Blockchain-based big data forecasting system for the marine environment

Luyao Han, Xiaoxiao Shi, Mingmei Dong, Guanghao Wei and Xiao Song
National Marine Data and Information Service, Tianjin 300171

1 E-mail: shixiaoxiaojt@sina.com

Abstract. In response to the difficulty in sharing and low utilization of marine forecasting resources such as marine environment data, marine environment big data forecasting methods and marine environment forecasting products, the author proposed a blockchain system architecture consisting of data layer, network layer, transaction layer and application layer. Based on key technologies such as FISCO alliance chains, incentive mechanism and smart contracts, a marine environment big data forecasting system is developed in order to manage marine environment data and marine forecast products such as measured data, marine information products, sea surface temperature, sea surface height, 3D thermohaline circulation, typhoon tracks, and red tide occurrence rate on the chain. The system is able to achieve the decentralized common maintenance and trusted data sharing of marine forecast resources. It also promotes the development of marine environment forecasting technology, and provides a concrete and workable solution for the integration of blockchain and marine technology.

1. Introduction
With the exponential growth of observation data, marine environment forecasting has also entered the era of big data. The study of big data-based marine environment forecasting technology is of great significance for the development of China's independent, safe and controllable marine environment forecasting. During the 13th Five-Year Period, based on the support of the National Key Research and Development Program, China established the ocean big data analysis and forecasting system, which introduced big data technology into ocean forecasting for the first time, realizing the data-driven forecasting of sea surface temperature, sea surface height, 3D thermohaline circulation, typhoon tracks and red tide occurrence. It provided a theoretical, technical and innovative practice for the improvement of ocean forecasting capability [1]. As the key output in the business chain of marine environment forecasting, the birth of marine environment datasets, marine big data analysis forecasting methods and marine environment forecasting products often requires the investment of huge time in research and implementation, which are valuable intellectual property rights [2]. The current mainstream marine environment forecasting systems often adopt a centralized architecture to realize the storage and management of operational data, which is subject to security risks such as data tampering and illegal copying, and is adverse to the protection of intellectual property rights. This greatly reduces the motivation of researchers to share their research findings leading to the slowing-down of the marine environment forecasting technology development.

Since the emergence of blockchain technology in the United States in 2015, blockchain has led many industrial changes like government applications, commercial applications, underlying technology and our daily life. As an emerging technology with great potential, blockchain is
decentralized, tamper-evident, open sharing, safe and secure [3], and is well suited to solve the challenges of sharing marine environment forecasting data resources. At present, many researchers at home and abroad have started using blockchain to explore and practice in the field of data protection and sharing [4]. In 2016, Azaria built a decentralized medical data management system using smart contracts to achieve blockchain-based medical data sharing and managing [5]. In November 2016, Cai Weide proposed a blockchain-based application system developing method. They designed a dual-chain model of account chain and transaction chain, which provided a new idea for blockchain system development [6]. In 2018, the U.S. Defense Logistics Agency (DLA) developed a blockchain-based disaster relief application that facilitates the sharing and exchanges of information transportation. In 2020, the U.S. Food and Drug Administration (FDA) used blockchain technology to address the lack of transparency and security problems in health data processing. In September 2017, by using blockchain technology, Xue Tengfei designed a medical data sharing model to achieve data sharing among different medical institutions [7].

In this paper, based on the full reference of previous experience, the author adopted the FISCO consortium chain technology to develop a blockchain-based marine environment big data forecasting system, exploring a new model of marine research resource sharing and aiming to solve the sharing problem of marine resources.

2. Key technology research
The key technologies to be considered in the development process of a blockchain-based ocean big data forecasting system include FISCO consortium chains, incentive mechanism and smart contracts.

2.1. FISCO consortium chains
According to access mechanisms, blockchains are divided into 3 categories: public, private and consortium chains, in which consortium chains are jointly managed by multiple organizations or individuals, where data on the chain are only allowed to be read, written and sent for transactions by a specific group who have obtained permission [8]. Compared to public chains, consortium chains are more private and underlining on decentralization, and are more suitable for building blockchain systems among the same industry.

The marine environment big data forecasting system is developed by the FISCO, an open source consortium chain underlying framework platform led by China, which supports multi-chain parallelism and cross-chain communication technology. FISCO can quickly form group architecture and build blockchain under limited resources. FISCO provides core modules like pluggable consensus mechanism, privacy protection algorithm, cross-chain communication protocols, distributed storage, visualization tool, which are easy to develop and maintain smart contracts. Its block-out speed can reach the second level, and the single-chain TPS (Transaction PerSecond) can reach 20,000 [9].

2.2. Incentive mechanism
Incentive mechanism is an economic concept introduced in the blockchain technology system, which mainly includes issuance mechanism and distribution mechanism [10]. The purpose of developing an incentive mechanism is to fully mobilize the initiative of the participants by rewarding tokens and points to encourage each node to upload intellectual property to the chain, in order to maintain the vitality of the blockchain in the long term, at the same time ensure the virtuous cycle of the whole system.

The intellectual property rights of the marine environment big data forecasting system mainly include 3 categories, like marine environment data, marine big data analysis and prediction methods, and marine environment forecasting products. On the premise of ensuring data security, each participant of the blockchain is encouraged to manage the intellectual property rights on the chain, and the automated reward and consumption of system points are realized through the formulation of a complete point allocation scheme to promote the sharing and use of marine environment forecasting resources [11]. Table 1 lists the point reward rules of the marine environment big data forecasting
system. For example, an ocean data provider can get 5 reward points for uploading ocean station observation data to the chain, and an ocean data user can get the on-chain access to that data by consuming 3 points.

Table 1. Integration rule of big data forecasting systems for the marine environment.

| Number | Type of behavior                  | Bonus Points | Spending Points |
|--------|----------------------------------|--------------|-----------------|
| 1      | Upload data to the chain         | 5            | -               |
| 2      | Use data                         | -            | 3               |
| 3      | Upload method to the chain       | 8            | -               |
| 4      | Use Method                       | -            | 5               |
| 5      | Upload forecast product to the chain | 3         | -               |
| 6      | Use forecast product             | -            | 1               |
| 7      | New User Registration            | 1            | -               |

2.3. Smart contracts

Smart contract is a core technology in the blockchain system. It is built on the basis of consensus and full trust of all nodes, supporting the process of signing contracts in the form of digitally anthropomorphized third-party institutions. It agrees on transaction participants’ rights, obligations and information in a computer-recognizable way, which is the reason why blockchain can be "decentralized" [12]. Compared with traditional contracts, smart contracts are tamper-evident, open, transparent, and run permanently, guaranteeing the trustworthiness and security of blockchain transactions.

The automatically and precisely execution of smart contracts generally include five main steps: contract formulation, state checking, identity verification, transaction execution, and transaction uploading [13]. Each participant on the blockchain jointly formulates a smart contract and regularly rotates its status. Once the agreed trigger conditions are reached, permission verification is carried out on the private key of the application account, and the account that passes the verification can complete the execution operation of the corresponding transaction and deposit the transaction result in the block [14]. Figure 1 illustrates the basic logic of smart contract execution in the marine environment big data forecasting system: the marine environment data provider and the marine big data analysis and forecasting method provider use the digital address of the private key to obtain the execution authority of the smart contract. They upload the intellectual property to the chain, and then obtain the reward points; the producer of marine environment forecasting product uses the private keys digital address to obtain the execution authority of the smart contract, spending points for using the data and methods on the chain.

Figure 1. Smart contract for the marine environment big data forecasting system.
3. System architecture design
The traditional blockchain architecture generally consists of a six-layer architecture of data layer, network layer, consensus layer, incentive layer, contract layer, and application layer [15], in which the technologies and mechanisms of the consensus layer, incentive layer, and contract layer need to be reflected in the transaction process. Therefore, the six-layer structure was simplified in the architectural design process of the marine environment big data forecasting system [16]. The consensus, incentive, and contract layers were merged into the transaction layer, and a four-layer system architecture consisting of the data layer, network layer, transaction layer, and application layer was constructed, as shown in Figure 2.

![Architecture of the big data forecasting system for the marine environment.](image)

3.1. Data layer
The data layer is mainly composed of 2 parts: on-chain data and off-chain data [17]. Figure 3 illustrates the data layer structure of the marine environment big data forecasting system.

(1) On-chain data
The data on the chain is stored in data blocks as the minimum storage unit, each block consisting of a block header and a block body. The block header encapsulates the hash value, current version number, timestamp, target hash value, and the Merkle root linked to the previous block. The block body data is stored in the form of Merkle tree and points to a unique Merkle root, which mainly includes 2 categories: smart contract status and system resource directory list. The smart contract status information includes binary code, number of transactions, account points status, etc. The system resource directory list includes resource unique identifier, resource type, resource access URL, etc.

(2) Off-chain data
The data under the chain is stored in a MySQL distributed database, which mainly includes marine environment data, marine environment big data analysis and forecasting methods, and marine environment forecasting products. Thereinto, the types of marine environmental data include: marine measured data, marine information products, human activity data and information collected on the Internet. The types of marine environmental big data analysis and forecasting methods include: classification algorithms, clustering algorithms, correlation algorithms and prediction algorithms adapted to marine disciplines. The types of marine environmental forecasting products include: sea surface temperature forecasting products, sea surface height forecasting products, three-dimensional...
thermohaline forecasting products, typhoon tracks forecasting products and red tide occurrence forecasting products. All these data and products can be discovered and accessed through the corresponding resource URLs.

Figure 3. Data layer structure of the marine environment big data forecasting system.

3.2. Network layer
The network layer is mainly used to transfer information between the nodes of the consortium chain, and the core elements include the networking method, data dissemination mechanism and data verification mechanism [18].

(1) Networking method
This system uses Peer-to-Peer network (P2P) to organize nodes involved in data verification and bookkeeping in the consortium chain. Each individual node in the P2P network is involved in propagating block data, verifying block data and undertaking network routing without the intervention of any centralized special node, which well reflects the decentralized feature of blockchain.

(2) Data dissemination mechanism
This system uses a tree broadcast mechanism based on TCP/IP communication protocol, where node generates block data and then broadcasts to all other nodes in the whole network through a hierarchical tree structure, spreading the pressure of directly connected nodes to subordinate sub-nodes, effectively increasing the broadcast rate and reducing the number of redundant message packets in the network.

(3) Data validation mechanism
When a node in a consortium chain cluster receives a new message, it will immediately verify the validity of the received data, which includes digital signatures, timestamps, and data legitimacy [19]. Only when the newly generated block data passes the validation of most nodes across the network, it can be credited to the blockchain system, and all nodes in the cluster will save this status information.

3.3. Transaction layer
The transaction layer, which encapsulates incentive mechanisms, smart contracts and consensus mechanisms, is the core in the architecture of the marine environment big data forecasting system.

(1) Incentive mechanism
In the marine environment big data forecasting system, all nodes on the blockchain are already consortium certified nodes, and each node will spontaneously participate in the operation and maintenance of the whole system even if no additional financial incentive is provided. However, in order to increase the motivation of each node to share resources, this system adopts the method of point rewarding. The nodes uploading marine environment data and big data forecasting methods will be rewarded by points. The specific allocation scheme has been described in detail in Section 2.2 of this paper.

(2) Smart contract

Smart contract is the programmable basis of the blockchain system. It specifies the transaction guidelines and processes of the marine environment big data forecasting system. The execution and bookkeeping of smart contracts requires all nodes on the blockchain to follow a unified consensus mechanism to accomplish this, and the smart contract design scheme of this system has been described in detail in Section 2.3 of this paper.

(3) Consensus algorithm

The Big Data Forecasting System for the Marine Environment uses the Practical Byzantine Fault Tolerance (PBFT) for multi-node parallelism. The core theory is given by the following equation.

\[ n \geq 3f+1 \]

In the equation, \( n \) is the total number of nodes in the system and \( f \) is the number of nodes allowed to have faults. It is clear that the prerequisite for the system to reach consensus is that the number of normal nodes on the chain is not less than \( 2f+1 \) [20].

3.4. Application layer

The application layer of the marine environment big data forecasting system encapsulates marine environment big data forecasting modular models, functional components, callable service interfaces, and visual interaction interfaces. Certified nodes of the consortium chain, such as marine departments, research institutes, colleges and universities and the social public, can interact with the system directly.

4. System implementation

In the process of implementing the big data forecasting system for the marine environment, firstly, with implementation of the smart contracts as the core, the business process of the system was studied and developed. Secondly, the main functional components of the system were identified. Based on this, the deployment and application of the system was completed.

4.1. System flow

The business process of the system is shown in Figure 4.

![Flowchart of the big data forecasting system for the marine environment.](Figure 4)
(1) Users logging into the marine environment big data forecasting system first need to verify their identity, and after their account information being verified, can they enter the main interface of the system.

(2) General authority users can select functions of data uploading, data querying or data using in the main interface of the system.

(3) After the user passes in the parameters that meet the requirements, the smart contract will be triggered to retrieve and execute the corresponding transaction protocol.

(4) After the transaction is executed, this user node will broadcast to all nodes on the entire blockchain. The nodes with consensus authority will perform consensus verification, and if the verification is passed, the points of the corresponding value will be rewarded or consumed according to the transaction category.

(5) After the point status changes, all nodes on the blockchain will record the new point status in the underlying ledger for preservation and write it to the log.

(6) After all transactions are completed, the system will automatically update the blockchain status.

4.2. Functional composition
The marine environment big data forecasting system mainly includes five functional modules, which are account management module, data uploading module, data query module, data usage module and status management module. The functional components of the system are shown in Figure 5.

(1) Account management module
The account management module is responsible for consortium chain node account registration, node authority management and user authentication. This system mainly has two kinds of authority, consensus and observation, in which consensus nodes can participate in the whole consensus process of the system, while observation nodes can only synchronize data.

(2) Data uploading module
The data uploading module is responsible for uploading marine environment data, marine environment big data forecasting methods and marine environment forecasting products. The block information will be updated in real time after being uploaded.

(3) Data query module
The data query module is responsible for implementing query operations on marine environment data, marine environment big data forecasting methods and marine environment forecasting products, and outputting the information of the resource directory list stored on the blockchain, which includes resource identification, resource name, cochain time and data description. The query operations do not generate new blocks.

(4) Data usage module
The data usage module can meet the user's demand for the use of marine environment data, marine environment big data forecasting methods and marine environment forecasting products stored on the blockchain. The system supports the use of the service interface to invoke the data on the chain.

(5) Status management module

The status management module is responsible for implementing functions such as account points management, system log management and blockchain status query to support monitoring and tracing of data.

4.3. System applications

Based on the Gradle architecture, the marine environment big data forecasting system is developed with an interactive front-end interface. Based on the FISCO WeBase platform, the system is developed and managed with back-end smart contract, as well as blockchain groups and nodes. Figure 6 shows the frontend data query interface and the backend blockchain state monitoring interface of the marine environment big data forecasting system.

![Figure 6. Interface of the big data forecasting system for the marine environment.](image)

This system builds a consortium chain in the ocean information and communication network, which can fully mobilize all sea-related departments to participate in data sharing, at the same time promote the circulation and use of precious research resources such as marine environment data and marine environment big data forecasting methods in the whole ocean industries. Now the system has been applied in the Zhejiang Smart Ocean Data Center Project. The data, methods, and forecast products provided by marine environment big data forecasting system have effectively supported the research of local marine research institutions and improved the marine environment forecasting capability in Zhejiang Province.

5. Conclusions

Combining the current situation and bottlenecks of data sharing in the field of marine forecasting in China, the author develops a marine environment big data forecasting system based on blockchain technology like FISCO consortium chain, incentive mechanism and smart contract, which realizes the decentralized management and exchange of marine environment data, marine environment big data forecasting methods and marine environment forecasting products. The system provides new methods and ideas to realize the value interconnection of academic resources in marine field, and is a useful exploration of the deep integration of blockchain technology and marine industry. Although it is still relatively difficult to promote the sharing of marine resources in China due to the restrict of social mechanism and policy, the successful application of the marine environment big data forecasting system in the project of Zhejiang Ocean Big Data Center has technically brought a new breakthrough for marine data sharing and is able to be further promoted for national and local use. It is foreseeable that with the continuous emergence of blockchain applications in marine industry, blockchain technology will bring great changes to the field of ocean information technology in the future.
References

[1] Yang J K and Miao Q S 2020 Research and development of marine big data analysis and prediction technology[J]. Management and Research on Scientific & Technological Achievements 15(1) 62-64

[2] Xu L B, Wang A X and Wang C Y 2020 Research on correction method of marine environment prediction based on machine learning[J]. Marine Science Bulletin (6) 695-704

[3] Beck R 2018 Beyond bitcoin: The rise of blockchain world[J]. Computer 51(2) 54-58

[4] Zhao Z 2018 Research and design of digital archive management system based on blockchain[D]. Hefei: University of Science and Technology of China

[5] Azaria A, Ekblaw A, Vieira T, et al. 2016 MedRec: Using blockchain for medical data access and permission management Proc. of the Int’l Conf. on Open and Big Data. IEEE 25-30

[6] Tsai W T, Yu L, Wang R, Liu N and Deng E Y 2017 Blockchain application development techniques Journal of Software 28(6) 1474-1487

[7] Xue T F, Fu C Q, Wang Z and Wang X Y 2017 A medical data sharing model via blockchain Acta Automatica Sinica 43(9) 1555-1562

[8] Shen X, Pei Q Q and Liu X F 2016 Survey of block chain[J]. Chinese Journal of Network and Information Security 2(11) 11-20

[9] Liu Y 2020 Research and implementation of intellectual property asset management and trading system based on blockchain[D]. Xi’an: Xi’an University of Electronic Science and Technology

[10] Yin T and Zhao S J 2021 Research on key technologies and applications of blockchain[J]. Economic Observation 8 43-45

[11] Li Y, Gao N, Sun J, et al. 2021 Research and exploration of big data transaction model based on blockchain[J]. Big Data 7(04) 37-48

[12] Zhang L, Liu B X, Zhang R Y, et al. 2019 Overview of Blockchain Technology[J]. Computer Engineering 45(5) 1-12

[13] Wang Z Z, Zhao X L, Peng Y, et al. 2020 Primarily Research for Sharing Mode ments Based on Blockchain Technology[J]. Advances in Marine Sciences 7(2) 25-30

[14] Cheng J C, Zhang C and He A Y 2020 Application research of blockchain technology in cross-departmental ocean data sharing[J]. Science & Technology Review 38(21) 60-68

[15] Zhuang C X 2020 Design and implementation of a collaborative management system for pharmaceutical supply chain based on blockchain[D]. Guangzhou: Guangdong Normal University of Technology

[16] Wei Y Z, Deng Z H, Guan Y R, et al. 2021 A Method of Secure Provenance of Scientific Data Based on Blockchain and Smart Contracts[J]. Journal of Modern Information 41(1) 32-38

[17] Han L Y, Liang J F, Song X, et al. 2021 Application and prospect of blockchain technology in marine fields[J]. Coastal Engineering 40(01) 76-82

[18] Lin S Y, Zhang L and Liu D S 2021 Survey of application research based on blockchain-smart-contract[J]. Application Research of Computers 38(09) 2570-2581

[19] Cui J D, Wang S W and Xin Y C 2020 Research on Technical Framework of Smart Grid Data Management From Consortium Blockchain Perspective[J]. Proceedings of the CSEE 40(3) 836-848

[20] Zhu Y, Wang Q S, Qin B H, et al. 2019 Survey of blockchain technology and its advances[J]. Chinese Journal of Engineering 41(11) 1361-1373