Journal of Computer Science and Electrical Engineering
ISSN: 2663-1946
DOI: https://doi.org/10.61784/jcsee3023

A BLOCKCHAIN-BASED SCHEME FOR CLOUD STORAGE DATA ACCESS CONTROL AND ANTI-COPYING

Jie Huang[1,2,*], JiangYi Yi[2]
[1]Hunan Provincial Engineering Research Center for Missile Maintenance, Changsha 410024, Hunan, China.
[2]Department of Aviation Electronic Equipment Maintenance, Changsha Aeronautical Vocational and Technical College, Changsha 410024, Hunan, China.
Corresponding Author: Jie Huang, Email: huangjie918@163.com

Abstract: This paper addresses data usage security issues and proposes a blockchain-based cloud storage data access control and anti-replication solution. The scheme introduces CP-ABE (Cipher Policy Attribute-Based Encryption) encryption technology and digital watermarking technology. By combining the two, a watermark embedding and CP-ABE encryption model based on orthogonal operation domains is proposed, and specific watermark embedding and CP-ABE encryption methods for image-type data based on orthogonal operation domains are provided. The solution stores request-confirmation records in the blockchain, further strengthening user access control rights and establishing a correlation between the requestor and the watermark in the data, preventing the requestor from denying the fact of their illegal copying. Through security analysis, the solution is shown to be secure and feasible.
Keywords: Blockchain; Cloud storage security; Access control; Anti-copying

1 INTRODUCTION

With the deepening of social informatization, whether it be enterprises, institutions, or individuals, a large amount of data is generated every day. Consequently, the demand for data storage is rapidly increasing. The emerging cloud computing technology can effectively address the issues brought about by storage. As one of the important services provided by cloud computing, cloud storage offers users remote data storage services. Thanks to its flexibility, pay-as-you-go model, freedom from hardware and software maintenance, and the ability for storage space to be continuously expanded, cloud storage services have become the preferred solution for data storage for enterprises, institutions, and individuals. However, there are also many data security issues [1]. This paper proposes a blockchain-based cloud storage data access control and anti-copying scheme by integrating blockchain technology, digital watermarking technology, and ciphertext policy-based attribute encryption technology[2].

2 ENCRYPTION MODEL

If data is subjected to both watermark embedding and CP-ABE encryption simultaneously, the latter operation will compromise the former. The CP-ABE encryption process will disrupt the structure of the data with embedded watermarks, making it impossible to extract the watermark from the encrypted data. Similarly, the watermark embedding process will destroy the structure of the encrypted data, resulting in the inability to decrypt the plaintext from the encrypted data[2]. To better integrate CP-ABE encryption technology with digital watermarking, a watermark embedding and CP-ABE encryption model based on orthogonal operation domains is proposed, building on the concept of selective encryption[3,4]. The data D that needs to be encrypted and have a watermark embedded is split into two complementary subsets, D_1 and D_2. Here, $D = D_1 \cup D_2 \left(D_1 \cap D_2 = \phi\right)$, where D_1 is the CP-ABE encryption operation domain, and D_2 is the digital watermark embedding operation domain. In the orthogonal operation domains, the CP-ABE encryption operation $CE(\cdot)$ and the digital watermarking operation $W(\cdot)$ satisfy the following relationship:

$$\int_D CE(x)W(x)dx = \int_{D_1} CE(x)W(x)dx + \int_{D_2} CE(x)W(x)dx$$

$$= \int_{D_1} CE(x)dx \cdot \int_{D_1} W(x)dx + \int_{D_2} CE(x)dx \cdot \int_{D_2} W(x)dx$$

$$= \int_{D_1} CE(x)dx \cdot 0 + 0 \cdot \int_{D_2} W(x)dx$$

$$= 0$$

It can be seen that *CE(·)* and *W(·)* achieve orthogonality in their respective operation domains. Not only does this allow for the embedding and extraction of digital watermarks in the encrypted data *D*, but it also ensures that the encryption and decryption operations performed on the watermarked data *D* are not affected by the watermark embedding and extraction operations[5]. In other words, the encryption and decryption processes of data *D* and the embedding and extraction processes of the digital watermark are independent of each other.

The schematic diagram of the watermark embedding and CP-ABE encryption model based on orthogonal operation domains is shown in Figure 1. Watermark Embedding and Encryption: The dataset *D* is orthogonally decomposed into two non-overlapping subsets[6-8], D_1 and D_2. CP-ABE encryption operation is performed on D_1, denoted as $CE(D_1) \rightarrow D_1'$, and digital watermark embedding operation is performed on D_2, denoted as $W(D_2, w) \rightarrow D_2'$. The results of D_1' and D_2' are merged to obtain the watermarked ciphertext D_{we}. Watermark Extraction: The data D_{we} is once again decomposed into orthogonal operation domains. The watermark extraction algorithm is applied to D_2' to extract the watermark. If the watermark is encrypted, the encrypted watermark ciphertext W_e must be decrypted to obtain the plaintext watermark *w*. Decryption: The data D_{we} is again decomposed into orthogonal operation domains. CP-ABE decryption is performed on D_1' to obtain the plaintext D_1. D_1 and D_2 are then combined to retrieve the watermarked plaintext D_w.

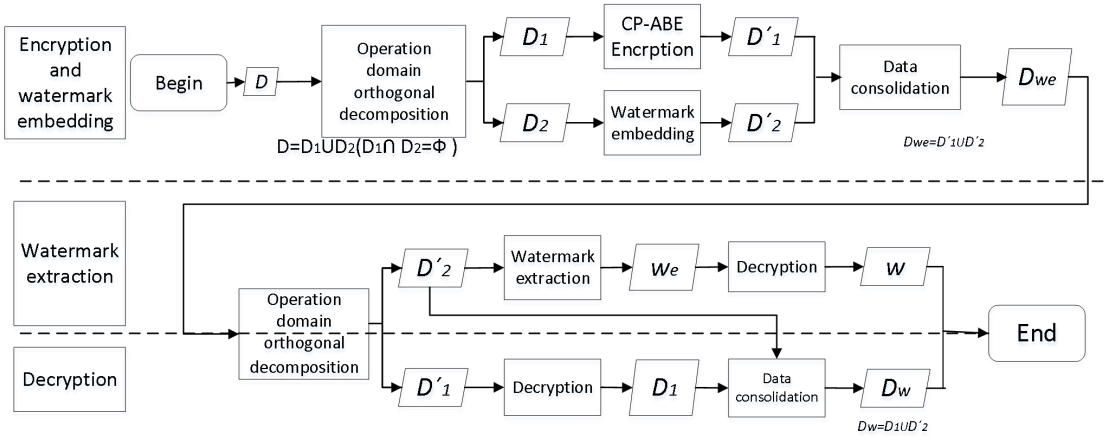

Figure 1 Watermark Embedding Based on Orthogonal Operation Domains and CP-ABE Encryption Model

3 A WATERMARK EMBEDDING AND CP-ABE ENCRYPTION MODEL BASED ON ORTHOGONAL OPERATION DOMAINS

3.1 Scheme Model

Based on the model introduced in the previous section, a blockchain-based cloud storage data access control and anti-replication scheme is proposed. Utilizing CP-ABE encryption technology can effectively control data access, while data embedded with digital watermarks can effectively prevent illegal copying[9]. The scheme model is shown in Figure 2, where there are four roles in the scheme: the data owner, the data requester, the CSP (Cloud Service Provider), and the BCN (Blockchain Network)[10]. The following is a detailed description of each role:

1) Data Owner

In the model of this scheme, the data owner holds the authority for data access control. The data owner primarily undertakes two types of tasks within the model. On one hand, they grant direct access authorization to the data requesters without the need for the CSP (Cloud Service Provider) to act as an intermediary for access authorization. On the other hand, before uploading data to the cloud storage server, the data owner preprocesses the data through a preprocessing algorithm.

2) Data Requester

The data requester, acting as the requesting party in the model, must adhere to the data request protocol. They need to submit a request to the data owner and also request the CSP to download the data.

3) CSP

The CSP (Cloud Service Provider) offers cloud storage services to the data owner and also provides a means for the data requester to download data. Before transmitting data to the data requester, the CSP adds the requester's identification information as a watermark to the data through a watermarking algorithm.

4) BCN

The BCN (Blockchain Network) in the model provides tamper-evident storage services for other roles. The request records of the data requester and the confirmation records of the data owner are stored in the BCN, which can provide effective proof for the data flow process.

The following section will give a detailed introduction to the preprocessing algorithm, the watermarking algorithm, and the data request protocol in the model.

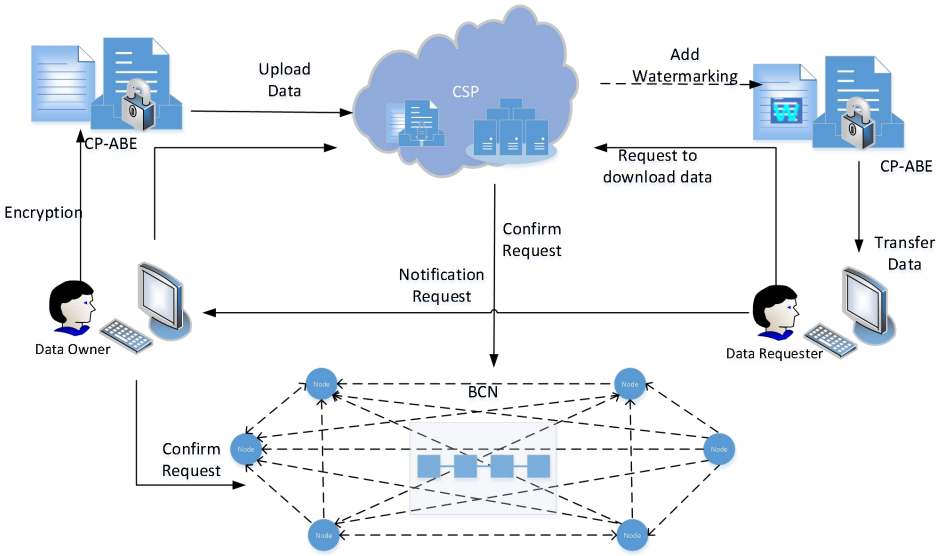

Figure 2 Model for Cloud Storage Data Access Control and Anti-Copying Based on Blockchain Technology

3.2 Preprocessing Algorithm

Before the data owner uploads data to the cloud storage server, they need to preprocess the data. The flowchart of the preprocessing algorithm is shown in Figure 3. First, determine the data type and generate a data type description, then perform orthogonal decomposition based on the data type. Next, encrypt the decomposed partial data with CP-ABE (Attribute-Based Encryption). Finally, merge the data and upload it to the cloud storage server. The specific steps of the algorithm are as follows:

Step 1：Initialization

The user inputs the security parameters, then applies to the authorization authority for the public key *PK* and the master key *MK* of the CP-ABE encryption algorithm, and constructs the access structure tree *T*.

Step 2：Generate data type description *K*

Generate a data type description *K* for data *D* based on its data type, which includes image data, audio data, text data, video data, and other data types.

Step 3：Orthogonal decomposition

Based on the data type *K*, the corresponding data orthogonal decomposition method is used to decompose *D* into two complementary subsets D_1 and D_2, which satisfy condition $D = D_1 \cup D_2 \left(D_1 \cap D_2 = \phi\right)$. D_1 is the domain for CP-ABE encryption operations, and D_2 is the domain for digital watermark embedding operations.

Step 4: Data encryption

Utilizing the CP-ABE encryption algorithm, input the public key *PK* and access structure *T* to perform the encryption operation *CE(D1, PK, T)* on D_1, obtaining the encrypted data D_1', which is denoted as $CE\left(D_1, PK, T\right) \rightarrow D_1'$.

Step 5：Merge data

Integrate D_1', D_2 and *K* to create the encrypted data D_e.

Step 6: Upload data

Upload the data D_e to the cloud storage service.

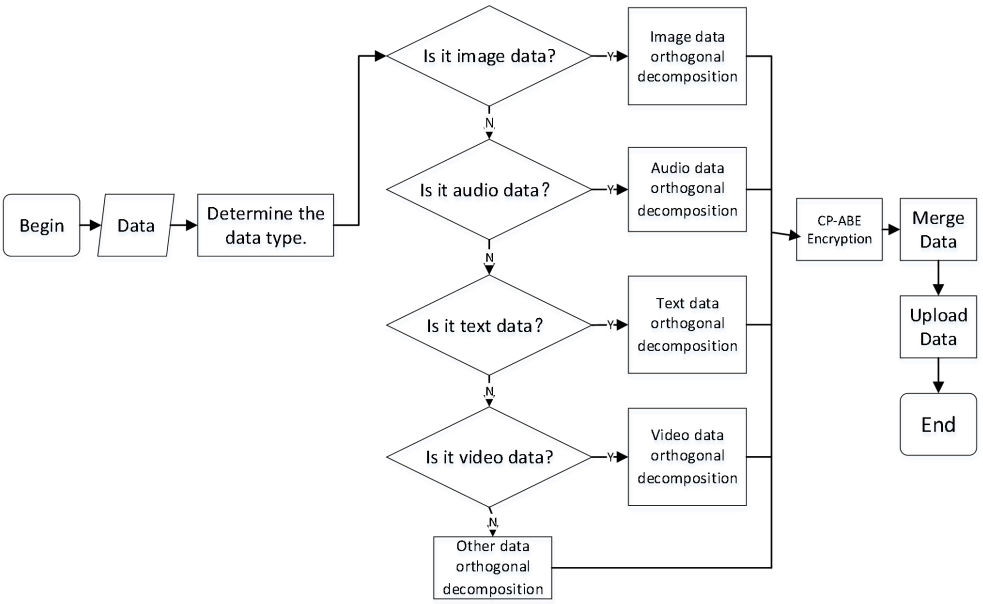

Figure 3 Preprocessing Algorithm Flow Chart

3.3 Watermarking Algorithm

Before CSP transmits data to the data requester, a digital watermarking operation needs to be performed on the data to be transmitted. The flowchart of the watermarking algorithm is shown in Figure 4. First, extract the data into which the watermark will be embedded, then embed the corresponding type of digital watermark based on the data type description, and finally merge the watermarked data with the data encrypted by the data owner. The specific steps of the algorithm are as follows:

Step 1：Data separation, extracting D_2 and K from D_e.

Step 2：Select the digital watermark embedding algorithm, specify K based on the data type, and choose the corresponding digital watermark embedding algorithm $W(\cdot)$ for that data type.

Step 3：Generate a digital watermark by creating a user-identifiable digital watermark w from the signature $Sign_R$ of the data requestor.

Step 4：Embed the digital watermark by embedding w into D_2 to obtain D_2', which is $W(D_2, w) \rightarrow D_2'$.

Step 5：Data merging, combine D_2'、D_1' and K to obtain the watermarked ciphertext data D_{we}.

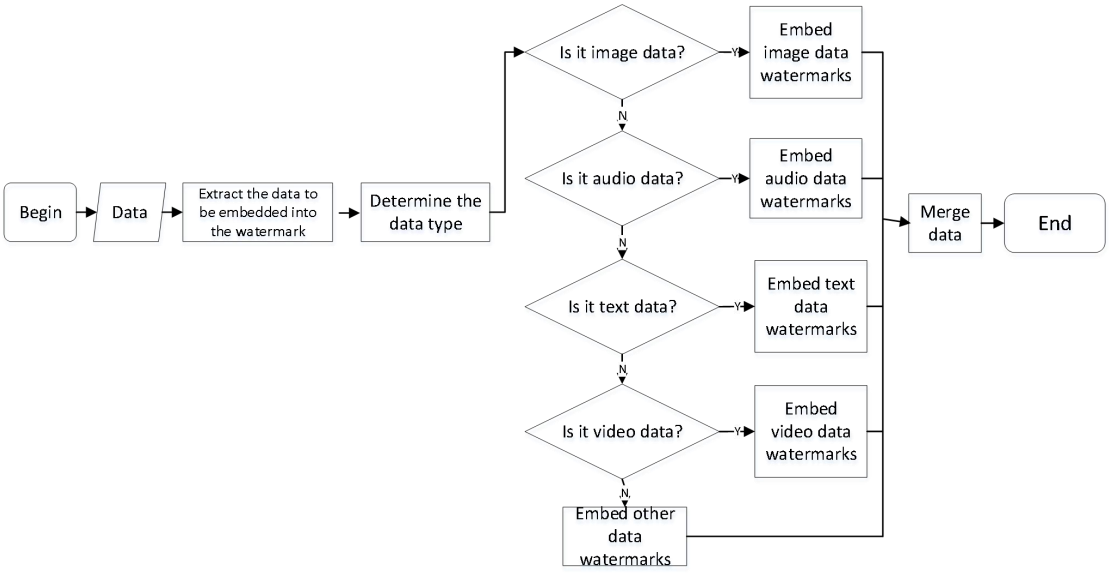

Figure 4 Add Watermark Algorithm Flowchart

3.4 Data Request Protocol

The data request protocol is initiated by the data requester, and is completed with the participation of multiple parties including the data owner, the data requester, the CSP, and the BCN. The data owner grants data access authorization to the data requester, the CSP provides the encrypted data with digital watermarking to the data requester. The BCN records the request process and authorization process between the data owner and the requester, and provides valid proof for this. The timing diagram of the data request protocol is shown in Figure 5. The specific process of the protocol is as follows:

Step 1: Initiating the request .

The data request party sends a data request $Q\{D_{id}, Sign_R\}$ to the BCN, confirming that the request is recorded on the blockchain.

Step 2: Confirming the request .

The smart contract within the BCN will send a data request to the data owner, and the data owner will initiate a confirmation request to the BCN.

Step 3: Initiating the CSP request.

The data request party sends the request $Q\{D_{id}, Sign_R\}$ to the CSP.

Step 4: Adding Watermark.

The CSP verifies whether the request $Q\{D_{id}, Sign_R\}$ is legitimate, which involves checking whether the request $Q\{D_{id}, Sign_R\}$ exists in the blockchain and whether there is a confirmation record. If the request is legitimate, the watermarking algorithm $W(D_e, Sign_R) \rightarrow D_{we}$ is used to obtain the watermarked ciphertext D_{we}, which is then transmitted to the data request party.

Step 5: Generating the private key SK .

The data request party sends their attribute set r to the authoritative authorization agency. The authoritative authorization agency inputs the system master key MK and the attribute set γ and $CK(MK, \gamma) \rightarrow SK$ to derive the private key SK, which is then returned to the data request party.

Step 6: Decrypting the data.

The data request party uses the CP-ABE decryption algorithm to decrypt the element D_1' within D_{we} from the ciphertext $CD(D_1', SK)$. If the attribute set γ satisfies the condition $\gamma \in T$, then $CK(D_1', SK) \rightarrow D_1$, the plaintext D_1 is obtained. The data request party merges D_1 with D_2' to obtain the watermarked plaintext D_w.

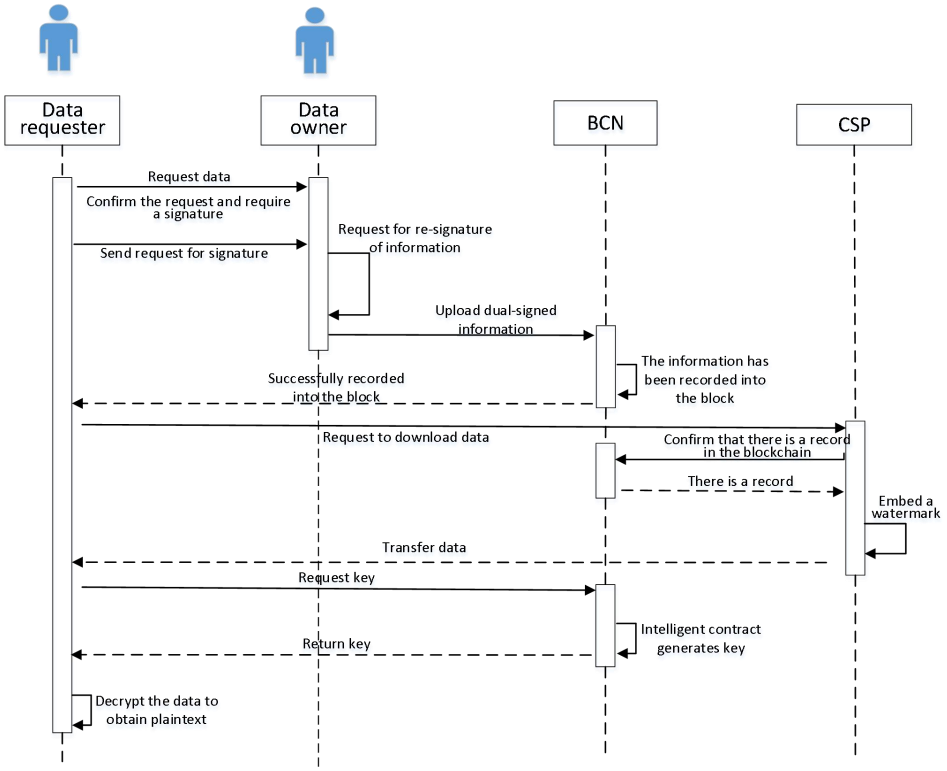

Figure 5 The Sequence Diagram of Multi-Party Participation in Data Request

4 SECURITY ANALYSIS

4.1 The Issue of Illegally Obtaining Data

Illegal users attempt to attack the system through unlawful means to obtain data. In the proposed solution, there are two layers of protection to ensure that data is not acquired by unauthorized users. The first layer of protection is that the unauthorized data requester cannot obtain data from the CSP. The CSP determines the legitimacy of the request based on whether a request-confirmation record exists in the blockchain, and the unauthorized data requester is unable to forge such a record. The second layer of protection is that, as a semi-trusted third party, the CSP poses a risk of colluding with unauthorized data requesters to attack the system. However, even if the unauthorized data requester acquires the data, it is in ciphertext form. Without the key, the unauthorized data requester cannot extract meaningful information from the encrypted data. Therefore, the solution can effectively prevent data from being obtained illegally.

4.2 The Issue of Illegal Data Copying

Malicious users, after obtaining the plaintext data, proceed to illegally copy it and spread it on the internet. In the proposed solution, the data obtained by the requester has already had their identification information embedded in it. In the event of a data leak, the watermark information within the data can be used to immediately trace back to the user who made the illegal copy, and hold them accountable. Therefore, for their own interests, users will strengthen the secure storage of the data after obtaining it, to prevent data leaks from occurring.

4.3 The Issue of Denying the Fact of Illegal Data Copying

After a user's act of illegally copying data is discovered, they may deny the fact of their unauthorized copying. Within the model, there are two pieces of evidence that effectively prove their act of illegal data copying. The first is the digital watermark contained in the data, which establishes the connection between the data and the watermark, proving that the data was disseminated by the user corresponding to that watermark. The second is the existence of a request-confirmation record with their signature in the blockchain, which establishes the connection between the user and the watermark, proving that the user indeed downloaded data containing their identity information watermark. These two pieces of evidence effectively prove the corresponding relationship between the data and the user, thus making it impossible to deny the fact of illegal copying.

4.4 The Issue of Discarding Watermarked Data

There are users who, in an attempt to hide their personal information, actively discard the portion of data in the orthogonal domain where the watermark is embedded and only use the data that has not been watermarked. To prevent the discarding of watermarked data, when selecting the characteristic positions for watermark embedding in the carrier data, positions that are more significant and have a substantial impact on the fidelity of the data are chosen. The following experiments reveal that the absence of the watermarked portion of the data leads to distortion of the image, rendering it useless. If a user discards the watermarked portion of the data, it will affect their ability to use the data, hence they will not actively discard the watermarked data.

5 CONCLUSIONS

Blockchain-based cloud storage data access control and anti-copying solution. In this solution, the data owner has two layers of protection to ensure they maintain actual control over data access. The first layer of protection is the request-confirmation record. Before the CSP (Cloud Service Provider) transmits data, it confirms whether there is a request-confirmation record between the data owner and the requester in the blockchain. The second layer of protection is the CP-ABE (Cipher Policy Attribute-Based Encryption) encrypted access control policy, where the data requester can only decrypt the data and obtain the plaintext if they meet the access control policy. In this solution, the data is embedded with a watermark that identifies the information of the data requester, which can effectively prevent the data from being illegally copied. At the same time, by using the request-confirmation records stored in the blockchain, it can effectively prove the corresponding relationship between the watermark information and the requester, preventing the requester from denying their act of illegal copying. This solution is secure and feasible.

CONFLICT OF INTEREST

We all declare that we have no conflict of interest in this paper.

DATA SHARING AGREEMENT

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable reque st.

FUNDING

This study was funded by Natural Science Foundation of Hunan Province in 2022 : " Design and application of cloud storage security architecture based on blockchain". (Research funder: Jie Huang, Grant number: 2022JJ60091).

ETHICAL APPROVAL

This article does not contain any studies with human participants or animals performed by any of the authors.

REFERENCES

[1] Sun X. Critical security issues in cloud computing: a survey. 2018 IEEE 4th International Conference on Big Data Security on Cloud (Big Data Security), IEEE International Conference on High Performance and Smart Computing, (HPSC) and IEEE International Conference on Intelligent Data and Security (IDS). IEEE, 2018, 216-221. DOI: 1109/BDS/HPSC/IDS18.2018.00053.

[2] Pierro M D. What Is the Blockchain?. Computing in Science and Engineering, 2017, 19(5): 92-95.

[3] Zhang Y, Xu C, Lin X, et al. Blockchain-Based Public Integrity Verification for Cloud Storage against Procrastinating Auditors. IEEE Transactions on Cloud Computing, 2019: 9(3): 923-937.

[4] Yining Q I, Yongfeng H. DIRA: Enabling Decentralized Data Integrity and Reputation Audit via Blockchain. Science China Technological Sciences, 2019, 62(004): 698-701.

[5] Li G, Sato H. A privacy-preserving and fully decentralized storage and sharing system on blockchain. 2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC). IEEE, 2019, 2: 694-699.

[6] Gao S , Piao G , Zhu J , et al. Trust Access: A Trustworthy Secure Ciphertext-Policy and Attribute Hiding Access Control Scheme Based on Blockchain. IEEE Transactions on Vehicular Technology, 2020, 69(6): 5784-5798. DOI: 10.1109/TVT.2020.2967099.

[7] Ma M, Shi G, Li F. Privacy-oriented blockchain-based distributed key management architecture for hierarchical access control in the IoT scenario. IEEE Access, 2019, (7): 34045-34059. DOI: 10.1109/ACCESS.2019.2904042.

[8] Zhang Y B, Cui M, Zheng L J, et al. Research on electronic medical record access control based on blockchain. International Journal of Distributed Sensor Networks, 2019, 15(11): 1-13.

[9] Mosleh M, Setayeshi S, Barekatain B, et al. High-capacity, transparent and robust audio watermarking based on synergy between DCT transform and LU decomposition using genetic algorithm. Analog Integrated Circuits and Signal Processing, 2019, 100(3):513-525.

[10] Nair U, Birajdar G K. Compressed domain secure, robust and high-capacity audio watermarking. Iran Journal of Computer Science, 2020, 3: 217-232. DOI: https://doi.org/10.1007/s42044-020-00059-x.