

# The Analysis of Big Data Distributed Processing Technology Based on Blockchain

Shuiquan Guo

Wuhan City Polytechnic, Wuhan, China

**Keywords:** Blockchain, Big Data, Distributed, Processing Technology

**Abstract:** Blockchain is a New Application Mode for Data Encryption Processing. Its Security, Independence and Other Characteristics Make It Widely Concerned and Integrated in Many Fields. in the Era of Big Data, the Protection and Processing of the Vast Amount of Data Information That Can Be Utilized Has Relatively High Security Requirements for Related Technologies. the Big Data Distributed Processing Technology under the Blockchain Can Greatly Improve the Security Performance of Data Information. Promoting the Development of Other Industry Sectors is Also Very Important.

## 1. Introduction

With the Passage of Time and with the Continuous Reform and Innovation of the Times, in Terms of China's Current Development, Whether It is China's Social and Economic Development, or the Improvement of Comprehensive National Strength, Compared with the Development of the Past, It Has Already Been Realized. a Very Important External Manifestation is That Domestic Information Technology Has Become Very Developed, and One of Them is the Network Construction of Distributed Collaborative Communication. through the Blockchain, This Technology Can Determine from the Basic Point That Digital Objects Can Be Applied for a Long Time. and It is Also a Very Effective Means of Implementing Format Management. from This We Can See That the Development of This Technology is Very Important and Can Contribute to the Current Social Development [1].

## 2. The Concept of Blockchain and Its Characteristics

### 2.1 Blockchain Concept

Blockchain is a new application mode of electronic computer technology such as distributed data storage, point-to-point transmission, consensus mechanism and encryption algorithm. It realizes a certain mathematical algorithm by realizing the trust relationship established between different nodes in the computer system, such shown as Fig.1. The information security of the blockchain is greatly improved. Therefore, the blockchain can be understood as a data structure in which data blocks are connected to each other in a chronological order to form a chain, and then a cryptographic method is used to ensure that the data cannot be falsified or forged [1]. The data operation in the blockchain comprehensively utilizes various advanced electronic computer technologies, which acquire a large amount of available information through the network communication technology to perform encryption operations, which is called "mining". The difficulty is high and the confidentiality is good. In the era of continuous development and progress of computer networks, the application scope will become wider and wider.

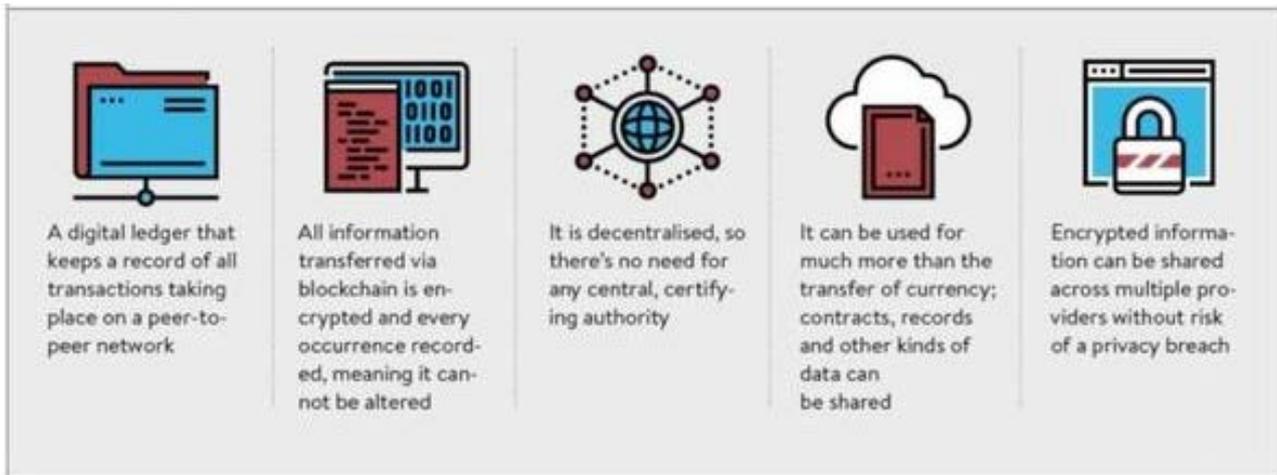


Fig.1 The Meaning of the Blockchain

## 2.2 Blockchain Characteristics

The birth time of the blockchain and the time of attention are relatively high, but in terms of the current development situation, there is no very complete definition of the blockchain in the world. Generally speaking, it is essentially a connected data structure. Each node in the blockchain can protect the entire blockchain through a set of consensus mechanisms during normal operation. Moreover, when these nodes work, the impact on each other is relatively small. It can be said that when any node has a problem, the entire blockchain can continue to work [2]. The blockchain has many basic structures, including not only the data layer and the network layer, but also the more important consensus layer, incentive layer and application layer, and the functions and functions of these structures are different.

Understanding blockchains requires attention to their specific characteristics, which together solve many of the problems of traditional data systems [2]. Among these features, the following five are particularly noteworthy:

### 2.2.1 Blockchain Ledgers Are Distributed

Unlike traditional databases, they are stored on a single computer or server. The blockchain is a distributed ledger that can be accessed directly by any party on the network. There is no single control point in the blockchain.

### 2.2.2 Point-to-Point Data Transmission

Since there is no single control point, blockchain data is transmitted directly from one user to another. All transmitted data immediately updates the entire ledger, so everyone on the network has the same single version.

### 2.2.3 Anonymous Transactions

Data transfers occur between addresses on the blockchain. This ensures transparency while allowing users to remain anonymous. However, if the blockchain is set to require proof of identity, no anonymity is required.

### 2.2.4 Transaction Records Are Immutable

Each blockchain transaction creates an irreversible and unchangeable record. The records are arranged in chronological order and are visible to all users in the chain.

### 2.2.5 Blockchains Are Programmable

Since blockchains can be bound to computational logic, functions and transactions can be programmed through rules and algorithms. Perhaps the easiest way to think about a blockchain is like a book(Fig.2) where each page (a block of data) references the previous page by page number (the

only block fingerprint). So, just as a book is essentially a series of pages that together provide a continuous flow of information, a blockchain is a bunch of packets that together provide a continuous, indelible transaction record.

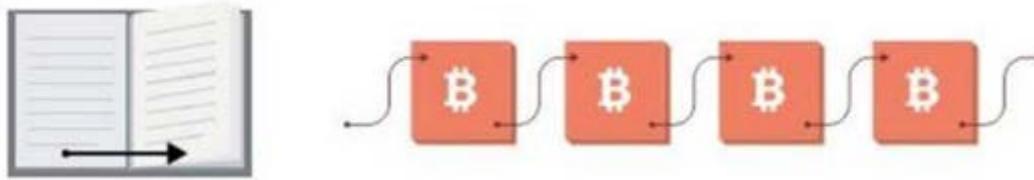


Fig.2 Blocks in a Chain Like Page Numbers in a Book

### 3. Big Data Characteristics

The so-called big data refers to a collection of data that cannot be captured, managed, and processed by conventional software tools within a certain time frame. Such data needs to be optimized by using updated processing modes and technologies, and eventually becomes a massive informational asset. In the era of information explosion, the amount of data information that can be obtained is very large. In order to process these data, it is necessary to comprehensively use a variety of technical means to make necessary screening of these low-density data information, and provide some industry fields. The most intuitive and valuable information to promote the development of the industry [3].

#### 3.1 The Comprehensive Openness of Big Data

In the era of big data, the construction of the financial sharing center makes the network environment more open and diversified. In the process of disseminating information through the Internet, the number of people who can participate in and access information is greatly increased, and it is vulnerable to hackers and viruses. Malicious attacks lead to the loss of important information and data of the enterprise [3]. Even hackers will tamper with the data in the system and cause irreparable damage to the enterprise. In the context of the full openness of big data, it will increase the possibility of information being stolen, attacked, destroyed, leaked, and falsified.

#### 3.2 Asset Attributes of Big Data

The rapid development of the economy has led to the continuous appreciation of big data, the increasing asset attributes of big data, and the continuous maturity of data mining and data analysis technologies, which has led to “excessive mining” of big data. In the process of analyzing big data mining, the risk of being illegally intercepted, attacked, tampered with or deleted is inevitably increased. Accounting information is one of the business secrets of enterprises [4]. In the fierce commercial competition environment, competitors have more unscrupulous access to commercial data, increasing the risk of corporate information being compromised and stolen.

#### 3.3 Reduced Controllability of Data

In the virtual Internet environment, the carrier of information is no longer a book or a company's computer, but a server and cloud storage distributed around the world. These Internet media are far beyond the company's control. Computer equipment is decipherable. Once these devices are maliciously attacked by hackers, it is difficult to take countermeasures [5]. The security maintenance of big data is far from being solved by turning off the power, and the risk of information leakage is greatly enhanced.

## **4. Blockchain Core Technology**

### **4.1 Distributed Accounting**

The biggest feature of distributed accounting technology is decentralization. No third-party trust institutions are needed in the blockchain network, which reduces the cost of credit. Every node in the system has the rights and obligations to book and store, and the entire blockchain is maintained by each node.

Each computer accessed in the system can be thought of as a single node, and the node is the biller in the blockchain network. In order to encourage node billing, the system will use incentives in accordance with the rules to encourage each node to book. Each transaction is reviewed and supervised by each node before being added to the account book [5]. Untrue transactions are rejected and recorded in the system to prevent tampering with the information. Each node stores a set of distributed book copies that cannot be tampered with and has a complete transaction history. Therefore, when a single node is attacked and fails, it will not cause the problem of information big data being destroyed or lost.

### **4.2 Time Stamp**

The timestamps connect the information of each block in time series to form a chain, and indicate the writing time of the data in the block, so that the data record is unique and continuous. If you want to tamper with the information and make a fake record, you must modify all the records in the blockchain. Each node saves the complete data of the blockchain network. Therefore, if you want to modify the past data records, you must construct it from scratch. A chain longer than before is used to cover the previous records, and the difficulty of fraud increases with time, and the cost of fraud is also very expensive [6]. The time stamp is based on the non-circularity of time, so that the data in the system can neither be modified nor revoked, ensuring the authenticity of the information and improving the anti-mite modification and security of the information.

### **4.3 Asymmetric Encryption Technology and Authorization**

In the blockchain system, asymmetric encryption technology is used to ensure the security and privacy of transaction records. Asymmetric encryption is an algorithm that selects different “keys” for data encryption and decryption. If the public key is used for encryption, it is decrypted with the private key; if it is encrypted with the private key, it is also verified with the public key. In general, public keys are publicly available. All participants can encrypt information with a public key. Only those who have a private key can gain access [6]. This technology can make information highly transparent and eliminate information. Symmetrical, it can effectively prevent the risk of information data being compromised, and ensure the security of information.

### **4.4 Smart Contract Technology**

Smart contract technology is a set of pre-made rules and a self-executing contract. If the application is in the system, the smart contract technology automatically processes, stores, verifies, and reports the information after a new transaction is added to the account book [7]. The autonomy of smart contracts can reduce the risk of information being tampered with.

## **5. Distributed Processing of Big Data Based on Blockchain**

### **5.1 Distributed Cloud Data Computing**

The Internet era has made data processing particularly important. In the process of informationization, enterprise business has precipitated a large amount of data, and the data covers a wide range, data is scattered, and the redundancy is redundant, which increases the data exchange and sharing among various departments of the enterprise [7]. The birth of the data center has greatly

enhanced the data security of enterprises, which is an inevitable trend of the deep development of enterprise information.

The data center built by blockchain technology will improve data integrity and accuracy, improve the uniformity of enterprise information systems and data consistency; eliminate information silos, realize the sharing of internal information resources, and improve the utilization of enterprise data. To realize deep development and utilization of data, to transform data into development resources, to realize data value-added services; to facilitate data management, improve data security and management effectiveness; improve information system disaster prevention capability and achieve sustainable development of enterprises [8].

## **5.2 Construction of Distributed Cloud Data Computing Center**

Data centers developed and built on blockchain technology provide more secure and efficient data support for enterprises, and can achieve more data integration capabilities, providing more employment and taxation for data computing center locations.

Since the data and content that the data computing center needs to process is too large and requires a large amount of server computing as a support, the data center generally chooses to be built near a power plant or a large campus. The data center is mainly composed of a software system and a hardware system [8]. A large amount of electrical energy is required to support the operation of the hardware system. Each data center is planned to consist of more than a thousand small data processors, which can reasonably utilize a large amount of surplus power energy in the Yunnan-Guichuan area, increase corporate income, provide jobs, and raise local government taxes. After the completion of the construction of the first data center, the data center will continue to be added according to the actual situation to achieve a better service for SMEs in the future.

## **6. Summary**

Blockchain and big data technologies are a kind of data processing technology developed based on developed computer technology and network technology. The two are different in characteristics and essence, but they are not completely opposite. The center of big data the decentralization of the block and the blockchain have their own advantages and disadvantages. It is necessary to seek common ground while reserving differences between the two, to integrate the places of comparative advantage, and then apply them to various fields of social development to continuously promote the further rapid development of society.

## **Acknowledgement**

Fund Project: This paper is the outcome of the study, Research on Effective Measures of Student Education and Management Based on “Smart Campus” Data, which is supported by the Foundation for Key Projects of the Planning Office of Hubei Province; the Project Number is 2017GA086. It is also the outcome of the study, Analysis on the Online Course Learning Process of Higher Vocational Students Based on Big Data, which is supported by the Foundation for Projects of Wuhan City Polytechnic; the Project Number is 2018whcvcS01.

## **References**

- [1] X. Shen and Q.Y. Zhai, Review of Blockchain Technology, Journal of Network and Information Security, vol.11, pp.11-20, 2016.
- [2] Y.Y. Yuan and F.Y. Wang, Current Status and Prospects of Blockchain Technology, Journal of Automation, vol.4, pp.481-484, 2016.
- [3] J.Y. Xu and X.T. Zhu, Research on the Impact of Blockchain Information Technology on Accounting Supervision, Friends of Accounting, vol.1, pp.156-158, 2018.
- [4] L.Ch. Zhang and L.L. Yu, Application Research of Blockchain Technology in the Field of Big

Data, Research and Information Technology and Application, vol.21, pp.23-25, 2018.

[5] P.Ch. Qiao, Information big data security protection from the perspective of blockchain technology, Modern business, vol.22, pp.45-47, 2017.

[6] W. Zhong and Y.Z. Jia, Application Prospect of Blockchain Technology, Journal of Network and Information Security, vol.17, pp.32-5, 2016.

[7] X.L. Dong and Y.J. Mao, Distributed Collaborative Technology Based on Blockchain, Library Theory and Practice, vol.11, pp.85-89, 2018.

[8] Y.T. Hu, Design and Implementation of Distributed System Automation Test Platform, University of Electronic Science and Technology, vol.3, pp.45-47, 2014.