ATLANTIS PRESS
Advances in Intelligent Systems Research, volume 161
International Conference on Transportation & Logistics, Information & Communication, Smart City (TLICSC 2018)

Research and Design of Electronic Evidence Preservation Platform based on Block Chain and Merkle Tree

Zhongyun Jiang
School of information technology, Shanghai Jianqiao University, Shanghai 201306, China.
jianqiao_jzy@163.com

Abstract. After a thorough study of the characteristics of electronic evidence, an electronic evidence preservation scheme based on block chain and Merkle Tree is proposed. In this scheme, a P2P network technology is used to realize the decentralization of the system, an encryption algorithm is used to ensure the security of user data, Merkle tree technology is used to ensure non tampering and traceability of data, an improved consensus algorithm is designed to ensure data consistency between computing nodes, timestamp technology is used to ensure the existence of data in a specific time period or before a certain time, and block chain data structure is used to store electronic evidence. An electronic evidence preservation platform based on block chain and Merkle Tree is designed and constructed. This platform can achieve the decentralized electronic evidence preservation in the cloud computing environment, and prevent any involved participants, including evidence investigators, cloud service providers, users and so on, from colluding to tamper with the digital evidence. Experiments show that the scheme can effectively preserve and verify the electronic evidence, and guarantee the integrity and timeliness of the evidence data.

Keywords: Electronic Evidence; Block Chain; Merkle Tree; Consensus Algorithm.

1. Introduction

With the advent of the era of data technology, both enterprises and individuals will face the problem of data sovereignty maintenance. Especially Internet finance, electronic contract, e-commerce, e-government, network medicine, intellectual property protection and other industries have a strong demand for electronic data preservation. Comparing with the physical evidence in common civil cases, electronic data does not have the physical appearance of the general physical evidence, and its collection, access, submission, display and other ways are quite unique. It is very easy to be objectively modified, annihilated or damaged. Especially in the cloud computing environment, multiple users share distributed heterogeneous virtual computing resources, a variety of temporary data and file access records are generated. It is difficult to ensure that the participants will not tamper with the original information. In view of the above problems, an electronic evidence preservation scheme based on block chain and Merkle Tree is proposed. In this scheme, the electronic evidence is stored on the block chain by the block chain technology principle, and the authenticity, integrity and validity of user data are guaranteed by combining encryption algorithm, digital signature, digital certificate and other technical means. The evidence preservation method based on Merkle Tree and the improved consensus algorithm are designed to realize the decentralization of evidence preservation, and to prevent the collusion of any participants. And the user information and timestamp are bound to the electronic records in the block to realize the verifiability and traceability of electronic evidence in the cloud computing environment.

2. Related Technology

2.1 Electronic Evidence Preservation Technology

Traditional electronic evidence preservation means mainly include printing out, backup, taking photos and videos, and making legal documents [1]. The electronic evidence preservation system designed by Gan Qixian and Song Xiuli uses chaotic block cipher to keep the electronic evidence secret, uses the piece Hashing technology to ensure the integrity of the electronic evidence, and uses the digital signature standard DSS to prevent the communication parties from cheating or denying

each other [2]. Li Xun and Huang Daoli introduced S/MIME signature technology to provide evidence preservation services for the outside world by online receiving electronic evidence, achieve a complete evidence chain through the electronic evidence preservation business process that meets the 17025 standards, and finally by a multi-signature appraisal report to ensure the legal effect of evidence preservation [3]. However, the above-mentioned electronic evidence preservation design mainly improves the security and reliability of electronic evidence preservation from the key technical aspects, does not consider the storage of electronic evidence. Investigators cannot guarantee the integrity and authenticity of the obtained information, especially in the case of dishonesty providers, it is difficult to prevent one or both sides from conspiring to tamper.

2.2 Block Chain Technology

Block chain [4] is a distributed database system participated by nodes, which is decentralized, trustless and collectively maintained. Based on a new decentralized protocol, it can safely store bitcoin transactions or other data (copyright, creditor's rights, equity and other digital assets). The digital information stored in the block chain is not forgery and tampering, and there is no need for any central organization to audit.

Block chain uses decentralized and trustless method to collectively maintain a reliable database technology scheme. The scheme allows multiple nodes of the system to compute and record all the data exchanged during a period of time into a data block by cryptographic algorithm, and generates the unique identifier hash value of the data block for linking the next data block and checking. All participating nodes of the system determine the true record by the execution of consensus algorithm, and maintain consistency of data. The block chain divides the data into several blocks, each of which links to the previous block according to specific information and presents a complete set of data sequentially. The block head of each block contains the hash value of the previous block, which is calculated by hash function on the block head of the previous block. The hash value of each block is connected with the previous block to form a chain.

2.3 Merkle Tree

The Merkle Tree structure [5] is the algorithm created by Ralph Merkle to synchronize data consistency. In this paper, Merkle Tree is used to record the transaction information data in the block. If the leaf node of the tree adds a forged malicious record, the changes caused by the node will lead to the changes of the upper node and the upper level nodes, and eventually lead to the changes of the root node and the block hash. When comparing or verifying in cloud computing environment, Merkle Tree is easy to locate.

3. Electronic Evidence Preservation Scheme based on Block Chain and Merkle Tree

3.1 Electronic Evidence Preservation Model

The electronic evidence preservation model based on block chain and Merkle Tree is shown in Fig.1. The participating traders can act as preservation nodes to provide electronic evidence preservation services to the whole network and maintain the overall electronic evidence database.

The core idea of this model is that in a P2P network composed of n nodes, each node maintains the same evidence chain, that is, the evidence information in the block chain. In this model, the block chain is equivalent to a point-to-point distributed database, which stores the user's security information. It uses an improved consensus algorithm to ensure the consistency of data preservation between nodes, and uses encryption algorithm and electronic signature technology to ensure the security and privacy of user data. At the same time, it uses the timestamp technology, Merkle Tree algorithm and block end-to-end linked chain data structure to form a decentralized, trusted, open, transparent, untouchable, verifiable and traceable electronic evidence preservation platform.

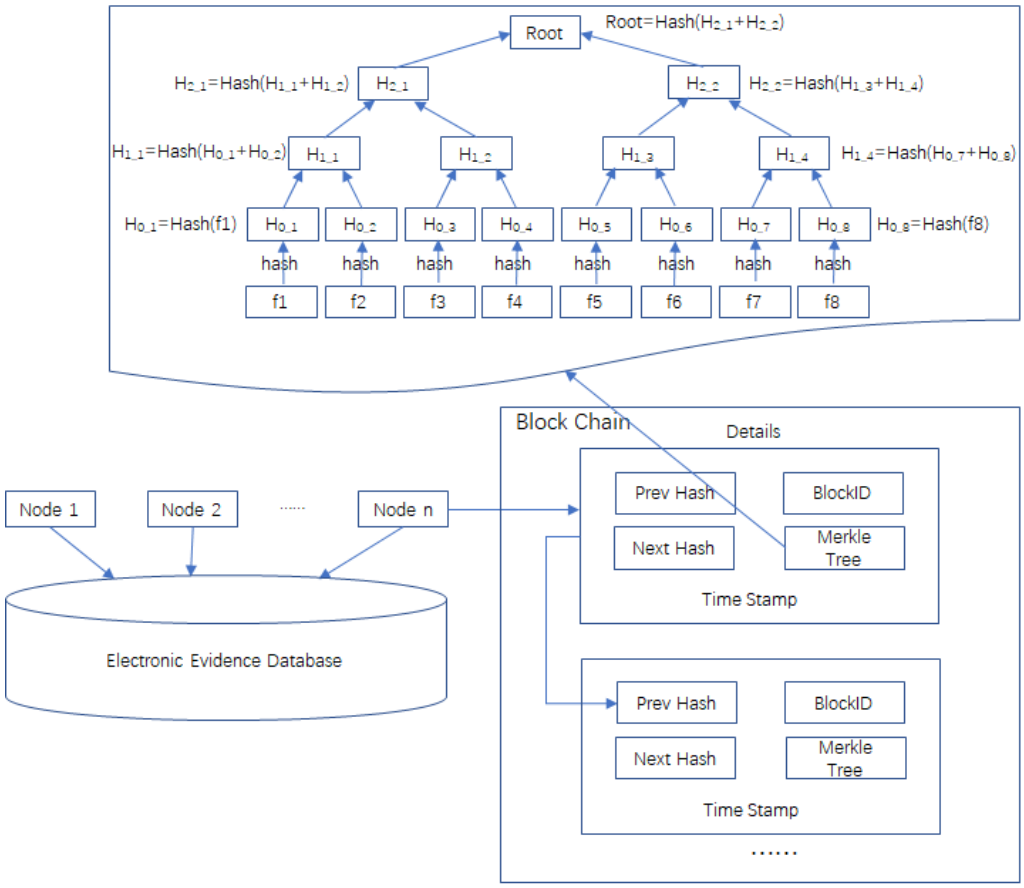

Fig. 1 Electronic evidence preservation model based on block chain and Merkle Tree

3.2 Node Management

In the electronic evidence preservation model, there are two roles of nodes in the network: the management node and the ordinary node, and the messages of the whole network node are sent by broadcasting. The behavior of nodes in the network can be arbitrary: they can join or leave the network at any time, they can discard or forge evidential information, they can stop working, and they may also have various human/non-human faults. The evidence information broadcast in the network may be lost, damaged, retransmitted, delayed and inconsistent with the order of transmission. In order to manage the nodes in the preservation network, the new nodes can join the preservation network only by authenticating the relevant data.

3.3 Evidence Preservation based on Merkle Tree

In the electronic evidence preservation platform based on block chain and Merkle Tree, the smallest unit of block chain is the evidence block, and the evidence block is a kind of data structure formed by the operation of evidential information in a certain period of time. Each evidence blocks records four items: the block ID, the head information of evidence blocks, the details of evidence, the total number of block preservation evidence, and its structure is shown in Table 1.

Table 1. Evidence block structure

Structure Name	Meanings
Block ID	The unique identifier of the current block.
Header information	Record the header information of the current block, and its Hash value is the parameter of the next block.
Details of evidence	Data information stored in the current block.
Total number	The amount of evidence stored in the current block.

The block size, previous block record, Merkle Tree root value, and timestamp are recorded in the evidence block header information, as shown in Table 2. The Merkle Tree root is the Hash value of the root node of all recorded evidence in the current block; the timestamp records the time the block was generated.

Table 2. Evidence block header information

Block Header Structure	Meanings
Block size	The current block size is recorded.
Previous block record	The Hash value of the previous block is recorded.
Merkle Tree root value	The Hash value of the root node of the Merkle of all the recorded evidence in the current block is recorded.
Timestamp	The current generation time of blocks is recorded.

3.4 Consensus of Evidence Blocks

The whole network nodes are composed of management nodes and ordinary nodes, in which ordinary nodes are divided into Speaker nodes and Members nodes according to consensus needs, including the new nodes added in the consensus process. The completion of the evidence block consensus means that evidence has been preserved. In this paper, a Byzantine fault-tolerant algorithm based on block chain is used to improve the consensus algorithm [6]. The flow of evidence block consensus algorithm is shown in Fig.2.

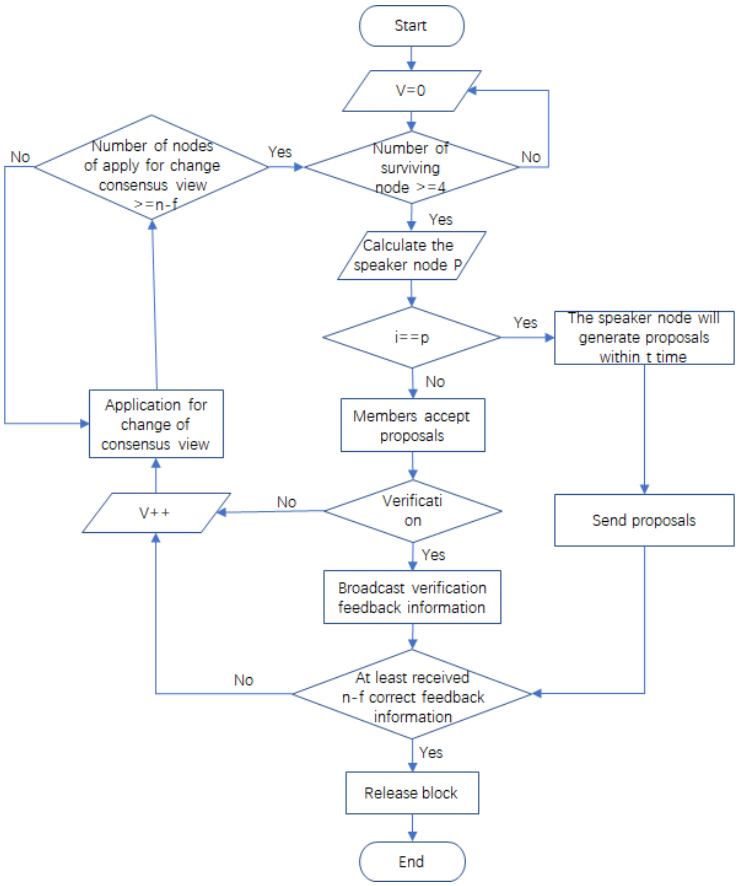

Fig. 2 Evidence block consensus algorithm flow

4. Design of Electronic Evidence Preservation Platform

4.1 Functional Structure Design

Based on block chain and Merkle Tree, users of the electronic evidence preservation platform are divided into enterprise users and individual users. Enterprise users can use the API to access the platform after authorized and authenticated, and complete the preservation of relevant electronic data for enterprises, so as to enhance the credibility and authenticity of enterprises. Individual users can register on the platform, successfully log in, and upload the electronic data information they need to preserve to this platform for preservation.

The electronic evidence preservation platform mainly consists of 8 functional modules. The main functions of the browser include: user registration, user login, real-name authentication, evidence management, etc. The main functions of the server include: authentication audit, certificate issuance, node management, block chain management, etc. The overall functional structure of the platform is shown in Fig.3.

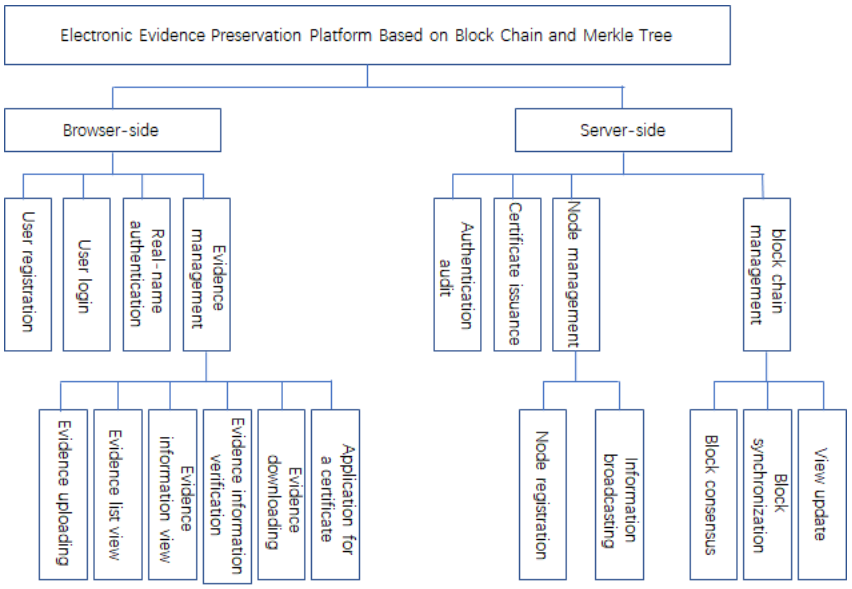

Fig.3 Overall function structure of electronic evidence preservation platform

4.2 Electronic Evidence Preservation Workflow

The main task of the electronic evidence preservation system based on block chain and Merkle Tree is to collect, preserve, store, verify and download the electronic evidence of the enterprise/individual. The main objects involved are system users, system administrators and node networks. Its main workflow includes system user upload, view, verify, download evidence information; network node management and evidence block consensus (evidence preservation), etc. The detailed workflow of evidence preservation is shown in Fig.4.

4.3 Implementation and Analyses

Aiming at the electronic evidence preservation scheme based on block chain and Merkle Tree, an electronic evidence preservation platform is developed and built. The development tool of this platform is Eclipse, the development language is Java, and the development, running and testing server is Tomcat 8.0. The platform uses SSM (Spring, Spring MVC, MyBatis) framework, uses DBCP/JDBC technology to realize the connection with Mysql and XML, and provides JSON data format support for the client. The overall framework provides log records by Log4. Using HTML, CSS, AJAX, Java Script, Jquery and Bootstrap technology to implement the GUI of system browser, the purpose is to provide a user-friendly interface for end users.

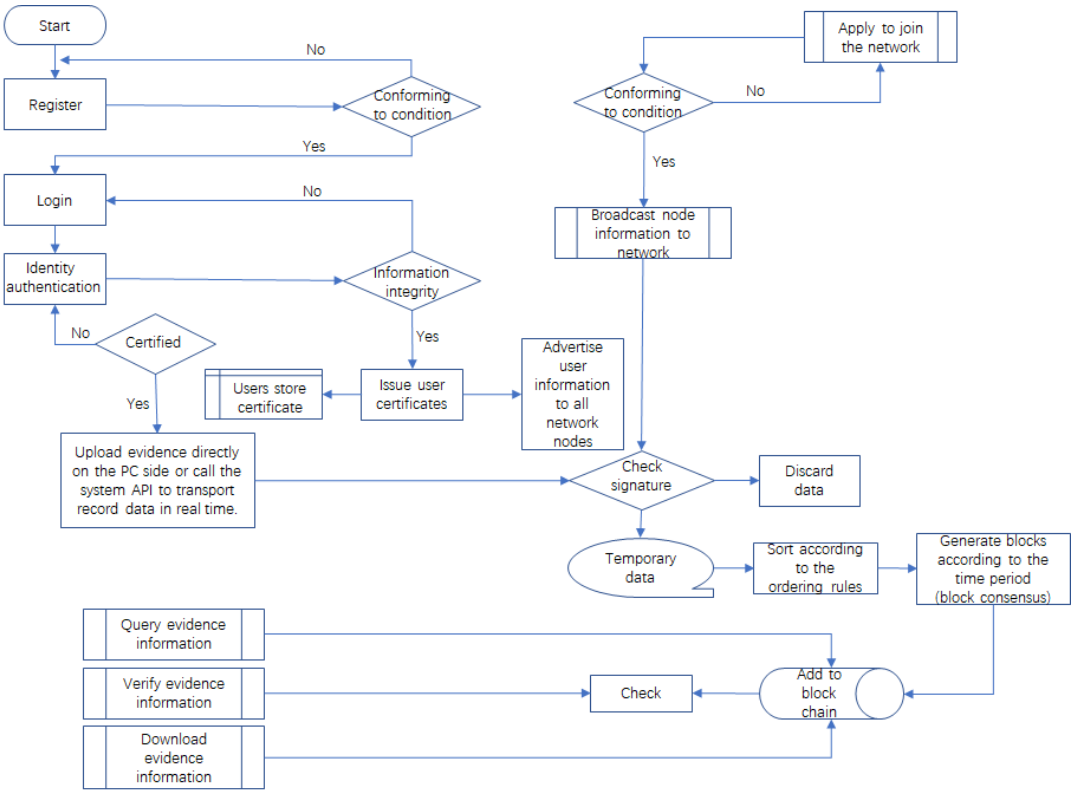

Fig.4 Electronic evidence preservation workflow

The Apache JMeter is used as the performance testing tool of the system. In the test, the block generation time is set to 60 seconds, and the preservation data of different time periods are selected. The electronic data information of the system is added to the block chain to preserve the block within the block generation time and timeout time (2*t). JMeter was used to simulate 200, 400, 600, 800, 1000 virtual users, respectively. User login, user evidence uploads, evidence information view, preservation verification, evidence download and other functions were tested, and test parameters were recorded. For example, preservation verification load test results are shown in Table 3. From the test results, we can see that the platform can respond to user's requests without error, and it is highly efficient, but it needs to be further optimized in a large number of concurrent cases.

Table 3. Evidence block header information

#Sample	Average response time	Minimum response time	Maximum response time	Error%	Throughput
200	923ms	10ms	1856ms	0.00%	111.3/s
400	1605ms	37ms	3389ms	0.00%	123.5/s
600	2289ms	34ms	4098ms	0.00%	161.9/s
800	3431ms	39ms	5338ms	0.00%	150.5/s
1000	4680ms	27ms	7691ms	0.00%	122.3/s
200	923ms	10ms	1856ms	0.00%	111.3/s

5. Summary

In order to avoid conspiracy tamper of the participants and realize decentralized electronic evidence preservation scheme, we propose a blockchain based evidence storage, Merkle Tree based evidence preservation and an improved consensus algorithm to reduce the time of consensus block generation. Experiments are carried out in fewer nodes network environment. Experiments show that the scheme can guarantee the integrity and timeliness of evidence data, but the algorithm efficiency needs to be further optimized to meet the concurrent requests of a large number of users.

References

[1]. Chen Xilin: Research and implementation of electronic evidence preservation platform based on Hadoop (MA, People's Public Security University of China, China 2017). p.2-7.

[2]. Gan Qixian, Song Xiuli. Design and implementation of electronic evidence preservation system. Information security and communication secrecy. Vol.33(2011) No.09, p.102-104.

[3]. Li Xun, Hunag Daoli. Research on a key technology of trusted online electronic evidence preservation. Police Technology. Vol.16(2010) No.04, p.33-36.

[4]. Swan M. Block chain: Blueprint for A New economy. CA.: O'Reilly Media, 2015, p.7-39.

[5]. Liu Tong, Wang Fengying. Solution of Origin Integrity Based on Merkle Tree. Journal of Shandong University of Technology (NATURAL SCIENCE EDITION). Vol.13(2012) No.03, p.68-71.

[6]. Information on: docs.neo.org/zh-cn/basic/consensus/whitepaper.html.