Data Circulation Blockchain Model Construction and Data Access Management

Yang Liu¹, Jincheng Huang²

¹ Department of Information Science and Technology, East China of Political Science and Law, No.555 Longyuan Road, Songjiang District, Shanghai, 201620, China
² Department of Information Engineering, Yancheng Institute of Technology, Hope Avenue Road 1, Yancheng City, Jiangsu Province, 224051, China
lyang@ecupl.edu.cn

Abstract. This paper proposes a controllable Blockchain-based data circulation model. In this model, smart contracts stipulate that data circulation is carried out according to law, and laws and regulations to be followed in data circulation are written into system hardware and software. The supervision and execution of the law is automatically completed by the system. Through the risk assessment of the model, it can be seen that the data circulation Blockchain model has the characteristics of regionalization, security, credibility, collective maintenance, and non-tampering. It is suitable for solving the problem of data sharing and changing the circulation of data. The model promotes the in-depth integration of law and computer technology and promotes the process of legal automation.

1. Introduction
In recent years, big data technology has developed rapidly, and the core is data circulation and sharing. The key to big data applications is to let data flow through, which includes government free data openness and enterprise paid data transactions [1]. However, in reality, all parties have great concerns about the application of big data. The main manifestations are as follows: First, the government is concerned about data security and other issues, so as to be open to the data; second, data transactions are mostly point-to-point, the size of the data meets the demand, but the dimension single. Therefore, the current data circulation is far from the expected results. The main reasons are as follows: First, the circulation mode of the central custody is difficult to meet the data requirements of all aspects; second, the characteristics of the data itself are easy to copy, so that the two parties and the custody center do not trust each other; the third is data security, including leakage of state secrets, corporate information and personal privacy; Fourth, data ownership is difficult to define; Fifth, the lack of relevant laws and regulations. At present, research at home and abroad mainly focuses on issuing policies and regulations to make up for these problems, but the results are not satisfactory.

The Blockchain technology which emerged in the past two years has attracted widespread attention. Its non-tamperable and traceable characteristics are of great significance for solving key problems in data circulation [2-4]. For one thing, the decentralized core advantage of the Blockchain can solve the problem of data circulation from the center to the surrounding divergence mode. For another, the Blockchain uses data encryption, time stamping, distributed consensus and other means to ensure the security, credibility of data circulation and problems such as ownership provide solutions.
2. Data circulation Blockchain model construction

2.1. Blockchain and share authorization certification mechanism
The generalized Blockchain technique is a brand new method of decentralized infrastructure and distributed computing paradigm which uses cryptographic chain block structure to verify and store data, uses distributed node consensus algorithms to generate and update data and uses automated script code (smart contracts) to program and manipulate data [5]. The narrow Blockchain technology is a data structure in which valence data blocks are combined in a chronological order, and cryptographically guaranteed non-tamperable and non-repudiation, attachable decentralized shared general ledger. The basic principles of the Blockchain consist of three main components: An electronic transaction component that records changes in the ledger. Any type of valid transaction information in the electronic transaction component is digitized or encrypted to ensure accuracy and authenticity [6]; Block components, that is, a data structure that stores all transaction information, including a block header and a block body; A chain component, which includes blocks generated in chronological order, and records state changes.

DPOS is known as a more efficient, decentralized, and more flexible consensus mechanism [7]. It allows all shareholder nodes to have voting rights and vote for 101 equity representatives in a fair and democratic way. DPOS can freely re-submit votes based on the performance of the delegates in the subsequent process, effectively reducing the number of participating accounting nodes and achieving rapid consensus verification. The basic working ideas of DPOS are:
- Each shareholder node needs to vote for the trusted representative. The first 101 nodes with the highest number of votes and willing to be represented will take turns to generate new blocks.
- If the authorized representative produces the wrong block or misses signing the new block, it will be replaced by the next representative and may be cast by the shareholder node.
- To be a representative, you need to register a 32-bit unique code on the Internet using the public key, and the header of each record will reference this identification number.
- Each shareholder has a performance indicator to record the behavior of the representative. If you miss the signing of many blocks or sign the wrong block, the indicator suggests to choose another representative.

2.2. Hash algorithm and Merkle tree
The encrypted hash algorithm is a single-item function, which makes it easy to calculate the hash value of the data, but it is difficult to derive the original data based on the hash value. This feature is very important for the Blockchain [8]. The Merkle tree is a data structure based on a hash algorithm. It is characterized in that each non-leaf node is a hash of its leaf node. In a peer-to-peer network, you can use the Merkle tree to verify that the data has been tampered with or that the received data is corrupt. All records generated in the Blockchain are generated by the hashing process of the Merkle tree to generate a unique Merkle root, which is stored at the head of the Blockchain.

2.3. Circulating Blockchain model
Constructing a data circulation Blockchain of data circulation, which is the sum of various types of data formed by physical space or administrative area, forming open, shared and connected within the "local area network", and each data link is connected to form a data circulation network. The system model framework is shown in Figure 1. Confidential data within the department can form a private chain, with dashed boxes indicating sensitive data. The main modules include: supervision audit module, responsible for the authorization and supervision of data circulation on the chain, providing certificate issuance, discussing the trust issues of the parties to the transaction; encryption module, encrypting the data in the chain, discussing the security of the data; distributed ledger Service module: Provides explanation and traceability of transactions, discusses data ownership issues; intelligent contract module, discusses legal and regulatory issues in data circulation.
In the narrow sense, smart contracts stipulate that data circulation is carried out according to law; in a broad sense, laws and regulations to be followed in data circulation are written into system hardware and software, and the system automatically completes legal supervision and execution.

The role of smart contracts in the data flow Blockchain can be divided into two points. First, the smart contract is a contract signed by the parties in the data circulation in accordance with the contract law. It consists of four parts: the contract subject, the contract terms, the arbitration platform and the law enforcement object. Second, the executable code of all specified legal rules in the Blockchain system belongs to the category of smart contracts, ensuring the normal operation of the Blockchain system from the perspective of security, management and efficiency. In a sense, the enforcement of the law is no longer an afterthought, but an event, and can be performed “automatically” without the third parties.

According to the distribution of data resources, the main information services and major data storage nodes are deployed in various government departments or enterprises. Two types of server groups are set up here, namely the department and organization alliance server group and the supervision audit alliance server group. High-level institutional nodes join the departmental and institutional alliance server groups to become agents, and second-level organizations join the regulatory audit server group as audit verification nodes. The model uses the improved Delegate proof of stake (DPOS) implementation. The data storage structure of the model is a hierarchical storage structure, which can be efficiently and conveniently propagated. Finally, the Merkle root of all data is anchored to the Bitcoin Blockchain to achieve true non-tamperability and non-repudiation. In this way, the encryption of the data file is stored in the distributed database system, which reduces the pressure on data storage and high frequency access on the Blockchain.

The data from government departments or major enterprises are basically built in key institutions. On the one hand, small-scale organizations produce less data and less data traffic, and their data center construction is generally not perfect and lacks data processing capabilities. Therefore, the model is designed for the public data and access restriction data. With the improved DPOS consensus mechanism, it can effectively use the current status of each organization to achieve decentralized, secure, fast and traceable data sharing.

In order to ensure that the recorded content is trusted and not falsified, the model records the summary of all the data and stores the data hash value in the Item structure by using the hierarchical structure storage, and then calculates the hash value of each Item into the Item block structure. This can effectively reduce the search space and speed up the user's verification of the record. A data block consists of multiple item blocks. Calculating the hash value at each layer will get the Merkle root of this data block [9], which is done every 1 minute as shown in figure 2. Since the Bitcoin Blockchain uses the Proof of work (PoW) to stabilize a block every 10 minutes, the data is frozen every 10
minutes in the model, represented by the values in the department and institutional server groups. Submit the generated Merkle root into the Bitcoin Blockchain in a form similar to submitting a Bitcoin transaction. Thus we can achieve real tampering [10]. Each Item block stores only the hash value of the Item and a header information, which not only facilitates the propagation of each block in the peer-to-peer network, but also reduces the cost of data verification. Each item can store 10 pieces of data, each of which contains three pieces of information: the data owner public key, metadata, and data summary.

![Figure 2. Merkle tree of the storage structure.](image)

3. Model risk assessment
The data circulation Blockchain model in this paper is based on a distributed network, using a hierarchical data structure and two-tier server federation storage. In this model, once a piece of data is recorded in any of the nodes of the chain of blocks, it is impossible to change again. A Blockchain is a chain of blocks. It is not managed by a central entity or a third-party database. Any government department or enterprise can join. The new block is issued to each node on the network, and each block is verified to ensure that it has not been tampered with. If all the content is verified, then this block is considered valid and can be allowed to join the end of the Blockchain, creating a new Blockchain to ensure that the chain is the longest valid chain. The accepted block is considered to be a random hash prior to the new block. This kind of distributed information preservation mechanism can effectively guarantee the security of information.

The combination of Blockchain technology and Markle algorithm, combined with the DPOS consensus mechanism, can effectively utilize the current state of the organization to achieve decentralized, secure, fast and traceable data sharing, so that the model can greatly reduce the risk of data circulation.

4. Conclusion
With the rapid development of Blockchain technology, the research on Blockchain in the legal field has attracted much attention. The data circulation Blockchain model proposed in this paper can help the transformation of data centers of government departments or enterprises and other institutions to meet the needs. The design of the two-level server group can effectively cooperate with the DPOS mechanism, and the Blockchain intelligent contract technology is used to change the data circulation
mode, which helps to establish a legal system to ensure the development of Blockchain intelligent contract technology, realize data circulation, promote the in-depth integration of law and computer technology, promote the process of legal automation and reduce the risk of data circulation.

Acknowledgments
This research was financially supported by the Ministry of Justice of China Research on the National Rule of Law and Legal Theory (Grant NO. 17SFB3001) and Ministry of Education of China "Clouds Integration Science and Education Innovation" Fund Project (Grant NO. 2017B06106)

References
[1] Kang J, Yu R, Huang X, Maharjan S, Zhang Y and Hossain E 2017 Enabling localized peer-to-peer electricity trading among plug-in hybrid electric vehicles using consortium Blockchains. *IEEE Transactions on Industrial Informatics* **13** 3154-3164
[2] Ma Z, Huang W and Gao H 2018 Secure DRM scheme based on Blockchain with high credibility *Journal of Electronic* **27** 1025-1036
[3] Fairley P 2017 Blockchain world-Feeding the blockchain beast if bitcoin ever does go mainstream, the electricity needed to sustain it will be enormous *IEEE Spectrum* **54** 36-59
[4] Hinarejos M F, Ferrer-Gomila J and Huguet-Rotger L 2019 A solution for secure certified electronic mail using Blockchain as a secure message board *IEEE Access* **7** 31330-31341
[5] Ma M, Shi G and Li F 2019 Privacy-oriented Blockchain-based distributed key management architecture for hierarchical access control in the IoT scenario *IEEE Access* **7** 34045-34059
[6] Li D, Du R, Fu Y and Au M H 2019 Meta-key: a secure data-sharing protocol under Blockchain-based decentralized storage architecture *IEEE Networking Letters* **1** 30-33
[7] Zhou T, Li X and Zhao H 2019 DLattice: a permission-less Blockchain based on DPoS-BA-DAG consensus for data tokenization *IEEE Access* **7** 39273-39287
[8] Liu M, Yu F R, Teng Y, Leung V C M and Song M 2019 Distributed resource allocation in Blockchain-based video streaming systems with mobile edge computing *IEEE Transactions on Wireless Communications* **18** 695-708
[9] Fernández-Caramés T M and Fraga-Lamas P 2018 A review on the use of Blockchain for the internet of things *IEEE Access* **6** 32979-33001
[10] Cai W, Wang Z, Ernst J B, Hong Z, Feng C and Leung V C M 2018 Decentralized applications: the Blockchain-empowered software system *IEEE Access* **6** 53019-53033