Exploration of Block Chain Technology in Data Security Field

Qiushuang Yan

Kunming University of Science and Technology, Kunming, China

Keywords: block chain technology; data security; credit certification system

Abstract: In order to explore the application of block chain technology in data security field, first of all, the definition of block chain technology is introduced. Secondly, related concepts of block chain technology, including block chain support technology, are described. Thirdly, taking the credit certification system as an example, the application of block chain technology in data security in education field is discussed. The system development process is introduced, the network topology is described, and the system process is designed. Finally, the system is implemented to test the operation status of the system, so as to ensure the data security of the system. The research results show that, with the help of block chain technology, the data security can be well guaranteed. It can be seen that the block chain technology is effective.

1. Introduction

The development of information technology has brought tremendous impact on people's life style, especially prominent in the aspects of data dissemination and information sharing. According to the estimation of the Ministry of Industry and Information Technology, the number of Internet users in China reached 614 million by 2014, and the Internet penetration rate reached 47.4% [1]. the data provided by the Ministry of Industry and Information Technology showed that, by the end of 2013, the total number of smartphones in China has reached 580 million, and the average domestic users spend 46% of their daily time working, studying and entertaining on the Internet [2]. The Internet has become the main way for people to obtain information and various communication methods are presented in front of people. Real-time and non-real-time communication tools emerge in endlessly. These communication methods have brought tremendous impact on the learning and communication. With the sharing and popularization of the Internet, the influence of the Internet has gradually increased, especially in the aspects of thinking and culture. The application of information technology in education makes the learning environment, learning methods, learning exchanges and other aspects change with the development of information technology. Network support for learning is not only static support for resources and media, but also highlights the support for interaction between learners.

At present, the application scope of block chain is expanding, and there are application scenarios in many important fields. In the military field, block chain has been widely used, in which the U.S. military uses the advantages of block chain in transaction anonymity and privacy, and applies block chain technology to its intelligence collection work. For the corresponding intelligence, the anonymous privacy payment of password currency is used [3]. And the U. S. military plans to use block chain technology to achieve the program design, frequency of use, effect analysis and other digital information records and analysis of nuclear weapons from the project demonstration, development and production, delivery to service and retirement [4]. Moreover, the attention, investment and research and development of block chain technology have reached an unprecedented level in society and academia. Educational practitioners and scholars begin to think about how to introduce block chain technology into education industry to further improve and reform the infrastructure of education information, and solve the problem that education management relies heavily on centralized management.

Based on the above background, firstly, the definition of block chain technology is introduced. Secondly, related concepts of block chain technology, including block chain support technology, are described. Thirdly, taking the credit certification system as an example, the application of block

chain technology in data security in education field is discussed. The system development process is introduced, the network topology is described, and the system process is designed. Finally, the system is implemented to test the operation status of the system, so as to ensure the data security of the system.

2. Block chain technology

Block chain is a special data structure which consists of data blocks according to the logical order. Through encryption algorithm, it ensures that the internal transaction information cannot be tampered with or forged. Because of its logical appearance of sequential connection, the block chain is displayed with a chain structure. Block chain abandons conventional key-value access to data, but uses chain structure and verifies block data through computation [5]. Blocks are generated by consensus and voting mechanism and new blocks are loaded. Encryption algorithms are used to improve the reliability of message transmission. Intelligent contracts formed by digitized code are subversive distributed control architecture that can manage data. It can store the virtual currency transaction records or other interactive data completely, and the related data cannot be forged and modified. Block chain technology solves the Byzantine General problem, greatly reduces the trust cost and accounting cost of the real economy, and redefines the property rights system in the Internet era.

2.1 Block chain

Block chain is a new technology system combining various technologies. The earliest definition came from the bitcoin paper published in 2009 [6]. Block chaining is a distributed account shared by each node of a distributed system. Each distributed node uses a specific hashing algorithm and a Merkle tree data structure to encapsulate the transaction data and code received over a period of time into a block and link to the longest block chain. Block chain has the characteristics of de-centralization and de-trust, which can establish trust transfer between peers without relying on third-party trusted institutions and help to reduce transaction costs and improve transaction efficiency.

The bottom of block chain is P2P (peer-to-peer) network, which mainly completes related network communication and related interactive tasks. On the basis of P2P network, the cryptographic ledger technology is used to form a distributed ledger. The distributed ledger here is fundamentally different from the existing bank ledger. In the existing bank ledger, the essence of electronic money, whether coins, banknotes, bank cards or Alipay, is a series of numbers, which is managed by the central bank [7]. In the block chain, the elements in the distributed ledger are code. To manipulate the elements in the distributed ledger, it is necessary to provide the relevant private key to complete a series of column programming operations. Generally speaking, the block chain system consists of data layer, network layer, consensus layer, incentive layer, contract layer and application layer.

The data layer includes block, timestamp, Hash function, Merkel and other data organization methods of block chain, as well as asymmetric encryption algorithm. Network layer includes P2P network, data dissemination mechanism and data validation mechanism; consensus layer includes consensus mechanism in block chain, and typical consensus mechanism mainly includes proof of workload (PoW), proof of stake (PoS), and delegated proof of stake (DPOS) [8]; incentive layer mainly includes block chain issuance mechanism and distribution mechanism; contract layer mainly includes scripting language, algorithm mechanism and consensus contract, in which intelligent contract is the core; application layer includes application scenarios of block chain, represented by programmable money, programmable finance and programmable society.

2.2 Block chain support technology

First, P2P network: The role of P2P network in block chain is to connect all nodes so that any pair of nodes can establish an interconnected communication without relying on a third party, and use the broadcast to transmit data information so that the system can run normally. In P2P network, there are the following two key concept:

Broadcasting mechanism. In P2P network, there is no centralized special node and hierarchical structure in theory. Each node will bear the task of network routing, verifying information, disseminating information, discovering new nodes and so on. Block chains publish information by broadcasting, and the transaction information generated in the bitcoin is broadcast to all nodes. In the broadcast process, the node will verify whether the message is legal, and then decide whether to broadcast it to the adjacent nodes. As long as the transaction information is guaranteed to be received by more than 51% of the nodes, the transaction is qualified and can be recorded in the new block [9]. If the node determines that the transaction information is incorrect, it discards the information and terminates the broadcast operation. Besides the verification of transaction information, the broadcast mechanism of P2P is also used to confirm the bookkeeping right of nodes. The nodes with dynamic random numbers and full records broadcast the new block data to the whole network. The other nodes abandon their own blocks, receive the block data obtained by broadcasting, and store the new block in the block chain after verifying the bookkeeping right.

Consensus mechanism. There is no centralized centralized management system in block chain network, and all the accounting nodes play a key role. How to form a simple, easy-to-use, low-cost and manageable storage system is a problem to be solved in block chain technology. Whether data consistency and data availability can be guaranteed in a data sharing network system are two factors to be chosen in system construction. Bitcoin adopts PoW consensus mechanism to ensure data consistency.

Second, encryption algorithm: Block chain technology is based on the principle of cryptography, so that any two nodes can complete communication, and solve the channel credit problem. In the field of traditional audio-visual education, if the school-running institutions and users want to establish secure communication, the introduction of security certification mechanism is needed: the third party needs to establish a secure channel to ensure the safety of communication lines and monitor the behaviour of both sides, to prevent the learners' personal illegal operation and the organization's information modification. Such certification institutions are usually natural or legal persons who sign electronic certificates and provide electronic signatures. They are responsible for keeping public keys and issuing electronic signatures so that new certificates can be verified by the institutions. In contrast, the point-to-point data transmission of block chain solves the problem of tripartite cost, and the data transmission range is only between nodes. The trust and authentication between unknown nodes are implemented by encryption algorithm.

Third, Hash algorithm: Secure hash algorithm (SHA) is a commonly used data encryption algorithm, published by the National Institution of Standards and Technology (NIST) in 1993 as the Federal Information Processing Standard (the first generation SHA algorithm SHA-0). In 1995, its improved version, SHA-1, was also officially released. SHA algorithm is the most commonly used secure hash algorithm and the most advanced encryption technology. The general idea of hash algorithm is to receive a plaintext, then transform it into a (usually smaller) cipher-text in an irreversible way, and convert them into a shorter and fixed-digit output sequence, known as hash value (called information summary). The algorithm generates a 160-bit message digest output for messages with a length not exceeding 264, and the input is processed in 512-bit packets.

3. System design and implementation

Block chain technology is widely used in education, finance, e-government, energy applications, medical and other fields. Taking block chain technology in the field of education as an example, the performance of block chain technology is studied. Taking credit certification system as the research object, the performance of block chain technology in ensuring the security of educational data is discussed.

3.1 System design and development

The subject of credit certification system is the learner node, and with the corresponding credit generation rules, through the realization of transfer including learning record information, it builds P2P network. The process involves the block generation, recording rules, data signature, sending

authentication, data recording and other functional modules. The design patterns follow the basic design principles of the network system, striving to make the data structure clear, transmission process smooth, low learning records generated delayed, and learning certificate traceable. The incremental development process is adopted, which is different from the linear structure of the traditional design pattern of software. The modification and iteration of system performance, logic, and interface modification accompanies with every development stage. After every stage of the research is completed, the current system is evaluated and analyzed, and through reference to traditional software development process, the block chain credit system development process is made.

The system development process is divided into five main steps: requirement analysis, data model design, system design, system implementation and system testing. Here, the system design and system implementation are mainly introduced.

System design: According to the requirements and data structure, the system is designed in detail. The design of system front-end framework, database model, each functional module, and communication interaction model specifies a complete strategy for the system development.

System implementation: In accordance with the functional module and system architecture, the system is developed and implemented. By using related technologies and languages, the system is constructed. Then, debug the system, solve vulnerabilities, and improve system performance.

The design process has clear thinking and clear division of labour in all stages, and specific research plans. Interaction between the stages of the development process is conducive to adjustments and improvements in the process of system development, and the structural distinctions between them are obvious but closely linked.

3.2 Network topology design

Compared with the design of distributed hash P2P network in Bitcoin system, the block chain credit certification system uses the P2P network design of index server, specifically uses the server to establish addressing routing, and links a certain number of network nodes. Each routing forms a self-made system management channel and connects the nodes to the P2P network. The detailed design of P2P network is introduced below. The three common topologies of P2P networks are vertical topology, ring-shaped topology and star-shaped topology. The P2P network in bitcoin can be regarded as a server free star-shaped topology. In star-shaped topology, it is not a central server that connects each other's nodes, but a common node, thus realizing a de-centralized system and making the nodes equal to each other. Figure 1 shows the star-shaped network topology, and Figure 2 shows the P2P network structure in the bitcoin. The difference between the two graphs can be clearly distinguished from the following two figures.

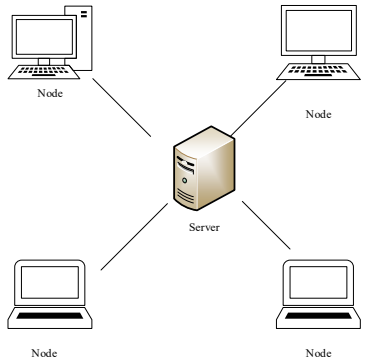

Figure 1. Star-shaped network topology

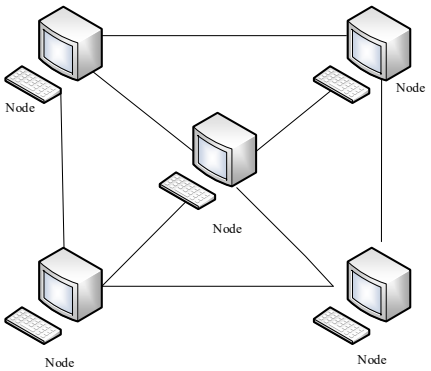

Figure 2. The P2P network structure in the bitcoin

Comparing the two figures above, it can be seen that the nodes in the star-shaped network are connected by servers. The advantage is that the network structure is simple and easy to centralized management. Bitcoin network topology is peer-to-peer, and the advantage is that all users are on an equal footing and there is no server. After joining the network, it downloads the history block chain automatically, not allocated or managed by the server. In addition, it is efficient in avoiding server

attacks caused by network failures, but also a great extent, realizing complete de-centralized design and ensuring normal communication between peer-to-peer networks.

The network topology of credit certification system adopts index network topology structure that is, deploying groups within the node cluster to divide and conquer the nodes. The server of the block chain system does not participate in the actual control of the node, and does not participate in the centralized management or supervision of the third party in the block chain network. The purpose of multi-group branching in the network is to verify the partial pressure of blocks, trust some samples and abstract a cluster of nodes as a whole to share the transmission pressure of block checking. Another advantage of this design of node dispersion management is that it does not need to worry about the location overhead caused by the inter-node transmission, and the relatively integrated management of the group is orderly and efficient. Even if a transit node has system errors and other problems, the nodes can be logically connected to other nodes, so the communication between the entire network nodes can be reliably guaranteed.

3.3 System process design

For user systems, learners can establish real-time communication connections through user login when using the system. When any learner logs in, the user will be activated in the server status. If someone logs in with the user status in other browsers, the user will be refused until the first user closes the page or exits the login status. The system checks whether the user has block chain data, returns NULL if it exists, and dispatches the original block if it does not. When the users are in the login state, they can learn. After the learning event occurs, the system generates corresponding data and records nearly a new block according to the corresponding calculation rules.

User identity checking: The main visible operation that users can do with this system is learning operation. As a new user, they will obtain the privileges as nodes in block chain network through a series of form submission and database operation. As an existing user, the system will enable parameter echoes to see if the user's public key is available. When the user's private key can normally decrypt the information and the identity check information in the central database is filled in correctly, the user's identity will be confirmed by the system. If the system only determines that the username and password are correct and the key is leaked, the user can only learn and cannot participate in the block generation, information recording, block broadcast, block storage and other operations as a normal node in the system. The flow chart of identity verification is shown in Figure 3.

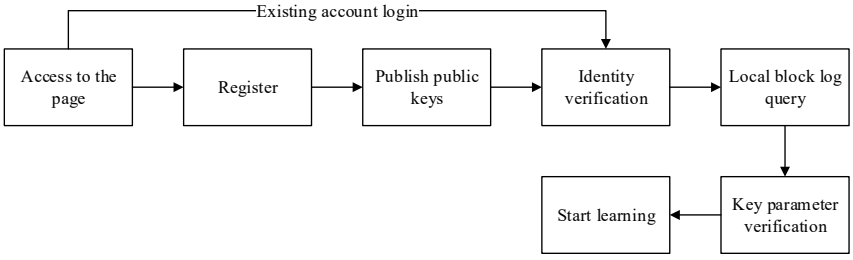

Figure 3. Identity verification process

Resource management: static resource management and script files constitute the interaction architecture of the front-end of the system. In order to consider compatibility and system stability when applying for resources, it is necessary to make corresponding matching strategies for the devices and browsers used by the client. Dynamic scripts also need to follow the adaptation principle to perform asynchronous loading. The related work flow chart is shown in Figure 4.

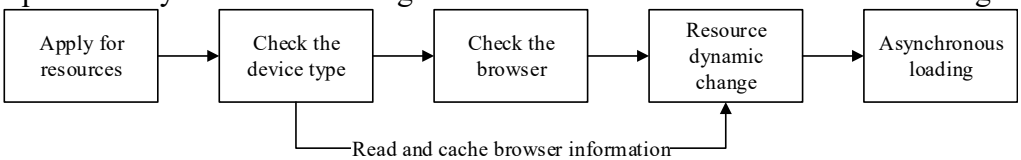

Figure 4. Resource management process

3.4 System implementation

The operating system is Win10, the database is mongodb, and the node version is 8.4.0. After compiling the program code, building the environment, and running as a whole, the system running state needs to be further checked. For the block chain system, first open the system home page to verify the normal presentation of the system as a learner. At the command line, input C:\Program Files\ MongoDB\ Server\3.4\bin>mongod.exe-dbpathc:\mongodata\db, open the local mongodb database for manager and user identity checking. In practical stage, it is necessary to input the extension command-auth for database management authorization verification.

The system checks the manager identity and initializes the corresponding hash ID in the process of construction, which proves that the server works smoothly and keeps running normally. In order to ensure the normal data extraction of the block chain system, the design strategy is specified to ensure that the system will automatically create a virtual node during the system construction process, so it can be seen that the number of users in the figure above exists first and only. When a user uses it, there are multiple nodes in the network, and the virtual nodes are released. The virtual nodes will not be rebuilt until the monitoring system cannot detect any other user nodes.

4. Conclusion

Block chain technology, as a distributed public account, is changing the development concepts and models of current global financial, commercial, public management and education. Therefore, major banks, stock exchanges, governments and educational institutions around the world have invested heavily in the development and application of block chain technology. Due to the security and transparency of block chain storage information, in addition to the virtual currency transactions, block chain in education and other fields are also gradually valued by the government and educational institutions. Building an information management system based on block chain to reduce trust has become the research focus of technological innovation and model innovation in various fields. The definition of block chain is firstly introduced, then the related concepts are described. Secondly, taking the credit certification system as an example, the system design and implementation are discussed. The results show that the credit certification system runs well aided by block chain technology. It is concluded that block chain technology is effective in protecting data security.

References

[1] Yli-Huumo J, Ko D, Choi S, et al. Where is current research on blockchain technology?—a systematic review. PloS one, 2016, 11(10), pp. e0163477.

[2] Peters G W, Panayi E. Understanding modern banking ledgers through blockchain technologies: Future of transaction processing and smart contracts on the internet of money. Banking Beyond Banks and Money. Springer, Cham, 2016, pp. 239-278.

[3] Apte S, Petrovsky N. Will blockchain technology revolutionize excipient supply chain management?. Journal of Excipients and Food Chemicals, 2016, 7(3), pp. 910.

[4] Brandon D. The blockchain: The future of business information systems. International Journal of the Academic Business World, 2016, 10(2), pp. 33-40.

[5] Park J H, Park J H. Blockchain security in cloud computing: Use cases, challenges, and solutions. Symmetry, 2017, 9(8), pp. 164.

[6] Tai X, Sun H, Guo Q. Electricity transactions and congestion management based on block chain in energy Internet. Power Syst. Technol, 2016, 40, pp. 3630-3638.

[7] Maurer B. Re-risking in realtime. On possible futures for finance after the blockchain. BEHEMOTH-A Journal on Civilisation, 2016, 9(2), pp. 82-96.

[8] Kupriyanovsky V, Sinyagov S, Klimov A, et al. Digital supply chains and blockchain-based technologies in a shared economy. International Journal of Open Information Technologies, 2017, 5(8), pp. 80-95.

[9] Buitenhek M. Understanding and applying Blockchain technology in banking: Evolution or revolution?. Journal of Digital Banking, 2016, 1(2), pp. 111-119.