Journal of Combinatorial Mathematics
and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Design and Implementation of a Secure Storage and Traceability System for Financial Documents Based on Blockchain Technology

Linlin Zhang[1,✉], Yuening Wang[1], Hui Lu[2]

[1] *Financial Sharing Service Center, Yunnan Power Grid Co., Ltd., Kunming, Yunnan, 650000, China*

[2] *Information Center, Yunnan Power Grid Co., Ltd., Kunming, Yunnan, 650000, China*

ABSTRACT

As the digital economy develops, the use of digital methods for the storage of financial documents is being commonly adopted. To protect the data security of financial documents and ensure the traceability of data, this article designed a secure storage and traceability system for financial documents based on blockchain. Firstly, a blockchain platform with a private chain that had better security and scalability was selected, and the data structure and smart contract design were then carried out. Secondly, an identity authentication and permission management mechanism was established, and data storage and transaction processing modules were designed. Asymmetric encryption was used to secure data, and the digital signatures were combined to ensure the integrity of financial documents. Finally, the traceability function was achieved through the immutability of blockchain technology. The task of storage and traceability was accomplished by designing a secure storage and traceability system for financial documents in conjunction with the blockchain technology. In the experiment, it only took about 50 seconds to process traceability tasks with an interval of 50 people, which was shorter than the traditional system's 180 seconds; it also maintained an accuracy rate of over 90% in traceability tasks with an interval of 100 people; in the face of 1000 network attacks in a short period of time, the financial management system based on blockchain technology was only invaded 20 times, while the traditional financial system was invaded 200 times. This system, in terms of time, traceability accuracy, and data security, were all improved over the traditional system. The design of a secure storage and traceability system for financial documents based on blockchain technology is conducive to strengthening the security of data and the accuracy of traceability.

Keywords: Blockchain Technology, Storage Security, Traceability System, Financial Document

1. Introduction

With the development of information technology, enterprises or organizations in the office process,

✉ Corresponding author.
E-mail address: zhanglinlin8356@163.com (L. Zhang).
Received 25 March 2024; Revised 05 June 2024; Accepted 15 December 2024; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-422

gradually from paper documents to save important information data replacement for the use of electronic documents. Paper documents after the formation of documents need to manually file, organize, save; and electronic documents are generated with the data, anytime, anywhere, you can automatically file the document information management, easy to operate the document at the same time also easy to share [1-2]. Of course, paper documents are not eliminated, because electronic documents are also facing some problems, such as electronic documents have a high degree of volatility, and in the process of its storage, transmission and processing operations are very easy to be modified; electronic documents in the process of sharing is easy to be stolen, tampered with, etc.; electronic documents will be due to the collapse of the system, poisoning and other problems caused by the lack of data [3-5]. Document traceability refers to the process of tracking and restoring document changes by recording the document modification history [6]. In daily work, often need to modify and update the document, but sometimes may encounter the need to retrace the previous version of the situation, document traceability can help us accurately find the historical version, and restore to the specified point of modification.

Financial industry due to the characteristics of the financial industry's own operations, in terms of information security requirements are generally higher, in the security design to fully take into account the factors affecting the business model, not only to ensure the higher availability of business networks, reliability, confidentiality, but also to the internal core data have a strong defense, control capabilities [7]. In order to improve the security and traceability of internal documents in the financial industry, the realization of internal working documents and application systems in the office documents in the internal flow can be safe and controllable, the realization of the document in the flow process and storage after being tampered with at will or attacked by viruses, the need to establish a set of perfect security storage and traceability system [8-9]. To achieve both to prevent the loss or tampering of documents, but not to affect the use of employees.

Blockchain is a decentralized distributed database, with decentralization, tampering and traceability [10]. The immutability of blockchain technology means that once a document is stored on the blockchain, no one can modify the content of the document without authorization [11]. This provides the highest level of security for electronic documents. Blockchain technology stores data on distributed nodes rather than a central server. Such decentralized characteristics ensure the credibility and reliability of the document, even if a node failure or attack, its impact on the system as a whole will be made up by other nodes [12]. The data on the blockchain is publicly available, anyone can view and verify each stored document [13]. This transparency can effectively reduce information asymmetry and fraud, and enhance the credibility of the entire document storage system. Blockchain technology provides an accurate timestamp for each document storage operation, which ensures the chronology and traceability of documents [14]. This is especially important for some documents that need to retain historical records and audit trails. In recent years, blockchain technology has attracted widespread attention in society, and the applications of blockchain in electronic document storage include digital signature, file sharing, copyright protection, etc [15-17].

The purpose of this article is to design a secure storage and traceability system for financial documents based on the blockchain technology, and to explore in depth how to design a financial document management system with the combination of the blockchain technology to improve data security and traceability, so that the current traditional financial documents management system can be improved. Ultimately, it was found that by combining the blockchain technology, the security and traceability accuracy of the data in the financial document management system was greatly improved.

2. Blockchain Technology and Secure Storage and Traceability System for Financial Documents

2.1. Blockchain Technology

Blockchain technology is a decentralized distributed ledger technology. It links data together in blocks, each of which contains a certain amount of transaction information, along with a hash of the previous block. With this technology, it is possible to construct a immutable database.

Traditional databases are prone to tampering, high risk and difficulty in traceability, whereas the problem of data security and reliability can be solved in blockchain by symmetric encryption combined with digital signatures.

2.2. Secure Storage and Traceability System for Financial Documents Based on Blockchain Technology

The combination of blockchain, distributed file encryption storage, and identity authentication management technology can solve the problems of data tampering, transmission, and insecure storage in financial documents, ensuring the authenticity and security of data.

Therefore, this article designs and develops a secure storage and traceability system for financial documents based on blockchain technology, which meets the needs of companies and individuals for secure data storage and full traceability of each financial document. A comprehensive financial document management system can record and track financial data, improve audit and approval efficiency, and better manage costs. The application of blockchain technology can protect financial document information from being tampered with by third parties, while also ensuring the authenticity and traceability of financial information.

In the financial document management system, data related to applicants, approvers, auditors, finance, and external regulatory departments are classified, and standardized processes and formats are established.

The information of financial documents includes date, order number, transaction description, amount, payer/payee information, payment method, tax information, project information, approval information, additional instructions, and related attachments. The system provides a unified data access interface to share the required data from other systems on the chain in a public key encrypted manner. Each node packages the collected data into data blocks and establishes an index. Through hash operations and random adjustments, hash values are formed and stored in the form of notes on each block.

3. System Design

3.1. ramework Design

In terms of system construction, it is necessary to consider the connections between various departments and the permissions of different departments, and establish standards to rely on blockchain technology to ensure data security and build traceability mechanisms. The framework design is shown in Figure 1.

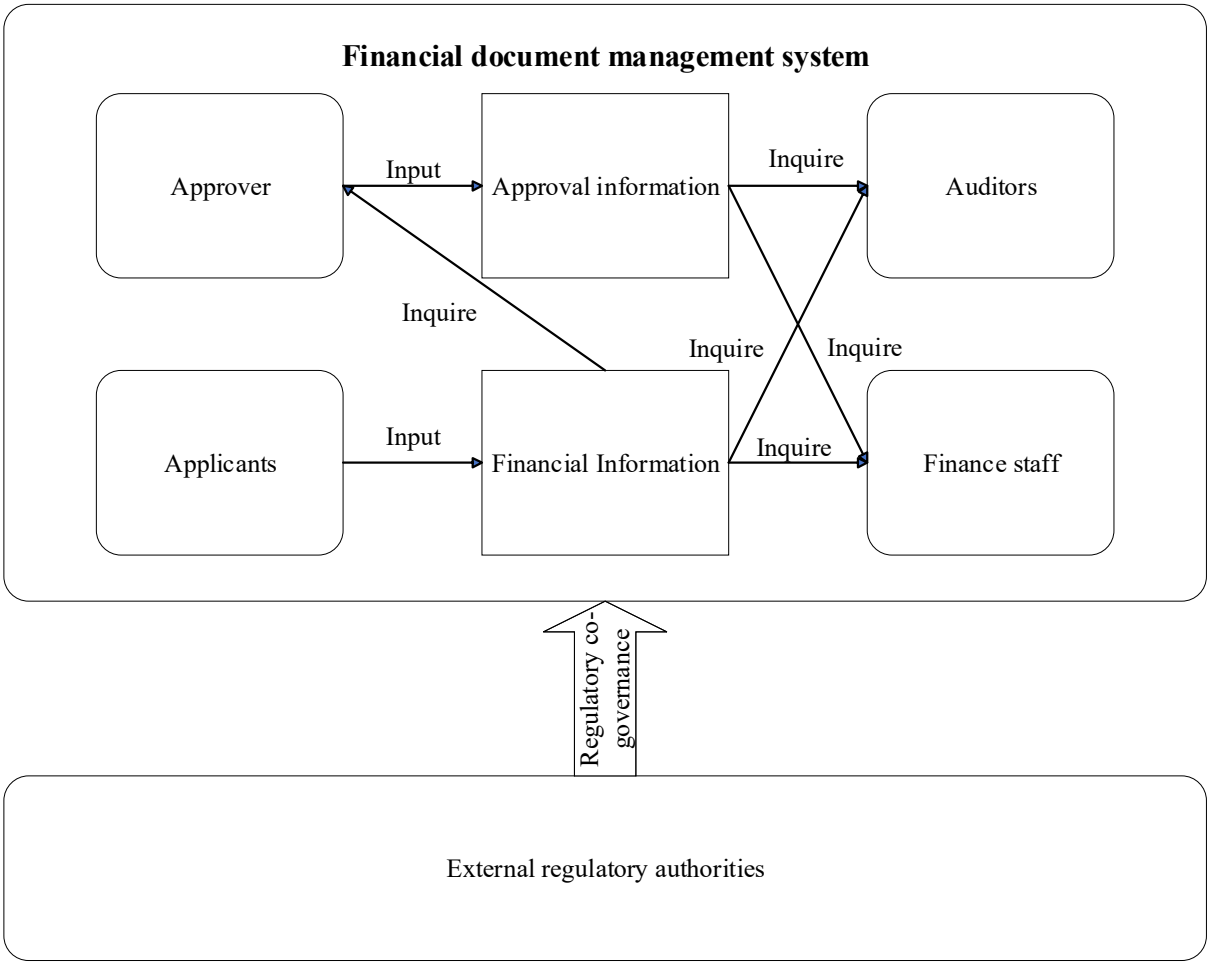

Fig. 1. System architecture design diagram

The designed system architecture mainly consists of two parts. The first is the data module, which is divided into financial information and approval information. The second is the external interface, which is composed of approvers, applicants, auditors, and financial personnel interfaces. Approval information is input into the approval information module. Auditors and financial personnel have the authority to view approval and financial information, but cannot input to ensure data security.

The design principles for establishing a secure storage and traceability system for financial documents based on the above framework are mainly as follows:

1) Information sharing utilization: Firstly, the business related needs of each department are sorted out, and a management system that can facilitate dynamic management is established to ensure that relevant work can be quickly completed when business intersections occur between departments; then, according to the management system, the functions of information management, data exchange, and approval certification are completed, and global information sharing is also achieved to achieve data coherence and financial information traceability.

2) Division of business and regulations on permissions: By dividing the business into different departments, the enterprise regulates and restricts the permissions of different departments, ensuring the correctness and continuity of data while ensuring data security. Based on the work content of the department, the responsible work business is divided, and the upload and viewing of data are restricted with permissions to prevent the mixing of erroneous and malicious information and the leakage of financial document information.

3) Standardization of data content: The system needs to interact and connect with multiple departments, and various financial data and approval opinions need to be provided by each

department. The content formats of various data are not the same, and it is necessary to establish unified standards to standardize the content.

3.2. rchitecture Design

The secure storage and traceability system for financial documents based on blockchain technology can securely store financial document data, directly connect with document issuing personnel, and obtain real and reliable traceability information. Therefore, it is necessary to build a fully functional blockchain system that integrates technologies such as symmetric encryption, hash functions, digital signatures, as well as privacy protection schemes and distributed ledger protocols.

The division of blockchain architecture is generally the application layer, contract layer, consensus layer, network layer, protocol layer, and data layer.

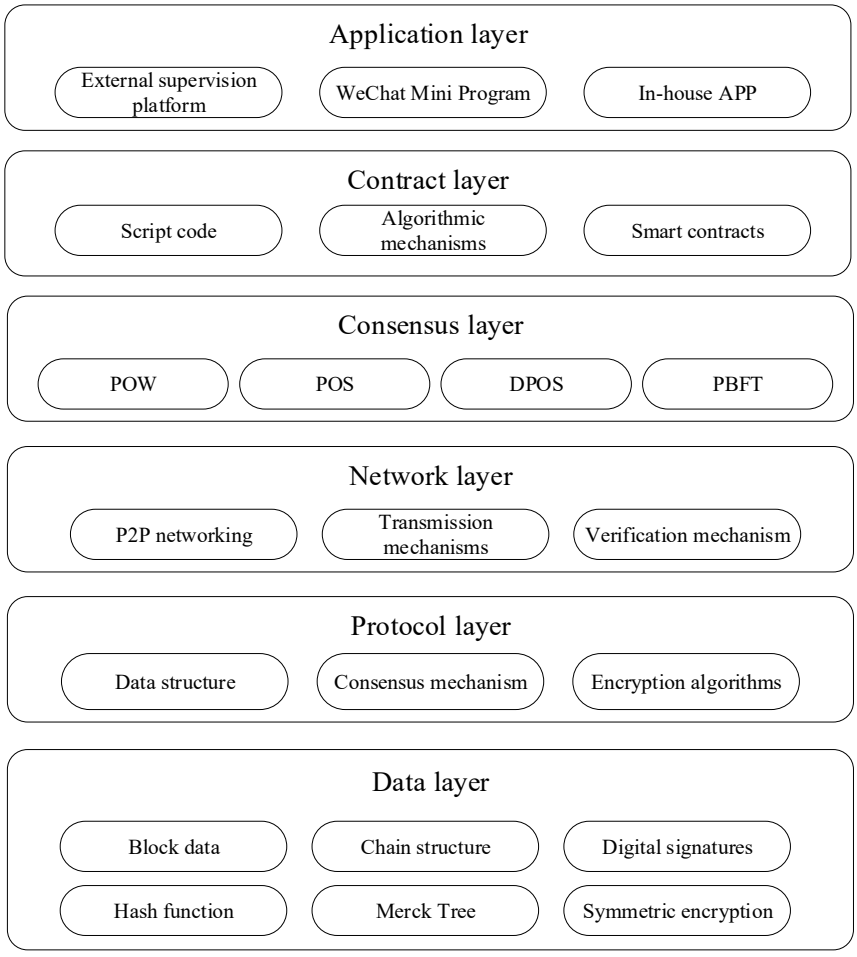

Fig. 2. Hierarchical Architecture of Blockchain

From Figure 2, it can be seen that the application layer is mainly designed for user oriented applications, including mini programs, apps, online web pages, etc. The design mainly consists of a user interface and account management, which can help users quickly and conveniently enter or query documents. The contract layer is the core of smart contracts, including data structures, business logic, permission control, and interaction interfaces, which together ensure the correctness, security, and reliability of contracts.

The consensus layer can help solve the problem of data consistency. Each node in the blockchain needs to maintain a completely consistent ledger data, but various problems may occur during formal use, leading to inconsistent data. Therefore, the consensus layer achieves ledger consistency in untrusted environments through consensus protocols such as POW, POS, DPOS, and PBFT. The

network layer is responsible for encapsulating the networking methods, information transmission protocols, and data verification mechanisms of blockchain. In traditional client/server communication, data is usually provided by the server to the client for use. However, in blockchain, the provider of technical data for each node is also a consumer. This means that every node can access the data, and the blockchain network uses the Gossip protocol to propagate the data. When a new block appears, the block information is transmitted to nearby nodes through broadcasting, and after verification, the new block is officially linked to the blockchain system. The protocol layer is the fundamental component for building the entire blockchain ecosystem, and solves trust issues in distributed systems by determining consensus mechanisms. The consensus mechanism is the basic rule for nodes in blockchain networks to reach consensus, ensuring the security and stability of the network. The data layer is the top priority of the blockchain architecture system, and its existence determines the correlation and construction of system data.

3.3. Software Design

In order to meet the secure storage and traceability requirements of financial document management systems, blockchain platforms need to have the characteristics of connecting all levels of departments. Therefore, in the construction of blockchain platforms, it is necessary to rely on the characteristics of blockchain technology to build a financial information collaborative sharing platform, connect various departments, and achieve full chain tracking of information.

3.3.1. Permission Design.The access and input of company financial data are divided into three levels of permissions, namely first level permissions, second level permissions, and third level permissions. Among them, the first level permission is the ability to obtain all financial information and authorize other users. Without the authorization of the first level permission, other levels of permission cannot access and upload any information. The second level permission can access financial document information and review opinion information under the authorization of the first level permission, while the third level permission can only input financial document information under the authorization of the first level permission.

To prevent users with different permissions from colluding with each other, each user who enters the system needs to undergo real name authentication and generate their unique ID. When a user logs into the system, they need to enter their ID and password for authentication. When users want to obtain certain information, they need to access it with corresponding permissions or apply for authorization from a first level authority. When uploading financial document information, the uploader first needs to select a label as the classification attribute to classify the file. Classification determines the level of permission and whether the uploader has the authority to upload modified information. Secondly, the format of the uploaded content is standardized to prevent information omission and ambiguity.

3.3.2. Algorithm Design.The following algorithms are used in designing the system:

1) Hash algorithm: The Hash algorithm is used to ensure the integrity and immutability of document data. The code is:

hash_value = SHA-256(data)

2) Symmetric encryption algorithm: The symmetric encryption algorithms encrypt and decrypt data with the same key (symmetric key). Only authorised users can access and view the contents of the document so that the privacy and security of the document can be ensured. Figure 3 presents an example of the codes:

```
from cryptography.fernet import Fernet
key = Fernet.generate_key()
cipher_suite = Fernet(key)
original_message = b"Hello, this is a secret message!"
encrypted_message = cipher_suite.encrypt(original_message)
print("Encrypted message:", encrypted_message)
decrypted_message = cipher_suite.decrypt(encrypted_message)
print("Decrypted message:", decrypted_message.decode())
```

Fig. 3. Example of symmetric encryption algorithm code

3) Traceability function: It is used to record the creation history of documents and can use on chain events or logs to record relevant information, such as:
event DocumentCreated(address indexed creator, uint indexed documentId, string documentType);
Here, creator represents the creator of the document; DocumentId represents the unique identifier of the document; DocumentType represents the document type.
4) Identity verification mechanism: It is used to verify the user's identity and permissions, and digital signatures can be used to achieve identity authentication. The code is as follows:
signature = sign(private_key, data)
5) Smart contract: A smart contract is a computing contract that runs on the blockchain and can automatically execute contract operations. However, the intervention of intermediaries can help the blockchain efficiently and securely execute transactions.

3.4. Blockchain Database Design

This article combines blockchain technology to design a blockchain database, ensuring that all operational information in the database is complete and reliable, and avoiding the possibility of tampering. The database management system based on blockchain combines the underlying database technology with blockchain technology, and writes user operation log records into the blockchain for distributed storage. During this process, users can store and query data through the client. The blockchain database framework is shown in Figure 4.

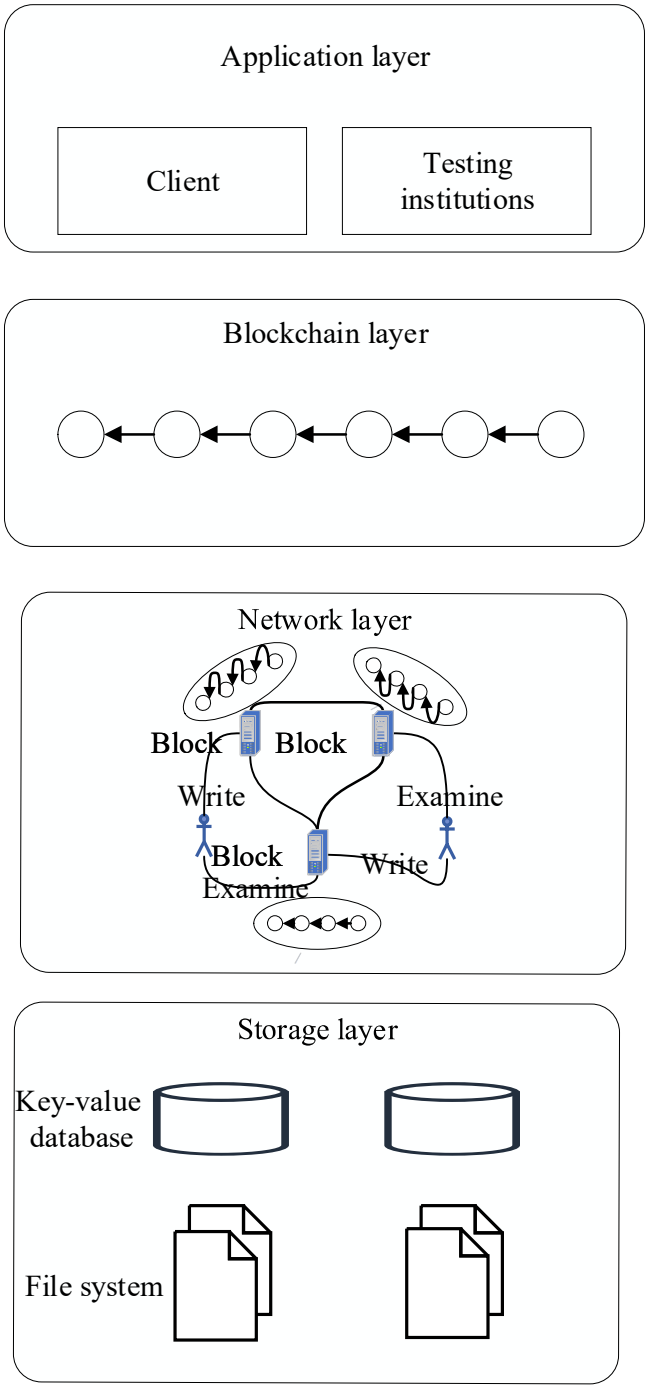

Fig. 4. Blockchain database system framework

3.5. Design of Security Traceability System

There are five modules in the system, namely trusted third party, data querier, data uploader, blockchain platform with smart contract, and blockchain database, as shown in Figure 5.

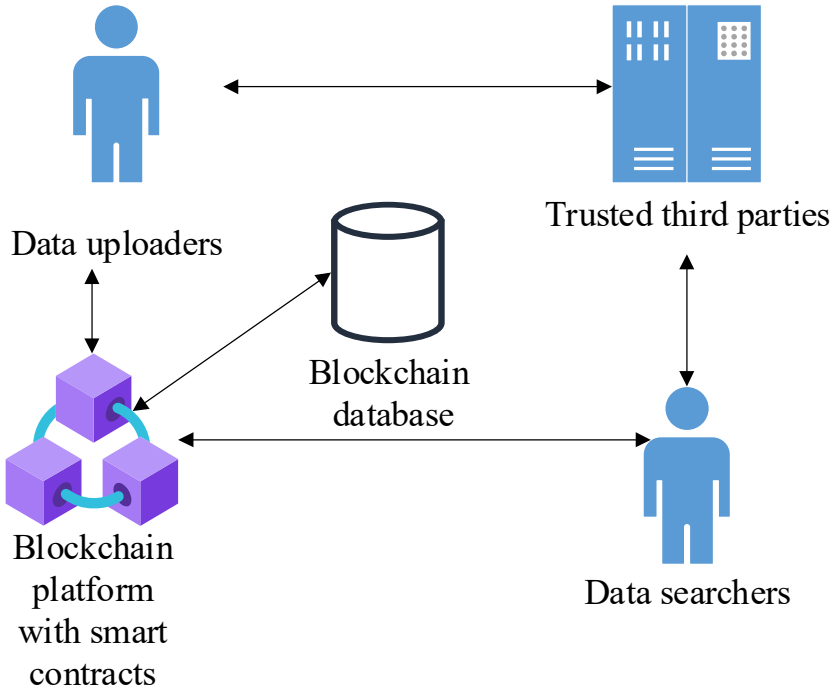

Fig. 5. Schematic diagram of system participants

Among them, trusted third parties have first level permissions, which can distribute keys and assign the highest level of security and management permissions to affiliated organizations. The data queriers have second level permissions and can access the system after being authenticated with first level permissions. The queriers can trace the payment method, execution location of the business, responsible person of the business, and relevant information of the products involved through the traceability system. The data uploader has three-level permissions, including the application initiator, product manufacturing organization, etc., in the system. After passing third-party authentication, they can access the system. The data uploader only has the authority to upload data and is supervised when accessing the system. If malicious data is found, punishment can be imposed. This system adopts a technology architecture based on blockchain technology. Each distributed node in the system can perform data uploading behavior, and each data has a unique label. Unique encrypted digits can be used for searching to ensure the security and accuracy of the data. This system stores financial document information through a blockchain database. Compared to traditional databases, blockchain databases have features such as distributed storage, immutability, decentralization, and transparency, ensuring data security.

4. Specific Implementation

4.1. System Initialization

Firstly, the system is initialized by a third party. Firstly, the init() function is called to initialize the contract and data. Then, after selecting a security factor, a security key is generated for subsequent access nodes to complete dynamic access authentication. Finally, the generated public key is distributed to each data upload node to encrypt the data.

4.2. Data Upload

The flowchart of the data upload function is shown in Figure 6.

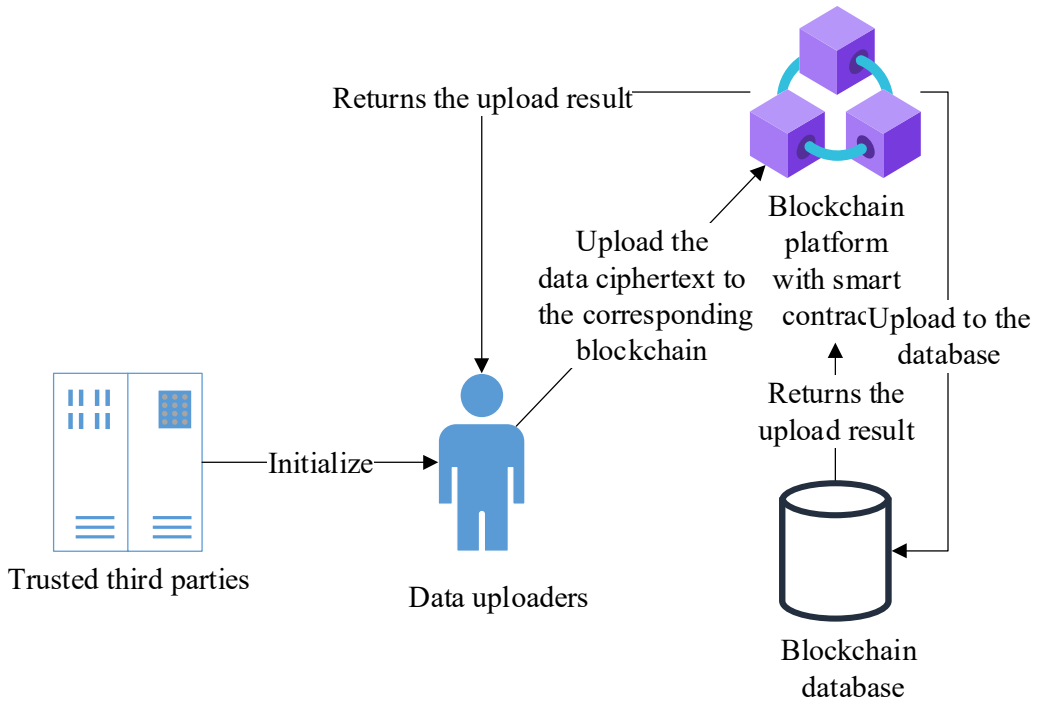

Fig. 6. Data upload process

The data uploader first registers and obtains authorization from a third party, and after authorization, obtains the master key and public key. Afterwards, the data uploader accesses the system and executes the key generation algorithm. As the file and keyword plaintext need to be encrypted separately, two key generation algorithms are used. A searchable encryption key is generated using the symmetric searchable encryption key generation algorithm to encrypt keywords, and the symmetric encryption key generation algorithm is used to generate an AES (Advanced Encryption Standard) key to encrypt files. The specific process of encrypting the dataset and keyword index after generating symmetric and searchable keys is to first ensure the security of the data by using symmetric encryption methods to input files and symmetric keys, and then encrypt and output the corresponding ciphertext. The searchable encryption methods are used to input keyword indexes, then encrypt and output the corresponding indexes. Finally, attribute-based encryption algorithms are used to input the public key, security index, ciphertext hash value, and access control structure, and then encrypt the security index and ciphertext hash value so that only specific users can decrypt the content, ensuring both security and integrity of the ciphertext.

The encrypted data is uploaded to different locations for storage.

4.3. Data Query

The functional flowchart of data query is implemented, as shown in Figure 7.

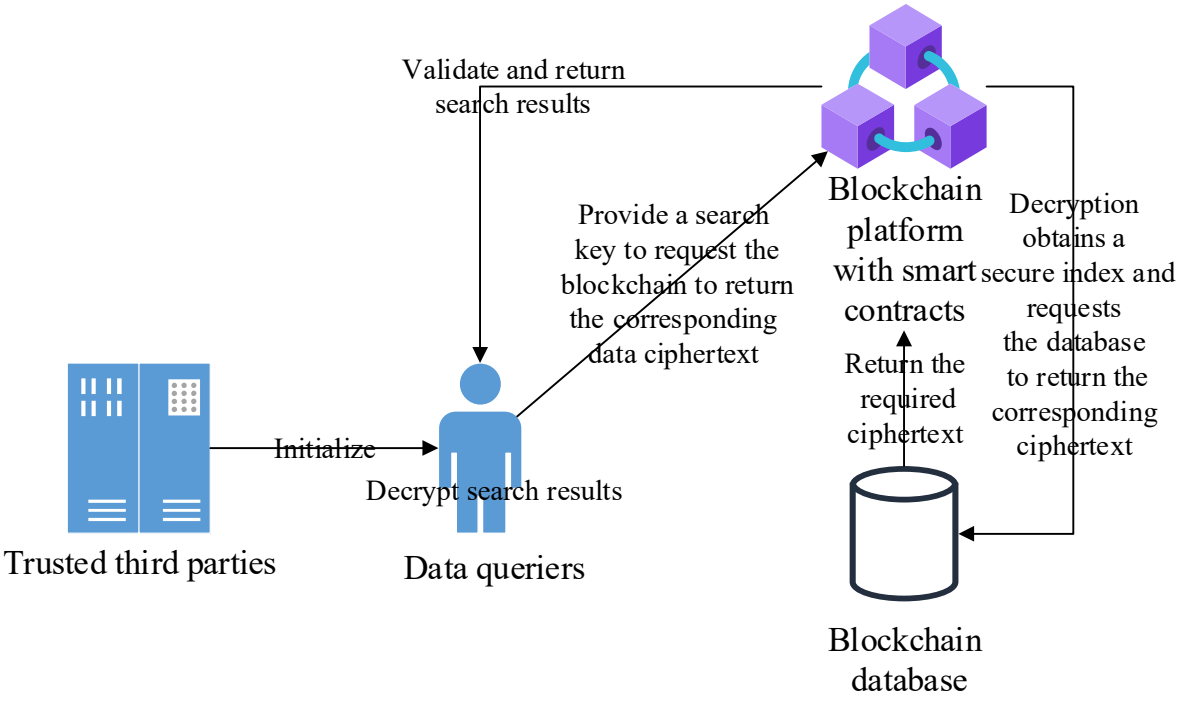

Fig. 7. Data query process

The data query first registers and obtains authorization from a third party, and after authorization, obtains the user attribute private key and security parameters. After data searchers obtain system authorization, they can access data in the blockchain. The properties returned by the system include the encrypted value of the base encryption, the security index value of the base encryption, and the hash value corresponding to the ciphertext. Afterwards, the system can automatically decrypt and obtain the security index and corresponding hash value of the encrypted data. This process is determined by the smart contract to determine whether the user attributes match the access policy and decide whether to proceed with the query process.

After decryption on the blockchain, data queries can obtain secure indexes from the blockchain, and the smart contract can automatically process query service requests. After receiving the query request and security index, the database retrieves the corresponding ciphertext and returns it. When the data query obtains ciphertext from the database, it is necessary to check the integrity of the ciphertext. This check is automatically invoked by the blockchain to verify the ciphertext after obtaining it from the database.

When the ciphertext is confirmed to be complete and not maliciously modified, it is decrypted and plaintext data is returned. After decryption, the plaintext dataset is obtained, and then the required content can be obtained through a smart contract.

4.4. Traceability Function

After the above series of operations, a smart contract can be used to split the string in the prefix field to obtain the unique identifier of the previous node. By using a unique identifier as the new search target for the next round of data queries, after multiple rounds of splitting and querying, a complete traceability query result can be obtained. Sorting according to the timestamp dictionary order can generate a complete traceability chain and complete traceability functions. The traceability process is shown in Figure 8.

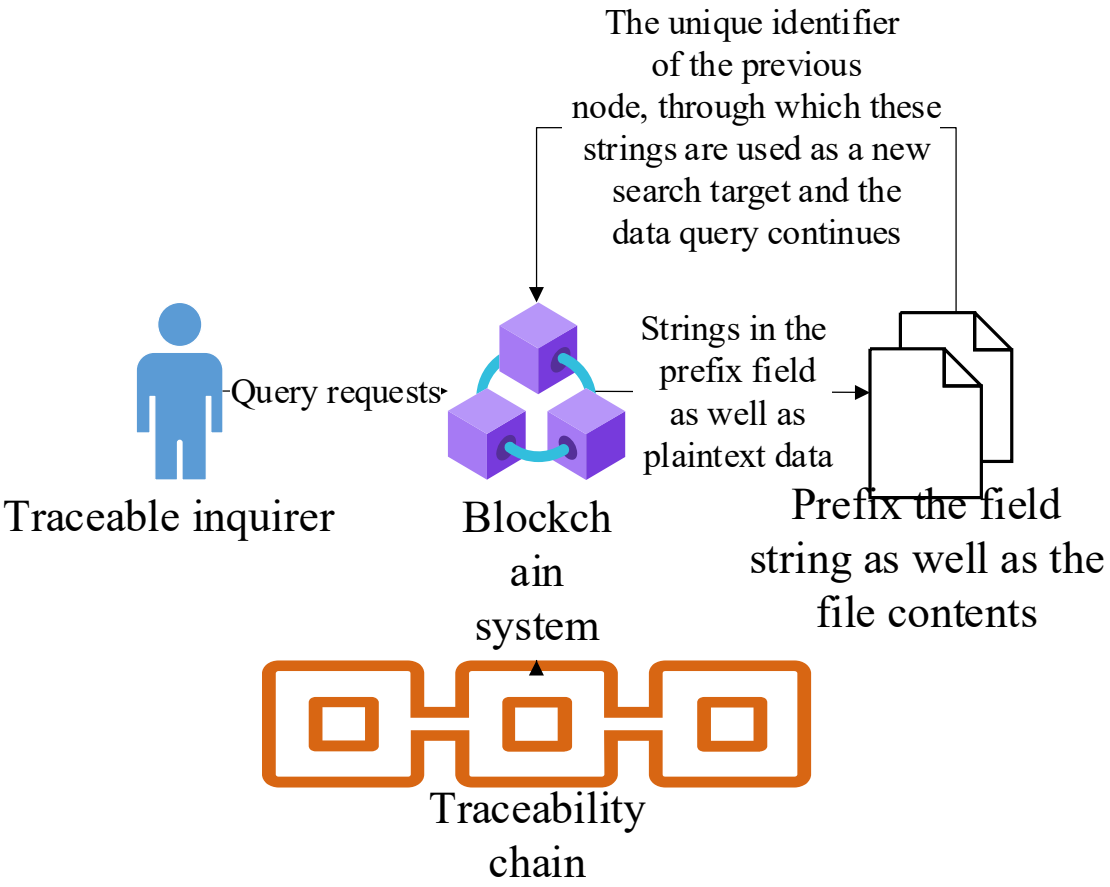

Fig. 8. Traceability process

5. Traceability Speed, Accuracy, and Data Security

Compared to traditional financial document management systems, the secure storage and traceability system for financial documents based on blockchain has strong advantages in terms of accuracy and speed of traceability. The comparison chart of traceability accuracy is shown in Figure 9, and the comparison of traceability speed is shown in Figure 10.

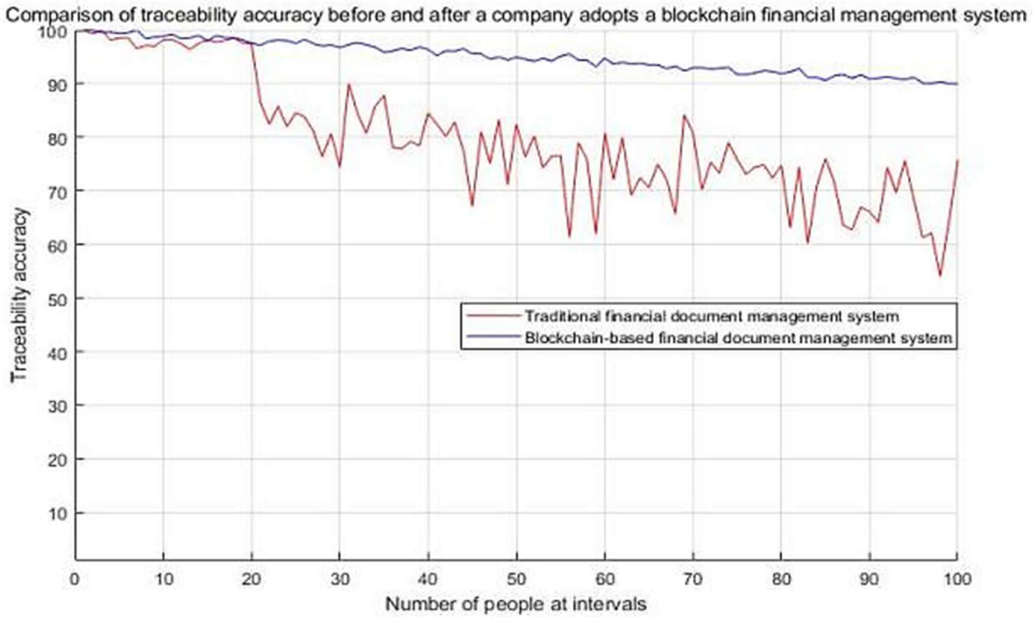

Fig. 9. Traceability accuracy for traditional financial systems and the secure storage and traceability system for financial documents based on blockchain

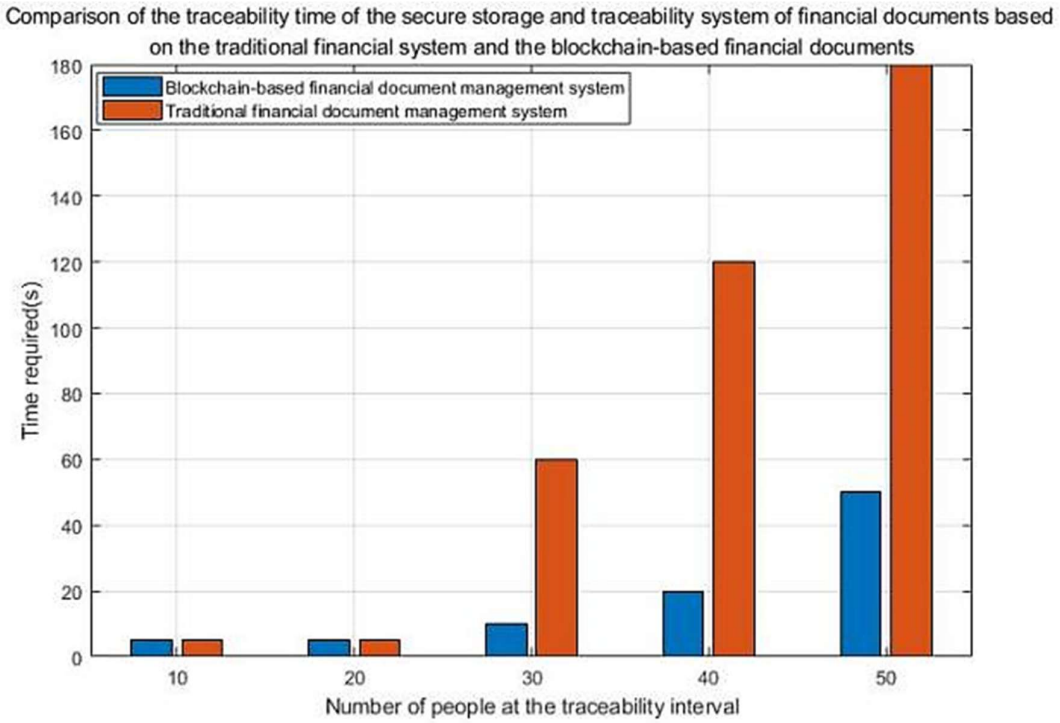

Fig. 10. Time required for traceability of traditional financial systems and the secure storage and traceability system for financial documents based on blockchain

From Figure 9, it can be seen that the accuracy of traceability was not significantly different when the number of people in the interval was within 20. However, after the interval exceeded 20, the accuracy of traceability in traditional systems decreased and became unstable. On the other hand, the accuracy of traceability in the secure storage and traceability system for financial documents based on blockchain was relatively stable and high, consistently maintaining above 90%.

From Figure 10, it can be seen that the traceability speed of both systems was very fast within an

interval of 20 people. However, after exceeding 20 people, the time required for traceability of traditional financial management systems significantly increased, while the time required for secure storage and traceability system of financial documents based on blockchain slowly increased, and the time required for 30, 40, and 50 people intervals was 10 seconds, 20 seconds, and 50 seconds, which was faster than the 60 seconds, 120 seconds, and 180 seconds required for traditional financial systems.

The security of data is crucial in the secure storage and traceability system of financial documents. Firstly, by using different encryption measures in different locations of data storage, attribute-based encryption technology was adopted to ensure the confidentiality of data on the data chain. Only users who meet the attributes can decrypt the information. At the same time, searchable encryption technology and symmetric encryption technology were used to ensure the confidentiality of the data, which can effectively ensure the confidentiality of the data.

In order to ensure the integrity of data, the combination of the feature that blockchain data cannot be tampered with and smart contracts ensures the integrity of database return data and blockchain data.

By controlling different authorization levels to prevent low-level users from accessing data beyond their authority and unauthorized users outside the company from accessing data, the security of data access is ensured and the risk of data leakage is reduced.

By attacking two financial document management systems, the number of attacks and the number of successful attacks were obtained, as shown in Figure 11.

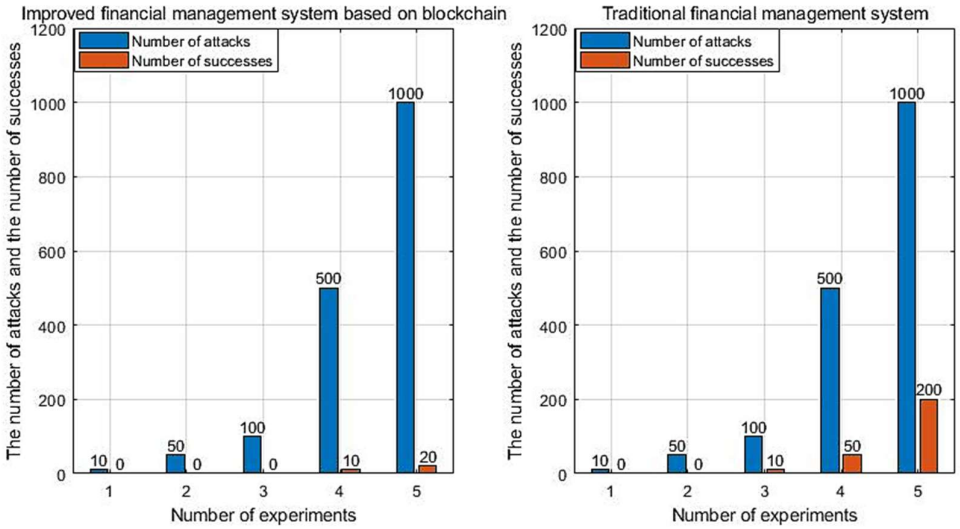

Fig. 11. Comparison of attack frequency and success frequency between traditional financial document management systems and the secure storage and traceability system for financial documents based on blockchain

It can be seen that the security performance of both systems was relatively effective in low-frequency attacks on financial systems. However, under high-frequency attacks, traditional financial document management systems have been successfully invaded 200 times out of 1000 attacks, while the secure storage and traceability system for financial documents based on blockchain have only been successfully invaded 20 times. From this, it can be concluded that the security of the secure storage and traceability system for financial documents based on blockchain was higher than that of traditional financial document management systems.

6. Conclusions

This article used blockchain technology to improve the design of a secure storage and traceability

system for traditional financial documents. Due to the shortcomings of traditional systems, such as low data security, long traceability process, and low accuracy, the design was combined with blockchain technology. By utilizing the decentralized, tamper proof, traceable, and secure features of blockchain, the secure storage of financial documents and the data security of the traceability system were improved, as well as the speed and accuracy of traceability. The application of three-level authorization ensured the security of data access, and multiple encryption mechanisms were used to encrypt information and channels to ensure the security of information. This meets the requirements for secure storage and traceability of financial documents. However, there are still some problems, as the continuous increase of data during the data storage process ultimately leads to the database being unable to accommodate the data or the database being too large, resulting in slower information retrieval.

Funding

This study did not receive any funding in any form.

Data Availability Statement

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Conflicts of Interest

These are no potential competing interests in our paper. And all authors have seen the manuscript and approved to submit to your journal. We confirm that the content of the manuscript has not been published or submitted for publication elsewhere.

References

[1] Akhu‐Zaheya, L., Al‐Maaitah, R., & Bany Hani, S. (2018). Quality of nursing documentation: Paper‐based health records versus electronic‐based health records. Journal of clinical nursing, 27(3-4), e578-e589.

[2] Jervis, M., & Masoodian, M. (2014). How do people attempt to integrate the management of their paper and electronic documents?. Aslib Journal of Information Management, 66(2), 134-155.

[3] Kuchta, T., Lutellier, T., Wong, E., Tan, L., & Cadar, C. (2018). On the correctness of electronic documents: studying, finding, and localizing inconsistency bugs in PDF readers and files. Empirical Software Engineering, 23, 3187-3220.

[4] Baratov, D. K., Aripov, N. M., & Ruziev, D. K. (2019, September). Formalized methods of analysis and synthesis of electronic document management of technical documentation. In 2019 IEEE East-West Design & Test Symposium (EWDTS) (pp. 1-9). IEEE.

[5] Noori, M. (2023). EVALUATION OF ELECTRONIC DOCUMENT MANAGEMENT SYSTEMS AN OVERVIEW. Computer Integrated Manufacturing Systems, 29(5), 203-225.

[6] Großer, K., Riediger, V., & Jürjens, J. (2022). Requirements document relations: A reuse perspective on traceability through standards. Software and Systems Modeling, 21(6), 1-37.

[7] Iqbal, J., Soroya, S. H., & Mahmood, K. (2024). Financial information security behavior in online banking. Information Development, 40(4), 550-565.

[8] Müller, J., Ising, F., Mainka, C., Mladenov, V., Schinzel, S., & Schwenk, J. (2020). Office document security and privacy. In 14th USENIX Workshop on Offensive Technologies (WOOT 20).

[9] Rajbhoj, A., Nistala, P., Kulkarni, V., Soni, S., & Pathan, A. (2022, August). DizSpec: Digitalization of requirements specification documents to automate traceability and impact analysis. In 2022 IEEE 30th International Requirements Engineering Conference (RE) (pp. 243-254). IEEE.

[10] Das, M., Tao, X., Liu, Y., & Cheng, J. C. (2022). A blockchain-based integrated document management framework for construction applications. Automation in Construction, 133, 104001.

[11] Kiu, M. S., Lai, K. W., Chia, F. C., & Wong, P. F. (2024). Blockchain integration into electronic document management (EDM) system in construction common data environment. Smart and Sustainable Built Environment, 13(1), 117-132.

[12] Nizamuddin, N., Salah, K., Azad, M. A., Arshad, J., & Rehman, M. H. (2019). Decentralized document version control using ethereum blockchain and IPFS. Computers & Electrical Engineering, 76, 183-197.

[13] Hofman, D., & Novin, A. (2018). Blocked and chained: Blockchain and the problems of transparency. Proceedings of the Association for Information Science and Technology, 55(1), 171-178.

[14] Zhang, Y., Xu, C., Li, H., Yang, H., & Shen, X. (2019, May). Chronos: Secure and accurate time-stamping scheme for digital files via blockchain. In ICC 2019-2019 IEEE International Conference on Communications (ICC) (pp. 1-6). IEEE.

[15] Bralić, V., Stančić, H., & Stengård, M. (2020). A blockchain approach to digital archiving: digital signature certification chain preservation. Records Management Journal, 30(3), 345-362.

[16] Kang, P., Yang, W., & Zheng, J. (2022). Blockchain private file storage-sharing method based on IPFS. Sensors, 22(14), 5100.

[17] Liu, Y., Zhang, J., Wu, S., & Pathan, M. S. (2021). Research on digital copyright protection based on the hyperledger fabric blockchain network technology. PeerJ Computer Science, 7, e709.