    end if
else
    Hospital refuses to provide the service
end if
return D_{pu_w}, D_{pr_w}

Step 3: In order to facilitate the user to verify the integrity of the EMR in the cloud server, the hospital calculates the hash value of C_{pu} and C_{pr}, which is represented as $H(C_{pu})$ and $H(C_{pr})$. The hash value is uploaded to the blockchain's data pool along with the hospital's public key, the user's public key, and the hospital's signature on the request. To store this data in the blockchain, consensus nodes take data from the data pool for processing. When consensus node reaches a consensus with other nodes, the data can be packaged into a block through smart contracts and stored in the blockchain.

6. Security and Performance Analysis

6.1 Security Analysis

Security is critical in EMR sharing systems. This part conducts a security analysis on the proposed scheme from five aspects: data segmentation confidentiality, attribute verification and privacy protection and collusion resistance.

(1) Data segmentation confidentiality

In the access control phase, patient separates the private part and the public part from the whole EMR and constructs the tree access structure for each part. Therefore, if the adversary's attribute set meets AT_1 but not meet AT_2, he cannot access the plaintext of the private part of the EMR whose plaintext of the public part can be accessed by the smart contract. Thus, data segmentation confidentiality is fully achieved.

(2) Attribute verification and privacy protection

In our scheme, only if the adversary submits the latest attribute set honestly can he request data from the smart contract. Because the smart contract will find the block which includes the user's latest attribute set by timestamp in the blockhead and verify the consistency of it once the user submits the request. If the attribute set is inconsistent, the smart contract will revoke its permission. Moreover, our scheme stores the attribute sets in the private blockchain which only can be accessed by specific people, so the attribute privacy is also protected. Therefore, the conclusion is valid.

(3) Collusion resistance

In our scheme, the cloud server is "honest but curious", which means that the cloud server will perform its duties but remain curious about the patient's EMR. For example, when user tries to collude with the cloud server to steal EMR, the cloud server cannot find the desired ciphertext by keyword w from the massive ciphertext without T_w, then the collusion cannot be reached. So, our scheme achieves collusion resistance.

6.2 Security Proof

In our scheme, if the adversary wants to tamper with the data in the blockchain, he must attack more than one-third of the normal nodes or all master nodes. In the blockchain network, assuring that the attack probability of the master node and normal node is p_a and p_b, and the number of the master node and normal node is n_a and n_b. The tamper probability of data in the blockchain is as shown in (3).

$$p_{bc} = {p_a}^{n_a} + {p_b}^{\frac{n_b}{3}+1} \tag{3}$$

Once the cloud server is attacked, all the data may be leaked or even tampered with. Therefore, the tamper probability of the cloud server is the same as the attack probability of the primary node as shown in (4).

$$p_{cs} = p_a \tag{4}$$

For the convenience of calculation, we assume that the attack probability of the primary node and normal node is 0.5% and 1% respectively. The tamper probability of index and decryption information in the two methods is shown in **Fig. 6**.

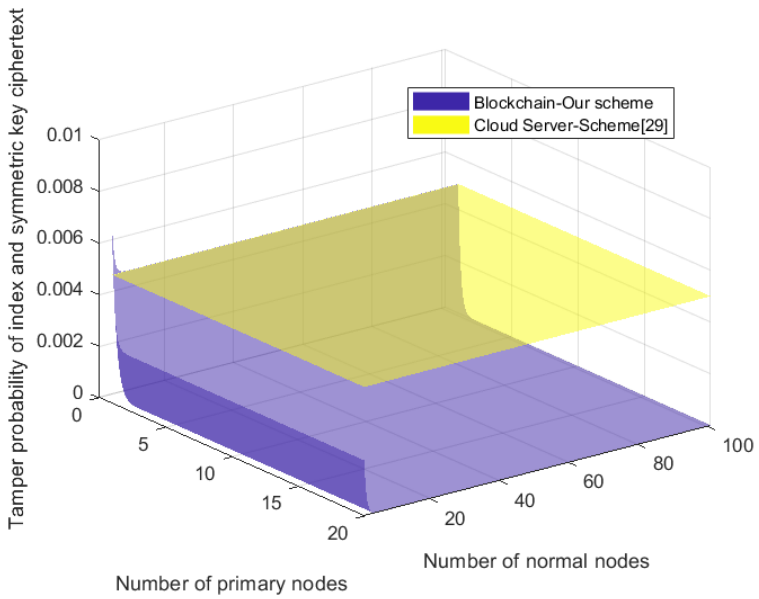

Fig. 6. Tamper probability of index and symmetric key ciphertext

In the above figure, when the number of primary nodes and normal nodes in the blockchain is small, the tamper probability of distributed blockchain is high relatively. The reason is that when the number of nodes in the blockchain is small, there is a high probability that the adversary attacks more than one-third of the ordinary nodes or all the master nodes successfully. When the number of ordinary nodes and master nodes increases gradually, it will be difficult for the adversary to attack more than one-third of the ordinary nodes or all master nodes simply. As a result, the tamper probability of distributed blockchains drops rapidly and approaches zero finally. However, since the tampering probability of scheme [29] is the tampering probability of the primary node, the tamper probability of cloud server will not be affected by the number of master nodes and ordinary nodes. That is, once the cloud server is attacked, the security of the data cannot be guaranteed.

In the real life, blockchain has a large number of normal nodes and primary nodes, so its tamper probability is quite low. Thus, our scheme achieves tamper-proof effectively.

6.3 Comparison of Computation Costs

In this part, we compute the computation costs of the EMR encryption and decryption to reflect the efficiency of fine-grained access control in our scheme.

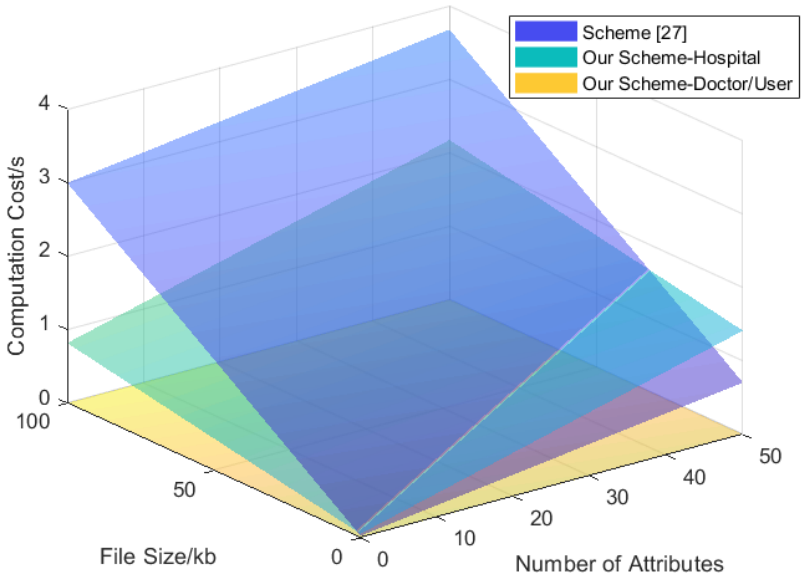

Fig. 7. The computation cost of EMR encryption

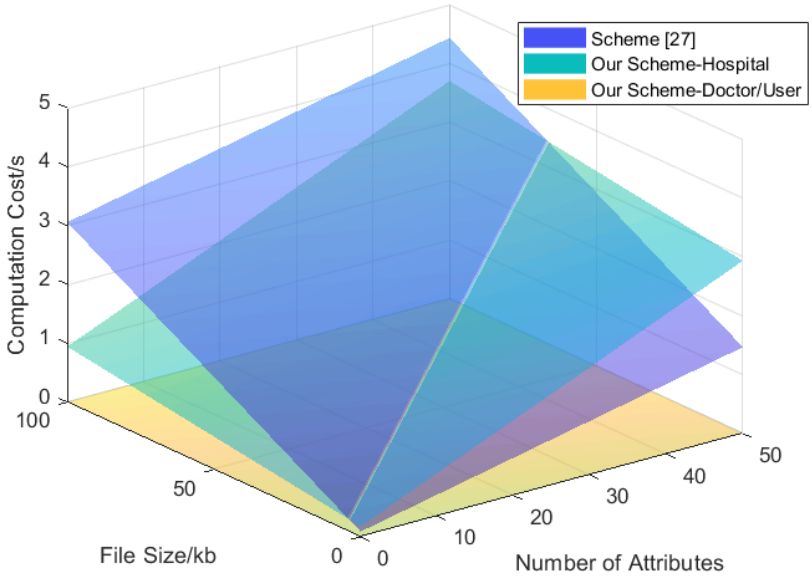

Fig. 8. The computation cost of EMR decryption

Fig. 7 and **Fig. 8** show the computation costs of EMR encryption and decryption of scheme [27] and our scheme, where we can see that the results of both schemes increase linearly with the file size of EMR and the number of attributes contained in tree access structure. It is worth

noting that the three-dimensional plane of the scheme [27] intersects with the plane of our scheme. However, the computation cost of our scheme is lower than that of the scheme [27] in most cases, so it is obvious that our scheme is more efficient under the premise of achieving fine-grained access control.

In addition, the encryption and decryption costs of doctors and users in our scheme are not related to the file size and the number of attributes. This is because our scheme outsources encryption and decryption work to the hospital and smart contract. Doctors and users who do not have large computing power do not need to encrypt or decrypt EMR and symmetric keys, so our scheme is user-friendly.

6.4 Comparison of Permission Revocation

In addition, we compare permission revocation manipulations with other schemes to reflect the flexibility and efficiency of our scheme in **Table 2**.

Table 2. Comparison of permission revocation manipulations

Entities	scheme [30]	Scheme [31]	Scheme [32]	Scheme [26]	Our Scheme
Blockchain	Update ciphertext of public key and transactions	-	Update user revocation list	Update ciphertext of symmetric keys	Update attribute sets
Access structure	-	-	-	Update access structure	-
Data owner	Re-encrypt ciphertext of public key	Generate attribute update-keys and attribute signing key	-	Re-encrypt ciphertext of symmetric keys	-

Researchers have adopted different methods to achieve permission revocation. Scheme [30] re-encrypts the public key and updates the ciphertext of the public key and transaction on the blockchain. Scheme [31] requires user to generate attribute update-keys and attribute signing key, which puts a heavy computation burden on the user. Scheme [32] controls the permission revocation by updating the user revocation list while our scheme controls the permission revocation by updating the attribute set. These two schemes do not need to re-encrypt the ciphertext and key, which ensures the update efficiency. However, our scheme stores the attribute set in the private blockchain and searches information on the public blockchain. This not only makes the search more transparent but also protects the privacy of user's attributes.

Scheme [26] re-encrypts ciphertext of symmetric keys, update access structure and ciphertext of symmetric keys, which involves multiple entities and huge space consumption. The space cost comparison of scheme [26] and our scheme is shown in **Fig. 9**. From this figure, we can see that the space consumption of scheme [26] grows faster than our scheme. With the increase in the number of users whose attribute sets were updated, the gap of space cost between the two schemes has widened.

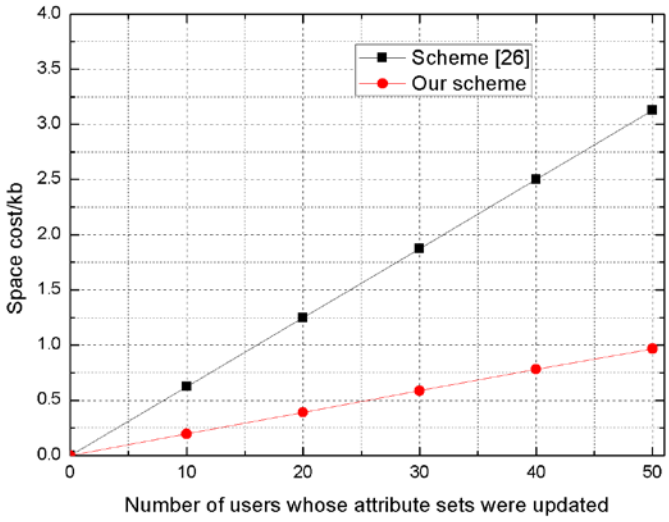

Fig. 9. The space cost of EMR permission revocation

7. Conclusions

This paper proposes a blockchain-based data sharing system, BDSS. Compared with scheme [27], the method of our scheme improves the efficiency of EMR encryption and decryption while achieving further fine-grained access control. In addition, simulation experiment proves that blockchain ensures low data tampering rate as a storage environment. Moreover, the permission grant and revocation mechanism in our scheme updates attribute sets by trust institution and verify attribute sets by smart contract without ciphertext re-encryption and key update, which is better than scheme [26] in time cost and space cost. In the future, we plan to improve our system in terms of search efficiency and search accuracy.

Acknowledgement

The authors would like to thank the reviewers for their detailed reviews and constructive comments, which have helped improve the quality of this paper. This work is sponsored by The National Key Research and Development Program of China No. 2021YFE0102100, the National Natural Science Foundation of China under grant number No. 62172353, Future Network Scientific Research Fund Project No. FNSRFP-2021-YB-48, Science and Technology Program of Yangzhou City No. YZU202003 and Six Talent Peaks Project in Jiangsu Province No. XYDXX-108.

References

[1] J. Qiu, Z. Tian, C. Du, Q. Zuo, S. Su and B. Fang, "A Survey on Access Control in the Age of Internet of Things," *IEEE Internet of Things Journal*, vol. 7, no. 6, pp. 4682-4696, Jun. 2020. Article (CrossRef Link)

[2] M. Shafiq, Z. Tian, A. K. Bashir, X. Du and M. Guizani, "CorrAUC: a Malicious Bot-IoT Traffic Detection Method in IoT Network Using Machine Learning Techniques," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3242-3254, Jun. 2021. Article (CrossRef Link)

[3] L. A. Tawalbeh, R. Mehmood, E. Benkhelifa and H. Song, "Mobile Cloud Computing Model and Big Data Analysis for Healthcare Applications," *IEEE Access*, vol. 4, pp. 6171-6180, Sep. 2016. Article (CrossRef Link)

[4] J. T. Sun and Y. G. Fang, "Cross-domain data sharing in distributed electronic heath record systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 21, no. 6, pp. 754-764, Aug. 2010. Article (CrossRef Link)

[5] C. Bösch, P. Hartel, W. Jonker and A. Peter, "A survey of provably secure searchable encryption," *ACM Computing Surveys*, vol. 47, no. 2, pp. 1-51, Jan. 2015. Article (CrossRef Link)

[6] G. S. Poh, J. J. Chin, W. C. Yau and K. R. Choo, "Searchable symmetric encryption: Designs and challenges," *ACM Computing Surveys*, vol. 50, no. 3, pp. 1-37, May 2018. Article (CrossRef Link)

[7] S. Li, M. Li, H. Xu and X. Zhou, "Searchable Encryption Scheme for Personalized Privacy in IoT-Based Big Data," *Sensors*, vol. 19, no. 5, pp. 1059, Jan. 2019. Article (CrossRef Link)

[8] D. X. Song, D. Wagner and A. Perrig, "Practical Techniques for Searches on Encrypted Data," in *Proc. of 2000 IEEE Symposium on Security and Privacy*, May 2000. Article (CrossRef Link)

[9] A. Sahai and B. Waters, "Fuzzy identity-based encryption," in *Proc. of 24th annual international conference on the theory and applications of cryptographic techniques*, May 2004. Article (CrossRef Link)

[10] S. Nakamoto, "Bitcoin: a peer-to-peer electronic cash system," 2008. [Online]. Available: http://bitcoin.org/bitcoin.pdf

[11] X. L. Yang, Y. Chen and X. H. Chen, "Effective Scheme against 51% Attack on Proof-of-Work Blockchain with History Weighted Information," in *Proc. of 2019 IEEE International Conference on Blockchain*, July 2019. Article (CrossRef Link)

[12] A. Azaria, A. Ekblaw, T. Vieira and A. Lippman, "MedRec: Using Blockchain for Medical Data Access and Permission Management," in *Proc. of 2016 2nd International Conference on Open and Big Data*, August 2016. Article (CrossRef Link)

[13] K. Fan, S. Y. Wang, Y. H. Ren, H. Li and Y. T. Yang, "MedBlock: Efficient and Secure Medical Data Sharing Via Blockchain," *Journal of Medical Systems*, vol. 42, no. 8, pp. 136, Jun. 2018. Article (CrossRef Link)

[14] Y. X. Ji, J. W. Zhang, J. F. Ma, C. Yang and X. Yao, "BMPLS: Blockchain-Based Multi-level Privacy-Preserving Location Sharing Scheme for Telecare Medical Information Systems," *Journal of Medical Systems*, vol. 42, no. 8, pp. 1-13, Jun. 2018. Article (CrossRef Link)

[15] Y. Tang, Q. Zou, J. Chen, K. Li, C. Kamhoua, K. Kwiat and L. Njilla, "ChainFS: Blockchain-Secured Cloud Storage," in *Proc. of 2018 IEEE 11th International Conference on Cloud Computing*, July 2018. Article (CrossRef Link)

[16] Q. Xia, E. B. Sifah, K. O. Asamoah, J. B. Gao, X. J. Du and M. Guizani, "MeDShare: Trust-Less Medical Data Sharing Among Cloud Service Providers via Blockchain," *IEEE Access*, vol. 5, pp. 14757-14767, Jul. 2017. Article (CrossRef Link)

[17] Z. Fu, X. Sun, Q. Liu, L. Zhou and J. Shu, "Achieving Efficient Cloud Search Services: Multi-Keyword Ranked Search over Encrypted Cloud Data Supporting Parallel Computing," *Ieice Transactions on Communications*, vol. 98, no. 1, pp. 190-220, Jan. 2015. Article (CrossRef Link)

[18] C. Liu, Y. Xiao, V. Javangula, Q. Hu, S. Wang and X. Cheng, "NormaChain: A Blockchain-based Normalized Autonomous Transaction Settlement System for IoT-based E-commerce," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4680-4693, Jun. 2019. Article (CrossRef Link)

[19] C. Cai, J. Weng, X. Yuan and C. Wang, "Enabling Reliable Keyword Search in Encrypted Decentralized Storage with Fairness," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 1, pp. 131-144, Feb. 2021. Article (CrossRef Link)
[20] X. Yang, G. Chen, M. Wang, T. Li and C. Wang, "Multi-keyword Certificateless Searchable Public Key Authenticated Encryption Scheme Based on Blockchain," *IEEE Access*, vol. 8, pp. 158765-158777, Sep. 2020. Article (CrossRef Link)
[21] L. Chen, W. K. Lee, C. C. Chang, K.K. R. Choo and N. Zhang, "Blockchain based searchable encryption for electronic health record sharing," *Future Generation Computer Systems*, vol. 95, no. 6, pp. 420-429, Jan. 2019. Article (CrossRef Link)
[22] T. Feng, H. Pei, R. Ma, Y. Tian and X. Feng, "Blockchain Data Privacy Access Control Based on Searchable Attribute Encryption," *Computers, Materials and Continua*, vol. 66, no. 1, pp. 871-890, 2021. Article (CrossRef Link)
[23] J. Niu, X. Li, J. Gao and Y. Han, "Blockchain-Based Anti-Key-Leakage Key Aggregation Searchable Encryption for IoT," *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 1502-1518, 2020. Article (CrossRef Link)
[24] J. W. Li, C. F. Jia, Z. L. Liu and J. Li, "Survey on the searchable encryption," *Journal of Software*, vol. 26, no. 1, pp. 109-128, 2015. Article (CrossRef Link)
[25] S. Wang, D. Zhang and Y. Zhang, "Blockchain-based personal health records sharing scheme with data integrity verifiable," *IEEE Access*, vol. 7, pp. 102887-102901, 2019. Article (CrossRef Link)
[26] L. J. Zhang, M. H. Peng, W. Z. Wang, "Secure and Efficient Data Storage and Sharing Scheme Based on Double Blockchain," *Computers, Materials and Continua*, vol. 66, no. 1, pp. 499-515, 2021. Article (CrossRef Link)
[27] S. F. Niu, L. X. Chen, J. F. Wang and F. Yu, "Electronic Health Record Sharing Scheme with Searchable Attribute-Based Encryption on Blockchain," *IEEE Access*, vol. 8, pp. 7195-7204, Dec. 2019. Article (CrossRef Link)
[28] S. C. Bunker, M. Barasa and A. Ojha, "Linear Equation Based Visual Secret Sharing Scheme," in *Proc. of 2014 IEEE International Advance Computing Conference*, February 2014. Article (CrossRef Link)
[29] Y. Hui, Q. Zheng, J. X. Zhang, H. Deng, F. M. Li and K. Q. Li, "A fine-grained authorized keyword secure search scheme with efficient search permission update in cloud computing," *Journal of Parallel and Distributed Computing*, vol. 135, pp. 56-69, Jan. 2020. Article (CrossRef Link)
[30] G. Yu, X. Zha, X. Wang, W. Ni, K. Yu, P. Yu, J. A. Zhang, R. P. Liu and Y. J. Guo, "Enabling Attribute Revocation for Fine-Grained Access Control in Blockchain-IoT Systems," *IEEE Transactions on Engineering Management*, vol. 67, no. 4, pp. 1213-1230, Nov. 2020. Article (CrossRef Link)
[31] Q. Su, R. Zhang, R. Xue and P. Li, "Revocable Attribute-Based Signature for Blockchain-Based Healthcare System," *IEEE Access*, vol. 8, pp. 127884-127896, Jul. 2020. Article (CrossRef Link)
[32] S. Liu, J. Yu, Y. Xiao, Z. Wan, S. Wang and B. Yan, "BC-SABE: Blockchain-aided Searchable Attribute-based Encryption for Cloud-IoT," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 7851-7867, Sept. 2020. Article (CrossRef Link)

Lejun Zhang received his M.S. degree in computer science and technology in Harbin Institute of Technology and the Ph.D. degrees in computer science and technology at Harbin Engineering University, where he was an professor at Yangzhou University. His research interests include computer network, social network analysis, dynamic network analysis and information security

Yanfei Zou is currently pursuing a M.S. degree in computer technology at Yangzhou University. Her research interests include blockchain, information security.

Muhammad Hassam Yousuf completed his bachelor (BS) degree in Software Engineering form Pakistan. Now doing his masters (MS) degree from Yangzhou University, Yangzhou, Jiangsu China. His research interests include wireless networking optimization and broadcast.

Weizheng Wang received the B.S. degree in software engineering from Yangzhou University, Yangzhou, China, in 2019, the M.S. degrees in computer science and engineering from the University of Aizu, Aizu-Wakamatsu, Japan, in 2021. Now he is currently a Research Associate in University of Aizu and pursuing the Ph.D. degree in computer science at the City University of Hong kong, Hong Kong. His research interests include applied cryptography, blockchain technology and IoT system.

Zilong Jin received the B.E. degree in computer engineering from Harbin University of Science and Technology, China, in 2009, and the M.S. and Ph.D. degrees in computer engineering from Kyung Hee University, Korea, in 2011 and 2016, respectively. He is currently an assistant professor of School of Computer and Software at Nanjing University of Information Science and Technology, China. His research interests include wireless sensor networks, mobile wireless networks, and cognitive radio networks.

Yansen Su received the B.Sc. degree from Tangshan Normal University, Tangshan, China, in 2007, the M.Sc. degree from Shandong University of Science and Technology, Qingdao, China, in 2010, and the Ph.D. degree from Huazhong University of Science and Technology, Wuhan, China, in 2014.She is currently an Associate Professor in the School of Computer Science and Technology, Anhui University, Hefei, China. Her main research interests include complex networks, computational biology, and multi-objective optimization.

Kim Seokhoon received the B.E. and Ph.D. degrees in computer engineering from Kyunghee University, Korea, in 2000 and 2004, respectively. From 2004 to 2006, he was with IPOne, Inc., Seoul, Korea, where he led various research projects as a Research Engineer. From 2006 to 2009, he was a Research Engineer at Neowave, Inc., Anyang, Korea, where he developed Mobile WiMAX (IEEE 802.16) devices. He was an Assistant Professor in the Department of Mobile Communications Engineering at Changshin University, Changwon, Korea. Since March 2016, he has been with the Department of Computer Software Engineering, Soonchunhyang University, Asan, Korea, where he is currently an Assistant Professor. His research interests comprise Cloud Computing, Internet of Things, Software Defined Networking, Mobile System/Communications, and Machine Learning based on Bigdata.