System Design of Archives Management Based on Blockchain Technology

Guoliang Li, Su Yan, Longzhou Yuan, Haibing Zhang
China Xiongan Group Digital City Technology Co.Ltd, Hebei, Xiongan, 071700

Abstract: Archive information is an important information resource for the economic and social development in new district. With the construction of the new district, archive information sharing has become a major technical and application problem that needs to be tackled urgently, due to rising archive data amid project process and separated management of them. The traditional system of archives management adopts a centralized architecture and centralized storage. However, different participants have various standards, so as to bring about many problems, such as, difficulties in data interaction, low authority, poor resistance to cyber attacks, and easy information leakage, which can’t meet the needs of archives management modes of personnel, enterprises, projects and other subjects amid the complex construction environment of the new district. With the rise of Bitcoin project, Blockchain technology has popularized. After long-term development, its application ecological environment has begun to form a scale. The concept of a new application system based on the blockchain technology architecture is in coincidence with that of the construction of archives management system of a new district.

1. Introduction
With the continuous development and improvement of the infrastructure and government functions in the new district, the simultaneous construction of digital twin cities has become an important part of the construction of the new district. The social environment of the new district is complex and changeable. Firstly, the composition of the personnel structure is complex, including local aboriginal residents, new district builders, and late settlers. Secondly, the composition of the enterprise is also complex, involving original enterprises, participating enterprises and later settled enterprises in the three counties. Thirdly, the construction projects are complex, including civil engineering projects, municipal projects, landscaping projects, informatization projects, comprehensive service projects, etc. For each individual personnel, individual enterprise, and various project subjects associated within the construction of a digital city, a set of unified data standards and archive management system fully trusted by all participants are required. The traditional archive management system adopts a centralized structure, which can’t effectively meet the actual needs of all participants when facing the constantly changing environment of the new district. Blockchain has the advantages of transparency, credibility, tamper-proof, traceability, privacy security, and high system reliability. Applying it to the construction of the archive management system in the new district can effectively solve the following series of problems encountered by the traditional centralized archive management system. The first one is the authority of data storage. The second one is the consensus mechanism-based system, which has strong resistance to network attacks and excellent stability. The third one is that the data is open and transparent to all parties and can’t be tampered with. The fourth one is based on smart contracts, which can ensure fairness, justice and openness in data usage. It can be seen that building a file management system with blockchain technology as the core is of great significance to digital archive
management in the new district.

2. Overview of Blockchain Technology

In 2009, the Bitcoin system came out and attracted more and more attentions, with increasing popularity. At the same time, the supporting technology behind it—blockchain technology is also well-known to the public. The general public, technology enthusiasts, professional practitioners, and domain experts conduct research and analysis on the blockchain and explore its development direction and application. Local governments actively support the blockchain, and domestic and foreign technology and financial giants have also entered the market. The ecology of blockchain application has begun to form its scale till now. Blockchain is indeed an emerging concept, but its technical foundation is very mature. The technology of hash algorithm, digital signatures, P2P networks, consensus mechanisms, smart contracts were all mature and widely used before the advent of the blockchain.

2.1. Hash Algorithm

A major feature of the blockchain is that the data on the chain can’t be changed, which is mainly guaranteed by the chain structure of hash values.

Hash Algorithm can be concluded that a fixed-length string is generated, if any length of input information (such as a binary file) is given and a certain algorithm (such as hash algorithm) is calculated. This string is called the hash value corresponding to the input under the hash algorithm. For an ideal hash algorithm model, the input sample space is infinite, while the output space is finite. Therefore, the hash algorithm must have a hash collision, that is, two different input mappings produce the same hash value.

Hash algorithms that can be applied to blockchain technology must have the characteristics of collision resistance, brute-force and hiding. The commonly used SHA256 algorithm has these characteristics and is used in the Bitcoin system.

The hash value is used as a pointer and is nested layer by layer. If you want to change with a piece of data on the recognized chain, it will affect the whole calculation and consume a lot of computing resources. At the same time, the hash value constitutes a Merkel tree, which can quickly react to changes in the content on the chain.

2.2. Digital Signature

The encryption algorithm has evolved from symmetric encryption to asymmetric encryption, which is the algorithm principle of digital signature. The digital signature achieves an effect similar to traditional physical signature.

In the network system of blockchain, there are many nodes with different permissions. The blockchain uses digital signatures to achieve permission control and identify whether the identity is legal.

Each node has its own public and private key pair. The public key, as the public identification of the node itself, will be broadcast in the blockchain network and recorded by all nodes.

In the blockchain system, when a node initiates a business logic process, it first encrypts the content with a private key and attaches the signature to the business logic. After other nodes receive the broadcast message, they first verify the attached digital signature after completing the integrity check and the verification of the validity of the sending node's identity, which will trigger the subsequent process.

2.3. Consensus Algorithm

In the blockchain system, all nodes participate in recording, so which node’s record shall prevail? In other words, how to reach a consensus between nodes is a problem that the blockchain system must solve. The traditional centralized system, with the endorsement of the central node, ensures the authority and consistency of the data.
In a decentralized system such as the blockchain, each node's own state and network environment are not the same, with its identity difficult to control. There may be malicious nodes that hinder message delivery or even send wrong information, thereby disturbing the consistency of the overall system. Therefore, the issue of consensus is a very important issue.

The consensus algorithms used in the blockchain are mainly divided into the following categories, that is, proof of work consensus algorithms, proof of work consensus algorithms, Byzantine fault-tolerant algorithms and combined consensus algorithm for trusted execution environment.

2.4. Smart Contract
Blockchain is integrated into various fields, with indispensable smart contracts. Generally speaking, a smart contract is a set of computer programs with a commitment nature defined in digital form. It has constraints on the parties to the contract and will be automatically executed when certain conditions are met.

A blockchain-based intelligence contract meets the needs including transaction processing and before, data storage mechanisms, and complete state machines. And transaction triggering, processing, and data storage must be carried out on the chain.

When the trigger conditions are met, the smart contract will read the corresponding data and perform calculations according to the preset logic, and finally save the calculation results permanently in the chain structure.

The current smart contracts are still lacking in completeness and are more prone to logical loopholes than traditional contracts. The high-level language for writing smart contracts has some "uncertain" instructions in this regard, which may cause system divergence and affect the consistency of the system.

3. Overall System Architecture Design
Based on the user's public key, the identity cloud is constructed, the user's registration, deregistration and special user will all go to the chain through consensus, so as to ensure that the identity can't be tampered with. The user's private key, verification logic algorithm, and archive information management operations based on smart contracts will exist in the terminal in the form of a lightweight service, with the identity cloud Hash to verify whether the operation conforms to the rules.

![System construction drawing](image)
4. The Convergence between Blockchain Technology and Archive Management System

4.1 Decentralization - the Promotion of Data Sharing and Integration
Decentralization means that each node in the archive management system is basically equal, in terms of function. Archive data can run and be stored independently on each node, no longer relying on additional centralized management systems. It can meet certain performance requirements via basic hardware facilities.

The multi-center governance of social subjects, citizens, and archives departments has changed the central position of archives departments in the traditional management system and shaped the benign interaction between the three. Then, peer-to-peer technology gets rid of the constraints of third-party platforms and promotes more convenient circulation and sharing of archive data. As an independent node, users can independently get access to and operate archive data on the blockchain, which is conducive to forming a pattern of diversified governance of archive data.

4.2 Stamp Encryption to Reduce Security Risks
To ensure the security of archive data, one is not to be leaked, and the other is not to be tampered with, which is the top priority in the entire archive management. Combined with the blockchain technology, the data is stamped and encrypted. On the one hand, it is identified by a timestamp, which is unique. When the block is packaged, a timestamp will be added to the header of the block to record the time when the data entered the block. On the other hand, in the blockchain, when the node packs the block data, it uses asymmetric encryption technology based on the hash algorithm to encrypt it, and the archive operation permissions are precisely divided and not easily tampered, which improves the security protection capability of the system.

In the big data environment, there exit problems, such as leakage, attack and abuse of data. The application of blockchain technology and the archive management system has built a safe sharing channel.

![Figure 2. Business model](image-url)
4.3 Multi-chain Cooperation to Promote Open and Transparent Data
There are many types of blockchains, which can be divided into public chains, private chains and consortium chains according to the degree of openness. The public chain is open to everyone, with a large number of nodes. Each node can participate in the recording and interaction of data on the chain. The alliance chain, as the name suggests, is open to members of a certain alliance or a group, while the private chain has the lowest degree of openness, with only open to a certain organization or individual.

Blockchain technology is applied to the archive management system. According to the different important levels of archive data, a multi-chain combination model can be adopted. The archive data disclosed by all members of the society can be managed by the public chain system, with the participation of the whole people, and the governance by the whole people, while it is not suitable for implementation at present. Cross-regional and cross-departmental archive data can be managed in a consortium chain, which is also a relatively mature application at present. Archive data involving personal privacy is managed by private chain.

5. Conclusion
The application of blockchain technology to the archive management system can achieve diversified management of it. Nowadays, the amount of data has surged and traditional centralized systems are faced with increasing difficulties in challenges of data governance. The development and application of blockchain technology can optimize the existed data environment and realize intelligence and modernization.

References
[1] Ante L., Steinmetz F. & Fiedler I.. (2021). Blockchain and energy: A bibliometric analysis and review. Renewable and Sustainable Energy Reviews(). doi:.
[2] Serrano Will. (2021). The Blockchain Random Neural Network for cybersecurity IoT and 5G infrastructure in Smart Cities. Journal of Network and Computer Applications(). doi:.
[3] Bouraga Sarah. (2021). A taxonomy of blockchain consensus protocols: A survey and classification framework. Expert Systems With Applications(). doi:.
[4] Khorasany Mohsen, Dorri Ali, Razzaghi Reza & Jurdak Raja. (2021). Lightweight blockchain framework for location-aware peer-to-peer energy trading. International Journal of Electrical Power and Energy Systems(). doi:.
[5] Stefanella Stranieri, Federica Riccardi, Miranda P.M. Meuwissen & Claudio Soregaroli. (2021). Exploring the impact of blockchain on the performance of agri-food supply chains. Food Control(prepublish). doi:.
[6] Niknejad Naghme, Ismail Waidah, Bahari Mahadi, Hendradi Rimuljo & Salleh Ahmad Zaki. (2021). Mapping the research trends on blockchain technology in food and agriculture industry: A bibliometric analysis. Environmental Technology & Innovation(). doi:.
[7] Pincheira Miguel, Vecchio Massimo, Giaffreda Raffaele & Kanhere Salil S.. (2021). Cost-effective IoT devices as trustworthy data sources for a blockchain-based water management system in precision agriculture. Computers and Electronics in Agriculture(). doi:.
[8] Carvalho Arthur, Merhout Jeffrey W., Kadiyala Yogesh & Bentley II John. (2021). When good blocks go bad: Managing unwanted blockchain data. International Journal of Information Management(). doi:.
[9] Xu Xiaojiong, Sun Gang, Luo Long, Cao Huilong, Yu Hongfang & Vasilakos Athanasios V.. (2021). Latency performance modeling and analysis for hyperledger fabric blockchain network. Information Processing and Management(1). doi:.
[10] Wang Jian, Liu Yongxin, Niu Shuteng & Song Houbing. (2021). Lightweight blockchain assisted secure roaming of swarm UAS networking. Computer Communications(). doi:.
[11] Hua Weiqi, Jiang Jing, Sun Hongjian & Wu Jianzhong. (2020). A blockchain based peer-to-peer trading framework integrating energy and carbon markets. Applied Energy(). doi:.