Research on access control method of Digital Archives based on blockchain

Rui Men
The Information Engineering College Long Dong University Qingyang City, China
531039764@qq.com

Abstract. This paper analyzes the security problems that need to be solved urgently in digital archives system: First, the centralized storage of digital archives makes it overwhelmingly dependent on the third-party storage organization. Second, every transaction has its own access proof and its own authorization proof, it’s not transparent to others. Then the blockchain technology is introduced to solve the existing problems. We store the digital file attachment and file attributes in the private IPFS cluster after encryption. The address of IPFS and digital fingerprint of the file are stored on the private blockchain. Additionally, this paper stores and checks the height and hash value of the latest irreversible block of the private blockchain, to guarantee the authenticity of data on private blockchain. To solve the second problem, we store the ABAC on public blockchain, add time and position information to the attributes of the subject, to achieve fine-grained access control. Finally, this model is implemented on fabric1.0.3.

1. Introduction
Archives are very important data records, which are the original information, with preservation value, and formed directly in various social activities. Different from the general library information and electronic document information, the essential attribute of archives is the original recording, which enables the archives to restore the real historical situation, so it has important preservation and reference value, and has legal effect[1]. The traditional paper archives management method has the disadvantages of slow searching and complicated management process, while the digital archives management system can improve the retrieval speed of archives through database query, simplify the management process of archives through online approval, reduce the human cost and improve the efficiency of business. However, its security issues are widely concerned, and file tampering incidents occur from time to time. How to protect the authenticity and security of digital archives from theft, tampering and destruction has become a hot issue in the field of digital archives construction, and strengthening access control of digital archives can avoid most of the security issues. The characteristics of the blockchain, such as decentralization, strong tamper proof and information traceability, are very suitable for the scene of digital file protection. Therefore, the access control process of the digital file is deployed on the blockchain, which greatly improves the security of the digital file system.

2. Related works

2.1. Access Control
In recent years, many people continue to explore and practice the blockchain technology in the access control of digital archives. Asaph Azaria and Ariel Ekblaw et al. [2] build a decentralized medical data access and authority management system by using smart contracts. The system realizes the ownership of patients' medical data and enables patients to independently access medical records, sharing and management. In [3], using blockchain to solve the problem of cross-organization access control in RBAC, realizing the authentication of users' roles in different organizations. In [4-6], using blockchain to ensure that user's identity attributes and access control policies cannot be modified by malicious users; The policies and authority exchanges are open on the blockchain to prevent refusing to implement the rights granted by the policies. In [7], using blockchain in a data storage system as a trusted database to access data. In [8-9], using blockchain to record the granting, using, circulation, and other operations of authority.

2.2. Blockchain

The blockchain structure is based on the data blocks and linked in chronological order. The data blocks are generated by the distributed nodes through consensus algorithm and guaranteed by certain economic incentives that all nodes are motivated to participate in the activities of the blockchain. All nodes in the distributed system have equal status, without any centralized special nodes, and each node will verify the blocks. Data, block data dissemination, so as to ensure that a small number of nodes will not affect the operation of the whole blockchain system. Each block includes five fields: block head, block size, magic number, transaction number and transaction. Among them, the transaction field records the specific transaction information list, the magic number is a fixed value, and the block head field is the extraction of all transaction contents in the block. For example, it is also the key to building blockchain. As shown in Figure 1:

![Figure 1 the figure of Blockchain structure](image)

3. Storage of Digital Archives based on Blockchain

IPFS (Internet File System) is a globally interconnected distributed file system, which integrates many excellent ideas of point-to-point system, including distributed hash table, block exchange, version control system and self-certification file system. It has the characteristics of content addressable, tamper-free and decentralized. When storing a file, IPFS calculates the hash value based on the file content and adds it to the global distributed hash table. When getting the file, the IPFS cluster looks up the storage node of the file from the global hash table according to the file address, then takes the file from the node and validates it and returns it to the user. IPFS can be divided into two types: private IPFS cluster and public IPFS cluster. Public IPFS refers to the worldwide IPFS distributed network. Anyone can participate in the network as a node and have permissions of write and read the data. Private IPFS is a distributed network limited to the internal users of a group or organization. Nodes trust each other through sharing swarm key, and other nodes are not allowed to participate to the network. In this paper, we designed the storage mode of digital archives management system, IPFS and distributed database technology are introduced to store and protect the digital archives. Data storage encapsulates Blockchain, IPFS and distributed database, and provides access operation for upper layer to call. The upper layer provides RESTful interface for legal users after access control calculation. In order to ensure the security and privacy of the file data, the file attachments and file attributes are encrypted and stored in the private IPFS cluster, and the IPFS address and digital fingerprint of the file are stored in the private blockchain. The Structure of the system is shown in Figure 2:
4. Access control method of Digital Archives based on blockchain

At present, the main access control methods are: RBAC (role-based access control), ABAC (attribute based access control), UCON (usage control) and CapBAC (capacity based access control).

The popularity of RBAC is because it is close to real life and convenient for people to understand and use. However, in complex scenes, RBAC is gradually not enough. It will produce many nihilistic roles and this is more difficult to manage and control:

ABAC can solve these problems very well. In ABAC, the following concepts need to be defined:

Subject: refers to the entity that can access and operate the protected resources. It can be user, process, etc.

Object: the object, the entity that can be accessed and operated.

Environment: refers to the environment, time, context information, etc. in the process of accessing resources.

Operation: read, write, modify, execute, etc.

The process of request resource as follow:

1. Set a request
2. Calculation: subject attributes + object attributes + environment conditions = allow or deny?
3. Access to resources.

The key step is (2), after setting the request, subject attributes, object attributes and environment conditions are taken as input, PEP obtains rules, PDP calculates them, and finally determines whether it has the right to request or not[10]. The control mechanism is shown as Figure 3[10]:

In this system, subject is the user who wants to access or operate the archive resources. The attributes include the company, age, role, application access time, position, approval authority and authority. Object is a digital archive resource, and its attributes include secret levels, allowed operation, allowed access time, allowed access position, required approval authority, etc. The environment is already included in the subject and object and is not listed separately. The rule of calculation process when applying for access to archive is as shown in Figure 4:
5. The implementation of ABAC based on fabric

The smart contract of fabric is called chaincode. Chaincode is the software running on ledger, which encodes user assets and executes transaction instructions to modify assets, so we set our ABAC principles on hyperledger fabric 1.0.3 blockchain platform. For example, The codes are shown in Figure 5:

These code show us that only permit the ID with organization name of "org1msp", "attr1" value of true, and organization name of "org2msp", "attr2" value of false can continue to execute the business logic code of the chain code. Users with other conditions have no permission to continue to execute. Chaincode developers can add similar code for access control before the business logic code of chain code, and implement different access control logic according to needs.

6. Conclusion

This paper combines the access control of digital archives with the blockchain technology, and stores the user's rights, contents and secret level information on the blockchain. After the user sends the resource access request, the authorization verification process is automatically calculated and concluded by the smart contract. This method is open, transparent and undeniable. On one hand, it can greatly improve the security of the digital archives system; on the other hand, it can greatly improve the security of the digital archives system. In addition, it expands the application field of blockchain technology and promotes its theoretical research.

References

[1] Mayssa JEMEL1, Ahmed SERHROUCHNI. Decentralized access control mechanism with temporal dimension based on blockchain. The Fourteenth IEEE International Conference on e-Business Engineering, 2017:177-182.

[2] Azaria A, Ekblaw A, Vieira T, et al. MedRec: Using Blockchain for Medical Data Access and Permission Management[C]// International Conference on Open and Big Data. IEEE, 2016:25-30.
[3] Cruz JP, Kaji Y, Yanai N. RBAC-SC: Role-based access control using smart contract. IEEE Access, 2018, 6:12240–12251. [doi:10.1109/ACCESS.2018.2812844]

[4] Alansari S, Paci F, Sassone V. A distributed access control system for cloud federations. In: Proc. of the 2017 IEEE 37th Int'l Conf. on Distributed Computing Systems (ICDCS). IEEE, 2017. 2131–2136. [doi:10.1109/ICDCS.2017.241]

[5] Alansari S, Paci F, Margheri A, Sassone V. Privacy-preserving access control in cloud federations. In: Proc. of the 2017 IEEE 10th Int'l Conf. on Cloud Computing (CLOUD). IEEE, 2017. 757–760. [doi:10.1109/CLOUD.2017.108]

[6] Maesa DDF, Mori P, Ricci L. Blockchain based access control. In: Proc. of the IFIP Int'l Conf. on Distributed Applications and Interoperable Systems. Cham: Springer-Verlag, 2017. 206–220. [doi:10.1007/978-3-319-59665-5_15]

[7] Hashemi SH, Faghri F, Campbell RH. Decentralized user-centric access control using PubSub over blockchain. arXiv preprint arXiv:1710.00110, 2017.

[8] Ouaddah A, Elkalam AA, Ouahman AA. Towards a novel privacy-preserving access control model based on blockchain technology in IoT. In: Proc. of the Europe and Mena Cooperation Advances in Information and Communication Technologies. Cham: Springer-Verlag, 2017. 523–533. [doi: 10.1007/978-3-319-46568-5_53]

[9] Ouaddah A, Abou Elkalam A, Ait Ouahman A. FairAccess: A new blockchain-based access control framework for the Internet of things. Security and Communication Networks, 2016, 9(18):5943–5964. [doi: 10.1002/sec.1748]

[10] https://blog.csdn.net/XiaoBeiTu/article/details/100773968