Government Data Management in the AI and Big Data Era: Research on Sharing and Governance Mechanisms

Xiao Jiong'en1,3 Wu Yingliang1,2

Abstract

Guided by collaborative theory, this study establishes a data sharing platform with blockchain as the core. Taking the application scenario of digital identity as an example, the data sharing platform adopts an authorization-based consensus mechanism (PoA) combined with Practical Byzantine Fault Tolerance (PBFT) to explore the collaborative innovation application of the government affairs system. The collaborative innovation application has achieved inter-departmental paperless approval among multiple government departments, realizing "zero running" for citizens in handling affairs in more than 20 government affairs scenarios.

Keywords Blockchain Technology; Government Affairs System; System Collaborative Application; Digital Identity

1. Introduction

Against the backdrop of the State Council General Office issuing Document No. 39, "Notice of the State Council General Office on Printing and Distributing the Implementation Plan for the Integration and Sharing of Government Information Systems," governments at all levels have also introduced relevant policies, actively promoting the integration and collaborative innovation of information systems. With the rapid development of government information systems, the data within these systems exhibit characteristics such as large data volume, rapid data growth, and data heterogeneity. The public service issues faced by modern governments are becoming increasingly complex, with public service issues in one field often involving multiple functional departments, requiring coordination between different government organizations to solve the problems [1]. The key to the integration of government information systems is the integration of data sources in various heterogeneous systems. Under the framework of system collaboration theory, many scholars believe that new technologies such as blockchain have high research value in solving data sharing issues [2]. Starting from the perspective of top-level planning, building collaborative public services [3], applying blockchain technology to the sharing of government data and the integration of government systems, and thereby promoting the efficiency of government systems and optimizing convenient services for the public is one of the focal points of research for scholars, experts, and government workers.

2. Research Status at Home and Abroad

(1) System Collaboration Theory

System collaboration theory, also known as "synergy theory" or "harmony theory," is an emerging discipline that has gradually formed and developed since the 1970s on the basis of multidisciplinary research. It is an important branch theory of systems science. Its founder is Professor Haken, a renowned physicist at the University of Stuttgart in the Federal Republic of Germany [4]. System collaboration theory studies the relationship between the whole and its parts. According to the different manifestations and results, system collaboration can be divided into two types: complementary effects and synergistic effects. In terms of the types of system collaboration, the ways of system

collaboration include: (1) integration of data; (2) integration of application systems; (3) integration of business processes. Introducing the collaborative business ideas from the field of enterprise management into the field of public management has given rise to the new concept of "collaborative governance" [5]. Collaborative governance refers to a new type of government work model under the background of informatization, where government departments use information technology to carry out cross-departmental business collaboration, ultimately changing the methods of administrative management to achieve the most full utilization of government resources [6]; it elaborates on the theory and practice of government collaboration, bringing public and private stakeholders together in the same system, coordinating the interests and power of all parties, and balancing the resources of all parties [7]. The focus of electronic government collaboration is the convergence of data on the backend, promoting the integration of systems on the front end, and finally optimizing the system processes to improve the efficiency of public services.

(2) Blockchain Technology and Applications

Definition of Blockchain [8]: Blockchain technology is a technical solution that enables all participants in a system to collaborate and collectively record and maintain a reliable database through decentralization and encryption technology under the conditions of multiple entities that do not need to trust each other. Blockchain technology allows any number of entities participating in the system to encrypt all information exchanges in the system over a period of time through cryptographic algorithms, calculate and record them into a data block, and generate a cryptographic key for the data block to link to a data block on the data chain for verification. At the same time, all participants and entities in the system can jointly determine whether the record is true.

Blockchain technology is considered to be the next disruptive core technology after the steam engine, electricity, and the Internet. If the steam engine released human productivity, electricity solved the basic needs of human life, and the Internet completely changed the way information is transmitted, blockchain, as a technology for constructing trust, has significant value.

Overseas applications of blockchain in e-government include: (1) The United Nations and the World Identity Network (WIN) announced a cooperation statement at the Humanitarian Blockchain Summit, aiming to use blockchain technology for identity authentication pilot projects. Storing digital identities on the blockchain can improve the chances of catching criminals and recovering children, and can protect privacy, trace and prevent human trafficking activities. (2) International organizations such as the United Nations pay attention to using new technologies for the operation and management of public affairs. In the management of refugee assistance in Jordan, the United Nations, through its technology partners, uses blockchain technology for financial aid and support. Refugees can receive financial assistance through an Ethereum-based payment platform, improving the traceability and credibility of transfer payments. (3) A report by the UK Police Foundation outlines how blockchain will change the UK's criminal justice system, providing a better experience for every service user. The report states that the UK's judicial system still relies on paper documents and traditional IT systems for work, leading to low work efficiency. For example, a large number of manual and manual handling of tedious procedures result in only half of the interrogations being held on time, and the cost of multi-party participation in accounting is high, and the error rate is high. The UK is currently conducting pilot projects for digital case file management and online appeal processes, which have initially highlighted the advantages of blockchain in reducing costs, optimizing work processes, and improving service quality. (4) The Dutch Ministry of Justice is using blockchain technology to build a digital legal system, achieving decentralized, automatic execution,

and human-computer interaction in legal enforcement through the form of smart contracts. This smart contract platform fully considers the automatic execution of machine language and the complexity of real situations in the design, ensuring the automatic execution of some tasks while placing the handling of some special cases under human control.

Blockchain Features: (1) Security [9], blockchain technology makes the distributed ledger difficult to be attacked by hackers because it does not store transaction records in a single database, but retains multiple copies in the blockchain, making the data very difficult to tamper with. (2) Scalability [10], according to the needs of business scenarios and participating institutions, nodes can be flexibly increased.

In the process of government collaboration, data is the foundation. Through the study of blockchain technology, from a top-level design perspective, the platform for data sharing and collaborative processing is planned [11-13], guided by electronic government governance theory [14], top-level planning is carried out, and the construction of data governance systems is built, to build an overall framework for blockchain in government data sharing, which can effectively solve issues such as data security and trustworthiness, and track the use of data.

3. Blockchain Sharing Architecture Solves the Problem of Government Data Sharing

(1) The Problem of Collaborative Data Sharing in Government Systems

Government systems have been initially scaled after many years of construction, but at the same time, the phenomenon of each bureau, commission, and office system being independent has also emerged. The phenomenon of "data islands" in government systems is serious. In order to eliminate data islands and achieve multi-dimensional data convergence and collaborative government systems, governments at all levels have tried different technologies. (1) Email or FTP upload form; (2) Database middleware form [15]; (3) Data interface web service method [16]; (4) Data service bus method [17]. However, the above four forms of data sharing also have many shortcomings, and the problems faced by government system collaboration can be summarized into three aspects.

1. The issue of data security, the problem of data tampering, in the system, if business handlers, system administrators, database administrators, and other parties collude to tamper with certain data in the system, it will severely affect the credibility of the data in the system.

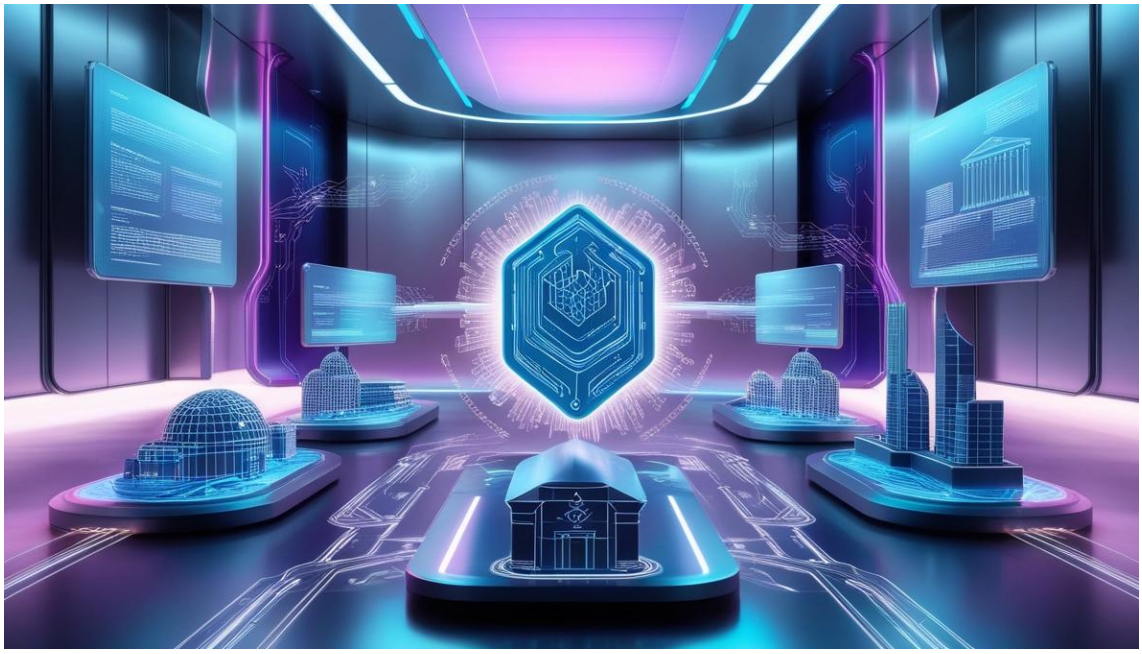
2. The scalability issue of the data convergence platform, when a new institution or system needs to submit data to the sharing platform, complex adaptation work is required.

3. After the data is converged, whether the data is used within the specified scope, whether it is used illegally, this puts forward higher requirements for the management methods and tools of the data management department.

In the government system, there are many problems in data convergence, and the emergence of blockchain technology has attracted the attention of scholars, government officials, and experts in the government system. Through research, it has been found that the technical features of blockchain are helpful in solving the problems faced in the process of data convergence.

(2) Blockchain Government Data Sharing Architecture

Based on the above analysis of the issues in the collaboration and data sharing of government systems, the integration and sharing of government information systems, and the construction of government informatization projects, it is necessary to adopt a reform-minded approach and innovative measures to transition from the construction of single systems to the construction of large integrated systems. Establish a new model of government informatization construction driven by data resources to drive government innovation, promote the integration and in-depth development of information resources, and facilitate the sharing and use of government information and the optimization and collaboration of business processes. Construct an integrated large platform, share and use big data, and build a collaborative and interconnected large system. Combining blockchain technology to achieve data sharing and system integration in government systems, the architecture is shown in Figure 1.



The collaborative innovation platform architecture of the blockchain is divided into three main parts: (1) the business system layer, which includes the information systems built by local departments and the directly managed systems constructed by the national and provincial governments. The provincial direct systems have accumulated a large amount of process and result data, such as the national Golden Tax and Golden Quality projects, some of which are sensitive and confidential data, and some can be publicly disclosed and reused. According to the standard system of data governance, the national direct management system, based on the Daas technology of the team led by Academician Mei Hong [18-21], establishes a data pipeline. Local areas can extract data within their jurisdiction, break the data islands, and achieve "on-chain" of data. Another part is the system built by the region, integrated using a microservices architecture [22]. At the same time, the data in each system is sorted out, and through the means of de-identification and declassification, ensure that the data is publicly available and usable, and then proceed with "on-chain".

(2) Data Sharing Platform, with blockchain as the core, the types of blockchain include private chains, public chains, and consortium chains [23]. In the application of blockchain technology in government systems, a combination of private and consortium chains is generally adopted. In the practical application of government scenarios, the construction of regional cloud platforms usually adopts a dual-site and triple-center model. The deployment of blockchain nodes is also based on the construction of the cloud platform and the deployment of government systems, using multiple servers in the cloud platform for the deployment of blockchain nodes. This deployment of private chains generally has no fewer than 20 nodes. Each government system organizes data resources, and according to the deployment of blockchain nodes, connects to blockchain nodes to promote the "on-chain" of data resources and establish a data resource directory. Data sharing based on the blockchain as the core is the core of the entire system platform, and its architecture is shown in Figure 2. The ① statistical bureau server cluster and ③ e-government cloud server cluster deploy blockchain nodes according to the system situation of various bureaus, commissions, and offices, with ② as the backup server cluster. A government data sharing platform is established based on the digital identity and corporate legal person data as the link, combined with the blockchain system, to achieve the goal of data sharing to the greatest extent. A secure network topology design is created to allow the government data sharing platform to operate in a safe, reliable, and efficient manner, ensuring smooth operation of government systems and blockchain systems.

(3) Application Layer, the blockchain-based data sharing platform has achieved the convergence of data, and innovative research on government affairs is carried out on this basis. The government data sharing platform, supported by the blockchain technology platform, breaks down the information barriers between relevant departments, establishes a horizontal linkage mechanism, and facilitates the linkage and sharing of data interfaces for public services such as convenience, public services, support for the elderly and disabled, medical care, human resources and social security, education, civil affairs, and housing and urban-rural development. A credit model is established to create a credit file for each citizen, constraining their social behavior norms. Through the analysis of the converged data, portraits of legal persons and natural persons are realized. At the same time, a data mining and analysis system is implemented, and the use of data is regularly audited.

4. Blockchain's Innovative Application in Government Collaboration

(1) Digital Identity Architecture

Based on the scenario of blockchain-based digital identity authentication, collaborative innovative application research is conducted to explore the sharing and innovative application of government data. The innovative application of blockchain technology in government collaboration mainly focuses on the collaboration of data within the system, using personal digital identity as the carrier to achieve the sharing of government system data, and on this basis, carry out innovative applications in government collaboration. The architecture of digital identity authentication based on blockchain is shown in Figure 3.

Firstly, front-end applications. Identity documents such as identity cards, household registration books, marriage certificates, and other valid credentials are uploaded via mobile devices, PCs, and web platforms to the relevant government systems for verification, and relevant data is accumulated within these systems.

During the participation of citizens in various government affairs, different data are accumulated in the personal data space of each government department. These personal

data are ultimately returned to the citizens through the APP. Therefore, the supporting APP allows citizens to register a digital identity and, through technical means, to access their personal data that has been accumulated in the personal data space of different government departments. Based on this data, citizens can complete higher-quality government services through the APP.

At the same time, the citizens' digital identity information is stored in the blockchain, but any access or use of the information by government units and business departments must be authorized by the citizens.

Secondly, the application service system. Each commission and office transforms government systems according to the service requirements for official duties, transforms communication services, encrypts the transmitted information, and develops an IPFS[24] (InterPlanetary File System) interface. IPFS combines the advantages of Git, BitTorrent, kademlia, SFS[25], and the web, providing an interface as simple as HTTP web. It ensures that personal space data can be stored in the blockchain's IPFS. IPFS is a distributed file system that seeks to connect all computer equipment with the same file system. In various government systems, non-structured large files such as images and ID photos are stored in IPFS, and at the same time, the hash values of the files are extracted and uploaded to the blockchain. By utilizing the undeniable and tamper-proof characteristics of the data stored in the blockchain, the uniqueness and accuracy of the shared data are ensured.

Lastly, the blockchain core system. The blockchain core system is the heart of the entire government blockchain system. The consensus mechanism is the core of the blockchain system. Common consensus mechanisms are listed in Table 1 below.

Improvements have been made in consensus and fault tolerance mechanisms. The generation of each block no longer uses PoW [26] (Proof of Work) but adopts an improved consensus mechanism that combines PoA [27] (Proof of Authority) with the Practical Byzantine Fault Tolerance PBFT [28] (Practical Byzantine Fault Tolerance). PoA, a consensus mechanism based on authorization, is more efficient, energy-saving, and hardware resource-saving in the application of blockchain in the government sector compared to PoW. The Practical Byzantine Fault Tolerance PBFT is used for fault tolerance processing, providing fault tolerance of (n-1)/3 under the premise of ensuring liveness and safety. In distributed computing, different computers attempt to reach a consensus through information exchange; however, in some special cases, the coordinating computer (coordinator/commander) or member computer (member/lieutenant) may have system errors and exchange incorrect information, affecting the final system consistency. The Byzantine Generals Problem seeks possible solutions based on the number of faulty computers. With PBFT, the blockchain network (n) nodes can withstand (f) Byzantine problem nodes, where f=(n-1)/3. In other words, PBFT ensures that at least 2*f+1 nodes reach a consensus before adding information to the distributed shared ledger. Through the improvement of consensus and fault tolerance mechanisms, it ensures that the block generation speed in the government blockchain is less than 1 second.

(2) Main Functional Process of Digital Identity

Based on the analysis of the digital identity architecture, the functional realization process of the digital identity is shown in Figure 4, mainly describing the "on-chain" process of the digital identity. After the data is "on-chain," users will not need to carry relevant documents in all authorized department systems of the blockchain. They only need to authorize through the user's mobile terminal, and the relevant business systems can obtain the required document information to handle the business.

The process is described as follows.

1. Create Digital Identity

(1) Send the identity creation request and the necessary information to the Server.

(2) Create a digital identity on the blockchain.

2. Declare Attributes

(1) The user declares attributes (such as identity card information and a portrait) and sends them to the server.

(2) Place the attribute information into IPFS.

(3) Store the hash information of the attributes on the blockchain.

3. Authenticate Attributes

(1) Authenticate directly with the relevant institution or through other means.

(2) Write the authentication information to the chain.

4. Log in to Website/Systems Requiring User Attributes

(1) Scan the website's QR code to access the login interface.

(2) Send the user's identity attribute information to the server through a mobile or web platform.

(3) The server checks on the blockchain whether the user's attributes and mobile phone match.

(4) The server retrieves the user's authenticated information from the blockchain, such as the identity card.

(III) The Effectiveness of Blockchain Government Innovation Applications

By following the above process, the "on-chain" of personal digital identity is achieved, and personal identity data is shared within the government system. The innovative effects of government achieved through blockchain include the following three aspects.

1. Blockchain Platform Achieves Information Sharing

The greatest value of blockchain technology lies in data trust and the establishment of data sharing based on data trust. It also builds an efficient collaboration and linkage government management system based on data trust/data sharing. Finally, a personal credit evaluation model and system based on trusted data are constructed through the government management system, thereby bringing government service capabilities and social governance capabilities to a new stage.

The system achieves data sharing and linkage with business systems such as government affairs, public services, elderly care, disability assistance, medical care, human resources and social security, education, civil affairs, and housing and urban-rural development, realizing an efficient collaborative government linkage management system. It establishes a working connection mechanism between various government departments, the platform achieves interconnectivity, office dataization, full coverage of supervision, tamper-proof real data, promotes the public, transparent, and standardized operation of government data work, and advances various departments' government services to a new level in an efficient collaborative manner.

2. The blockchain platform can achieve cross-departmental paperless approval

Cross-departmental affairs approval is usually carried out through paper correspondence, using manual approval and reply methods to complete the approval of various affairs. The characteristics of blockchain technology ensure the immutability and traceability of data. By digitizing system approval materials and storing them on the blockchain, and incorporating blockchain electronic signature technology for staff and approvers, electronic archives are stored on the blockchain and data sharing is achieved. This realizes a paperless approval process across departments, greatly improving work efficiency and facilitating the linkage capabilities of various departments.

The government data sharing platform, supported by the blockchain technology platform, breaks down information barriers between relevant departments, establishes a horizontal linkage mechanism, and creates a credit model. It establishes credit files for each citizen, constraining their social behavior norms, and provides linkage and shared data interfaces for work information in government affairs, public services, elderly care, disability assistance, medical care, human resources and social security, education, civil affairs, and housing and urban-rural development.

3. "Zero-visit" convenience service upgrade

Based on the government system collaboration of blockchain, the control of digital identity is transferred from the central server to the individual, giving individuals control over their digital identity. Focusing on the individual as the subject, a data set related to the individual's data and its relationships is constructed around three dimensions: data, business, and security, to create a "personal data space." On this basis, various government systems can access the trusted personal digital space on the blockchain, combined with facial recognition, real-name phone certification, electronic ID cards, etc., to achieve "zero-visit" government services. This has transformed the government service model, changing conditional approval to trust-based approval, and passive services to active services. Trust-based approval under the conditions of big data is the trend of modern social governance. Upgrading "zero-visit" convenience services with blockchain technology can achieve a total of 2 "lead-run" items, as shown in Table 2.

5. Conclusion

With the rapid development of government systems and the diversification of their user base, the functional requirements of the systems also become increasingly varied. This demands a high degree of collaboration between multiple government departments, extensive data sharing, and system integration. This paper addresses the pain points of traditional government systems, which are distributed across various departments, leading to data silos with low levels of data sharing and uncontrollable data after sharing. By leveraging the core advantages of blockchain technology, this paper

proposes a collaborative innovation application for government systems based on blockchain technology. Starting from a city-level perspective, it coordinates the establishment of a blockchain data sharing platform. Government systems managed by municipal departments directly "put data on the chain," while systems managed by provincial and national authorities extract data within their authority and "put it on the chain" using Daas technology. This breaks the "data silos" of government data. In the application scenario of citizen personal information digital space, the paper analyzes the architecture and implementation process of establishing digital identity based on blockchain technology. It solves the problem of data sharing among various business systems, improves the efficiency of data use, and enables collaboration and innovative applications between government systems of multiple departments. However, the application of blockchain technology in government system collaboration also faces some challenges, such as huge hardware investment in government systems, high resource consumption, and resistance from some departments during implementation. As these issues continue to be resolved in practice, blockchain technology will play an increasingly important role in the collaborative innovation applications of government systems.

References

1. Chuns A, Luna-Reyes LF, Sandoval-Almazán R. Collaborative E-government [J]. Transforming Government People Process & Policy, 2012, 6(01): 5-12.

2. Yuan Yong, Wang Feiyue. Current Status and Prospects of Blockchain Technology Development [J]. Acta Automatica Sinica, 2016(04).

3. Zhang Kailin, Wang Na, Huang Lei, et al. Constructing Collaborative Public Services: Research on the Top-Level Design Method of Governmental Informatization [J]. Management World, 2013(08).

4. Haken H. Advanced Synergetics: Instability Hierarchies of Self-organizing Systems and Devices [M]. Springer Science & Business Media, 2012.

5. Vella A, Nicole C. Creating Public Value Through Collaborative Governance - Case Study: The Strategic Development of the Bays Precinct, Sydney Transformation Plan [M]. Sustainable Development Research in the Asia-Pacific Region. Springer, Cham, 2018.

6. Emerson K, Nabatchi T, Balogh S. An Integrative Framework for Collaborative Governance [J]. Journal of Public Administration Research and Theory, 2012, 22(01): 1-29.

7. Ansell C, Gash A. Collaborative Governance in Theory and Practice [J]. Journal of Public Administration Research and Theory, 2008, 18(04): 543-571.

8. Yassine SM, Latifa A, Sarah B. Blockchain and Bitcoin: The Revolution of the Financial Industry [J]. European Journal of Economics, Law and Politics, 2017, 14(04): 98-106.

9. Puthal D, Malik N, Mohanty SP, et al. The Blockchain as a Decentralized Security Framework [Future Directions] [J]. IEEE Consumer Electronics Magazine, 2018, 7(02): 18-21.

10. Kim HM, Laskowski M. Toward an Ontology-driven Blockchain Design for Supply-Chain Provenance I. Intelligent Systems in Accounting, Finance and Management, 2018, 25(01): 18-27.

11. Wang Maolu, Lu Jingyi. Research on the Application of Blockchain Technology in Government Governance [J]. E-Government, 2018(02).

12. Gao Guowei, Gong Zhanli, Li Yongxian. Research on the Collaborative Sharing Model of Government Basic Information Based on Blockchain [J]. E-Government, 2018(02).

13. Jiang Yuhao, Jia Kai. Research on the Transformation of Public Decision-Making Responsibility Mechanism under the Blockchain Technology Path Based on Big Data [J]. E-Government, 2018, (02): 26-35.

14. Sun Baowen, Wang Tianmei. E-Government Governance: Theoretical Origin and Theoretical Framework Definition [J]. Economic Management, 2012(05).

15. Li Shun, Xu Fuchun, Kong Yimin, et al. Discussion on the Construction of the National Environmental Data Center [J]. Environmental Protection, 2006(02).

16. Shi Xin. Design and Implementation of Data Integration Platform for Shared Data Center [D]. Suzhou: Soochow University, 2006.

17. Li Xiaodong, Yang Yang, Guo Wencai. Data Sharing and Exchange Platform Based on Enterprise Service Bus [J]. Computer Engineering, 2006(21).

18. Tong Nannan, Chaolemun. The Change of Data Management Concept in the Big Data Era: From Result School to Process School [J]. Information Theory and Practice, 2017(02).

19. Huang Gang, Wang Qianxiang, Mei Hong, et al. Research on Reflective Middleware Based on Software Architecture [J]. Journal of Software, 2003(11).

20. Huang Gang, Mei Hong, Yang Fuqing. Runtime Software Architecture Based on Reflective Software Middleware [J]. Science in China Series E: Technological Sciences, 2004(02).

21. Song Hui, Huang Gang, Wu Yi Han, et al. Modeling and Maintenance of Runtime Software Architecture [J]. Journal of Software, 2013(08).

22. cerny T, Donahoo M J, Trnka M. contextual understanding of Micro service Architecture: current and Future Directions [J] . Acm sigapp Applied computing Review, 2018, 17(04): 29—45.

23. wu L, Meng K, Xu s, et al. Democratic centralism : A Hybrid Block chain Architecture and its Applications in Energy Internet [c] . Energy Internet (IcEI), IEEE International conference on. IEEE, 2017.

24. Koonce L. The wild, Distributed world: Get Ready for Radical Infrastructure changes, from Block chains to the Interplanetary File system to the Internet of Things [J] . Intellectual property & Technology Law Journal, 2016, 28(10): 3.

25. Lai Roy, Lee Kuo chuen David. chapter 7— Block chain———From public to private [M] . Handbook of Block chain, Digital Finance, and Inclusion, 2018.

26. vukolitm. The Quest for scalable Block chain Fabric: proof—of—work vs . BFT Replication [M] . International workshop on open problems in Network security. springer, cham, 2015.

27. Goranovit A, Meis el M, Foti adis L, et al. Block chain Applications in Microgrids an overview of current projects and concepts [c] . Industrial Electronics society, IEcoN 2017—43rd Annual conference of the IEEE. IEEE, 2017.

28. sukhwani H, Martinez J M, chang X, et al. performance Modeling of pBFT consensus process for permissioned Block chain Network (Hyperledger Fabric)[c] Ⅱ Reliable Distributed systems (sRDs), 2017IEEE 36th symposium on. IEEE, 2017.

Author Biographies

Xiao Jiong'en1,3 Wu Yingliang1,2

1. School of Economics and Trade, South China University of Technology;

2. Research Institute of Modern Service Industry, South China University of Technology, Guangdong Guangzhou 510006; 3. Guiyang Municipal Big Data Development Management Committee, Guizhou Guiyang 550081