Medicine®

OPEN

# Research on medical system based on blockchain technology

Qu Jia, EngD*

**Abstract**

Medical data sharing, anti-tampering, and leakage prevention have always been severe problems that plagued the pharmaceutical industry. When a patient is referred, he often cannot provide information about previous visits because the medical information of each hospital cannot be shared in most cases, but only through Medical records, test sheets, and other easily lost paper information are used to share some medical information. At the same time, patient medical information is easily leaked, and the medical information provided in the event of a medical dispute cannot guarantee authenticity and impartiality. This article designs a consortium medical blockchain system based on a Possible Byzantine Fault Tolerance algorithm. This system is a medical system that is maintained and shared by multiple nodes and can prevent medical data from being tampered with or leaked. It can be used to solve these medical problems. Compared with the existing medical blockchain system, this system has certain advantages and better applicability.

**Abbreviations:** AFS= auditing federate servers, CIS = clinic information system, DPOS = Delegate Proof of Stack, EMR= electronic medical record, HIS = hospital information system, LIS = laboratory information system, MIFS = medical institution federate servers, PACS = picture archiving and communication systems, PBFT = Possible Byzantine Fault Tolerance, POI = proof of information, POS = proof of stack, POW = proof of work, RBAC = role-based access control.

**Keywords:** consensus algorithm, medical blockchain, medical data, medical information sharing

## 1. Introduction

Medical information is valuable information for patients. However, in the current medical systems of various hospitals, most of this information cannot be used interchangeably, which leads to the need for new medical care for each patient to record a patient's medical information. Data can often only be obtained through vague memory. Although most hospitals will use paper medical records, paper medical records are straightforward to be damaged or lost, which is a very unreliable medical information recording method. On the other hand, using traditional Databases to achieve medical information sharing is often leaked due to reselling by some unethical staff, resulting in further losses to patients. Therefore, medical staff and patients urgently need a way to achieve medical information sharing between hospitals and to ensure that a system where patient information will not be leaked, and blockchain is currently a great way to implement this system.

Blockchain is a distributed database system with multiple independent nodes,[1] which can safely store Bitcoin[2] transactions or other data, and ensure the security of these data or information, preventing tampering and forgery Blockchain is generally deployed in P2P networks, which is different from common relational databases and nonrelational databases. Blockchains use digital signatures, hash algorithms, and other encryption algorithms and distributed consensus algorithms-tampering, destroying, or erasing database operation logs. Blockchain technology has the characteristics of decentralization, time-series data, collective maintenance, programmable and secure, and reliable.[3]

According to the different participants, the blockchain can be divided into a public chain, alliance chain, and private chain. The participants of the public chain can be anyone; all those who want to participate in the maintenance of the public chain can join and serve Bitcoin. The blockchain is a public chain. A private chain refers to a blockchain that is used internally by an entity and whose information is not public. The entities here can be companies, banks, hospitals, etc., which are currently being researched by domestic banks. Most of the chains are private chains. The alliance chain is a blockchain composed of multiple entities and with access restrictions. Compared with the public chain, the alliance chain is not arbitrarily joined but requires specific permission before it can be accessed, and the stored information access rights are restricted by these entities, which can be disclosed to the outside world only under certain conditions. Compared with private chains, the difference between the alliance chain is that the participating entities are multiple different companies or groups. These entities jointly

Editor: Bernhard Schaller.

The author received no financial support for the research, authorship, and/or publication of this article.

The author has no conflicts of interest to disclose.

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Information Center, Chengde Petroleum College, Chengde, China.

* Correspondence: Qu Jia, Chengde Petroleum College, Chengde, He bei 067000, China (e-mail: 15696966@qq.com).

How to cite this article: Jia Q. Research on medical system based on blockchain technology. Medicine 2021;100:16(e25625).

Received: 8 April 2020 / Received in final form: 14 September 2020 / Accepted: 1 April 2021

http://dx.doi.org/10.1097/MD.0000000000025625

maintain the blockchain and share the information in the blockchain.

The entities in the medical blockchain are hospitals and medical institutions. These entities are entirely independent in terms of administration and finance. They are different entities. At the same time, these entities are subject to government supervision and management and have strict admission and classification systems. There are certain access restrictions. Medical data is not only a patient's privacy but also involves state secrets, so access is strictly restricted. According to the above characteristics, it can be seen that the medical blockchain is an alliance chain.

Most of the current medical blockchain systems use the POX consensus algorithm to reach a distributed consensus. The POX algorithm currently mainly includes proof of work (POW), proof of stack (POS), and delegate proof of stack (DPOS). Practical Byzantine fault tolerance( PBFT) The problem to be solved by the blockchain consensus algorithm is the Byzantine Generals problem.[4] The reason why this problem is difficult to solve is that there may be multiple proposals in the system at any time, and it is tough to complete the final consistency confirmation. The POW algorithm is generally used. The public chain requires more nodes and higher computing power to maintain.[5] The process of generating blocks by the POS algorithm is determined by the digital currency held by the node.[5] The DPOS algorithm requires the holder of the digital currency to select a certain number of block generators, and block generators will be reelected every once in a while.[6] They do not meet the needs of medical blockchains. The medical blockchain does not require large computing power to maintain and does not require the generation of electronic money, and the number of nodes is small and flexible. This paper uses the Practical Byzantine fault tolerance (PBFT) algorithm for the first time[7] to construct a consortium medical blockchain that can start and run with fewer nodes and does not require a lot of computing power to maintain.

Castro and Liskov proposed the PBFT algorithm[7] in 1999 to solve the problem of the original Byzantine fault tolerance algorithm's inefficiency. Compared with the innovative Byzantine fault tolerance algorithm, the algorithm's complexity has been reduced from exponential to polynomial level,[7] making the Byzantine fault tolerant algorithm (PBFT) that can be used in practical applications.

Byzantine General Problem is a famous and intractable problem in distributed systems.[8] Another counterpart is Crash Failure Problem, which is simpler and more common. The Crash Failure Problem assumes that all nodes are honest. By contrast, the Byzantine General Problem implies a situation that there may be dishonest nodes in a distributed computing system. Specifically speaking, a dishonest node can send different or even contradictory messages to other nodes, aimed to prohibit a system from reaching a consensus or reach a false consensus. In the Byzantine General Problem, a system needs to reach information consensus among honest members and dishonest members. As a solution to solve the problem, PoW is used in various blockchain systems, such as Bitcoin and Litecoin. However, PoW requires high computational power to maintain correctness of consensus.

The PBFT algorithm is a consistency algorithm based on state machine replication. The service acts as a state machine and replicates in different nodes of a distributed system. Each copy of the state machine saves the state of the service and the operations implemented. This algorithm can ensure the regular operation of the system when the proportion of nodes with errors does not exceed one-third of the total number of nodes. The idea is to let every node that receives a message asking about the content of the message received by other nodes.

Compared with the PoW algorithm, PBFT algorithm[9] is more lightweight and effective. It ensures correct consensus decision if the number of malicious nodes is less than one-third of total nodes. The workflow of PBFT algorithm can be divided into a succession of views. Three phases are involved in a view to commit a request: pre-preparation, preparation, and confirmation. In each view, there is only one node that can be selected as the primary, and other nodes are called backups. In the pre-preparation phase, primary node broadcasts the pre-preparation message to each backup node. If a backup node accepts the pre-preparation message after verification process, it enters the preparation phase and multicasts the preparation message to all other nodes. The verification process mainly compares messages from different nodes, and it is considered valid if a node receives messages from more than two-thirds of total nodes and these messages contain a consistent data. Similarly, once a node (both the primary and backups) accepts the preparation messages, it enters the confirmation phase and broadcasts the confirmation message to all other nodes. Once the collected confirmation messages are considered valid, the node will give response to client. The client will make final decisions based on all the collected responses.

The consensus process of the PBFT algorithm is mainly divided into 3 stages: pre-preparation, preparation, and confirmation.

(1) pre-preparation stage:
After the master node receives the service request message and verifies it is correct, it generates a pre-preparation message according to the service request message and broadcasts it to the slave nodes.

(2) Preparation stage:
After receiving the pre-preparation message from the master node, the slave node verifies the message content to ensure that the message content has not been tampered with during transmission. After the content is verified correctly, the slave node will generate a preparation message according to the preparation message and broadcast it to all replica nodes.

(3) Confirmation stage:
When a node receives at least (2n+1)/3 preparation messages from different nodes (including itself), and the verification messages are correct and valid, the node enters the confirmation phase, generates a confirmation message based on the preparation message, and broadcasts it to all replicas node. At the same time, it will continue to receive and verify confirmation messages from other nodes. When it receives (2n+1)/3 valid confirmation messages (including its own), it says that the request has reached the committed state on this node. At this time, only through this node, it can be judged that the requested service has been verified by most of the replica nodes.

When the request reaches the committed state, the request can enter the commit phase, and then the request is executed by all replica nodes.

The primary 3-stage data transmission process (pre-preparation, preparation, confirmation) in the PBFT consensus process is shown in Figure 1.

At present, blockchain is mainly in finance, and there are relatively few applications in healthcare because the focus of

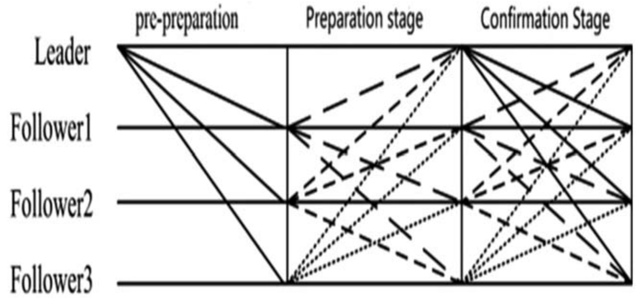


Figure 1. The 3 main stages in the PBFT consensus process.

blockchain is limited to digital currency blockchain systems such as Bitcoin. Xue Tengfei et al[10] made use of an improved DPOS consensus mechanism that proposes a medical blockchain system MDSM combining medical institution federate servers (MIFS) and auditing federate servers (AFS). Azaria et al[11] use Ethereum Blockchain, which realizes a medical information sharing platform MedRec.[11,12] Ivan,[13] which combines medical blockchain with big data, analyzes the use of blockchain as a novel method to protect medical health data storage, implementation obstacles, and A plan for the gradual transition of current technology to blockchain solutions. Shrier et al[14] used the combination of the OPAL/Enigma encryption platform of MIT and blockchain technology to create storage and analysis of healthcare information, a secure environment. Kuo et al[15] adopted a combination of privacy protection, online machine learning, and private blockchain technology. Witchey[16] introduced medical transactions single (transaction) verification system and method. It can be seen that there are relatively few applications and researches on the blockchain in the medical field, and most of them are at the application level.

In the existing medical blockchain system, the consensus algorithms used to belong to the POX series of algorithms. Among them, the medical blockchain researched by Xue Tengfei requires 121 hospitals or medical institutions to participate in the maintenance of the blockchain at the same time. Therefore, the startup cost is relatively high, and it is not suitable for the gradual research process from early exploration to later large-scale mature application. The consensus mechanism adopted by the Ethereum blockchain researched by Azaria is the POW algorithm. Everyone on the Internet can participate in or withdraw from the maintenance process at any time, which is a waste of computing power, and each operation needs to pay a certain token as a reward. It is not suitable for the use of medical blockchain. ModelChain[15] is not a blockchain designed specifically for medical care. Its consensus algorithm proof of information (POI) combines machine learning with a proof-of-work algorithm. The power will be greater, so it is not suitable for medical blockchain.

The PBFT algorithm only needs 4 or more nodes to start. Compared with the POX algorithm-based blockchain system, the PBFT algorithm-based blockchain system has a small startup cost, is suitable for early exploration and later expansion, and does not require a large amount of computing power to maintain. This article will use the PBFT consensus algorithm to implement a blockchain suitable for medical systems.

## 2. Methods

The blockchain system in the medical chain mainly includes 3 parts: storage management, node management, and user management. Storage management refers to how to store medical data on the blockchain logically, and how the blockchain is stored in this kind of storage device. Node management is the management of each node running the blockchain system. User management refers to the authentication and authority management of participants in the medical chain.

### 2.1. Blockchain storage management

Blockchain storage management mainly includes the management of blocks, transaction orders, and medical data storage. It is the most fundamental component of medical blockchain.

***2.1.1. Medical blockchain and medical block.*** The medical blockchain is mainly composed of 2 parts: Block and Transaction. A blockchain consists of blocks that record the previous block ID, and each block contains several transaction orders. These transaction orders are actual A carrier for storing Blockchain data. For example, a blockchain can be regarded as a database, each block constituting the blockchain can be viewed as a table in the database, and a transaction order can be regarded as each table A record on the (Record). The composition of a blockchain is shown in Figure 2

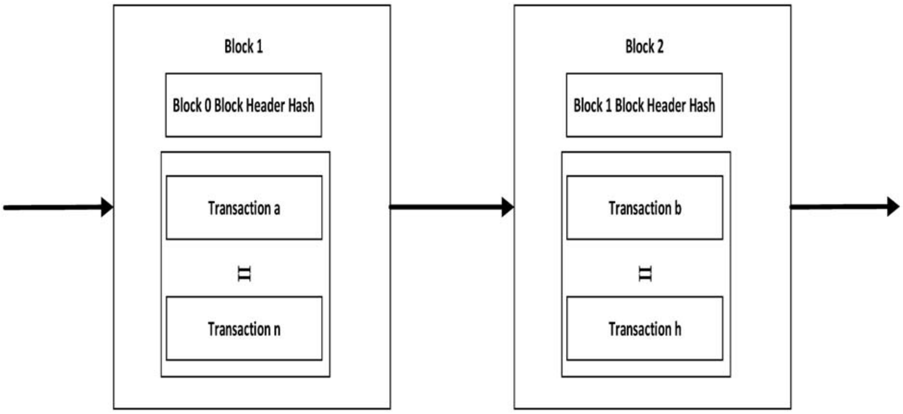


Figure 2. Composition of a blockchain.

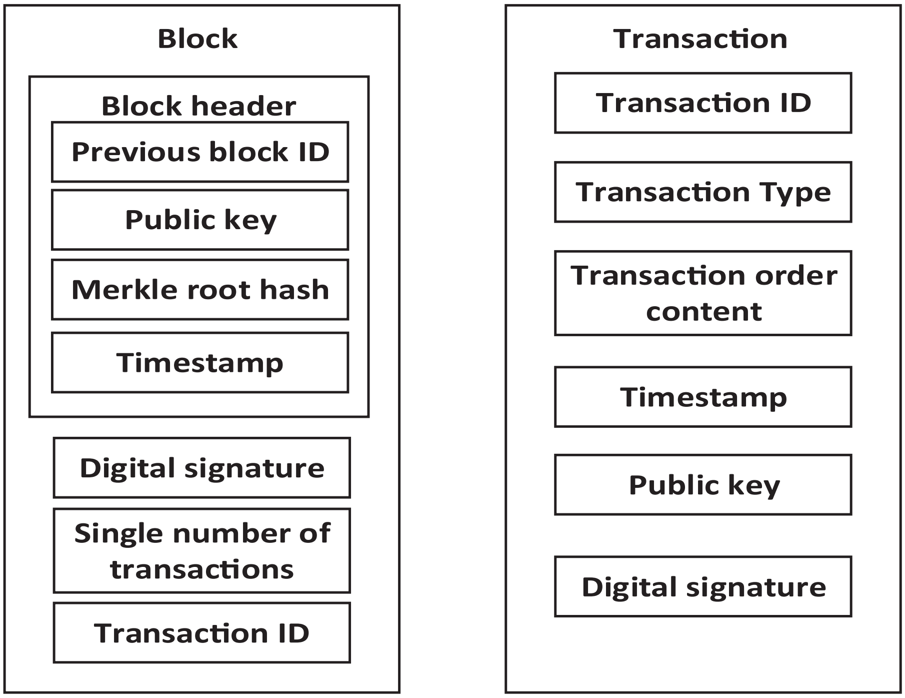


**Figure 3.** Composition of block and transaction.

The specific structure of each block is shown in Figure 3. A block is mainly composed of a block header and content other than the block header. The block header indicates the part that needs to be digitally signed. The block header contains the ID of the previous block, the public key of the block generator, the Merkle tree root hash value generated by the transaction ticket ID, and the timestamp of the created block. The content other than the block header includes the block generator's digital signature for the block header, the transaction ticket ID number, and all transaction order IDs stored in this block. The digital signature is to ensure that the content of the block cannot be tampered with and to ensure that the block producer cannot be denied after generating a malicious block. Also, only the block is saved in the block. The ID of the transaction order, that is, only the index pointing to a particular transaction order, is protected, but not the transaction order itself, which can reduce the capacity of each block and facilitate synchronization and backup. Blocks and transaction orders are physically stored in the database; it is logically stored in the form of a blockchain. In the transaction order design storage, the transaction order ID, transaction type, timestamp, public key, and number are added to the data generally stored in the database. Signature single transaction information field, the information to be stored as single transaction content, is formed on a single logical transaction, the data is generally not much difference between its storage on physical storage.

The content of each transaction ticket is shown in Figure 3. The transaction ticket type indicates the type of transaction, such as adding, deleting, querying, and modifying, to instruct the validator to perform the corresponding operation. For the introduction of the validator, see Section 2.2.3. There are 2 reasons for using transaction tickets instead of directly accessing the nodes that own the blockchain when adding, deleting, and modifying operations. First, a node may conduct malicious operations, expose patient information in violation of regulations, or tamper with information, etc. On the other hand, to record the operator's operation in the blockchain, the operator needs to use his private key to digitally sign the transaction ticket when operating, so that the operator cannot deny the operation he has performed. The content of the transaction order is the content stored in the transaction order, such as the medical information of the patient. The timestamp indicates the time when the transaction order was generated, the public key is the public key of the transaction order generator, and the transaction ID is the type of the transaction order, the content of the transaction ticket, the timestamp and the hash value generated by hashing the public key. The hash algorithm and encoding algorithm can choose the SHA-256[17] hash algorithm or BASE64[18] encoding algorithm. Its reliability has been verified in various blockchain systems. A digital signature is the signature of the transaction order generator to the transaction order ID, preventing the transaction order from being tampered with.

### 2.1.2. Medical information transaction slip.

The content stored in the transaction order includes patient information, doctor information, medical record information, information of each node, etc. That is, the transaction order is the carrier of each data record in each table of the traditional database, and the content of the transaction order is equivalent to each record. The contents of transaction orders are mainly as follows:

Entity information category: It is mainly used to record the detailed information of entities such as patients, medical personnel, such as the patient's ID number, name, gender, age, marital status, contact information, and other personal information, and the key held by the patient public essential details.

1. Medical information: It is mainly used to record the relevant medical information of the patient. If a patient "P" arrives at the hospital "H" to accept the doctor "D" at a specific time, an

outpatient record is generated, mainly including time, place of consultation, particular conditions of the consultation, etc. If the patient has a picture or video check like B ultrasound, the generated image or video is hashed to obtain a hash value and stored in the transaction slip.

2. Entity-information association information: This type of information is mainly used to associate entities with medical information or other sensitive information because this type of data requires encryption operations to prevent the entity's privacy from leaking.
3. Add, delete, update, and query classes.
4. Permission category: The transaction type in the transaction ticket is "permission," and the content of the transaction ticket is specific permission information.

The contents of the above 5 types of transaction orders can ensure that the patient's privacy is not violated, and various medical data generated during the patient's consultation can be found in time after being tampered with. At the same time, it is also possible to add previous medical data to the blockchain.

***2.1.3. Medical data storage.*** The storage of medical blockchain needs to be combined with medical information systems, that is, digital hospitals to coordinate storage arrangements. Digital hospitals refer to the use of the computer, network, database, and other information technologies to organically combine hospital business information and management information to achieve text, image, and voice hospital information system for the digital collection, storage, reading, and retrieval of information such as data, charts, and diagrams. Its main components include hospital information system (HIS), clinic information system (CIS), picture archiving and communication systems (PACS), laboratory information system (LIS), electronic medical record (EMR), etc.

HIS is an application information system that automatically collects, processes, stores, transmits, and utilizes relevant information inside and outside the hospital using computers and their network communication equipment and technologies.[19] CIS is an application. The information system in the clinical treatment process mainly includes doctor workstation system, nurse workstation system, blood transfusion management system, surgical anesthesia tube system, and clinical decision support system. PACS is a system used to manage medical images. LIS refers to the use of computer technology, network technology, a software system that realizes the collection, storage, processing, transmission, and query of clinical laboratory information, and provides analysis and diagnosis support. EMR electronic medical records are created, stored, and used electronically by medical institutions. The data integration system of clinical diagnosis and treatment and guidance intervention information for inpatients (or health care objects) is a complete and detailed clinical information resource generated and recorded by individual residents in the medical institution. The relationship between them is shown in Figure 4.

### 2.2. Node management

The most critical nodes of a blockchain system are the validator used to verify the correctness of the Transaction and Block, the generator used to generate the transaction order, and the block Blocker used to create the block. They work together under the specification of the consensus algorithm. Take Bitcoin as an example; its mining client will receive all transaction orders on the Internet, verify it, calculate random numbers, generate blocks, and broadcast them to the entire network. Similar to the Bitcoin blockchain components, the nodes in the chain system also

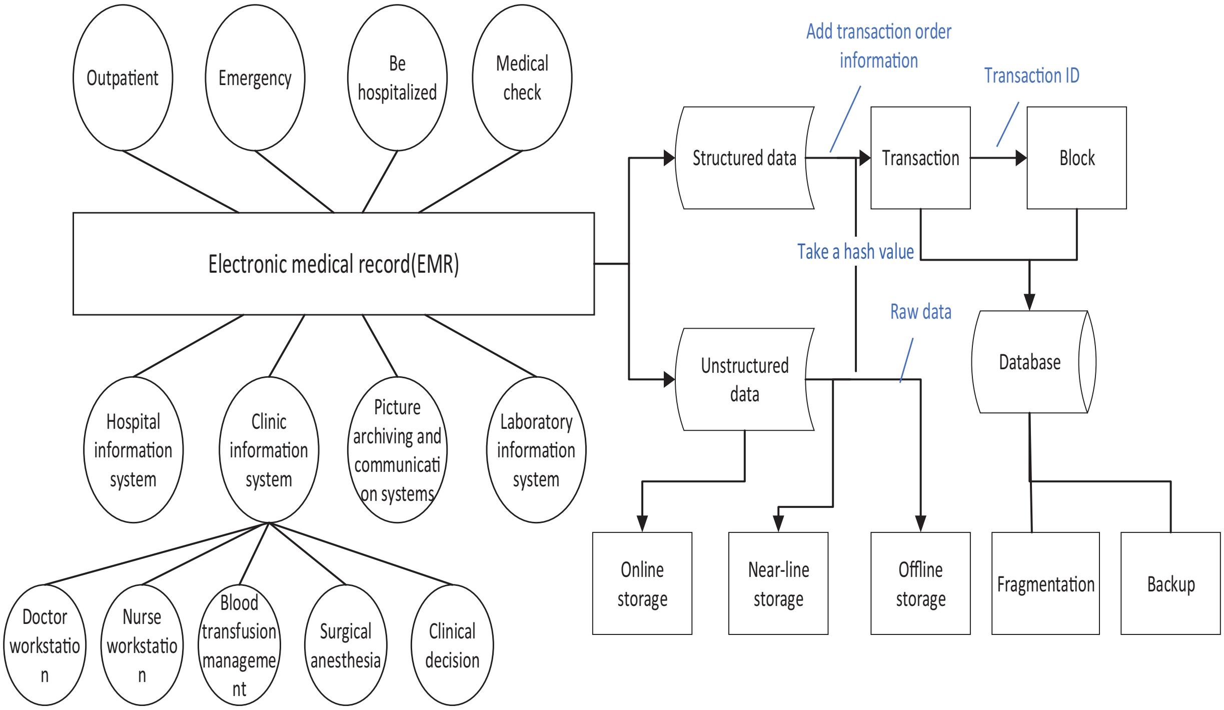


**Figure 4.** Medical chain data storage architecture.

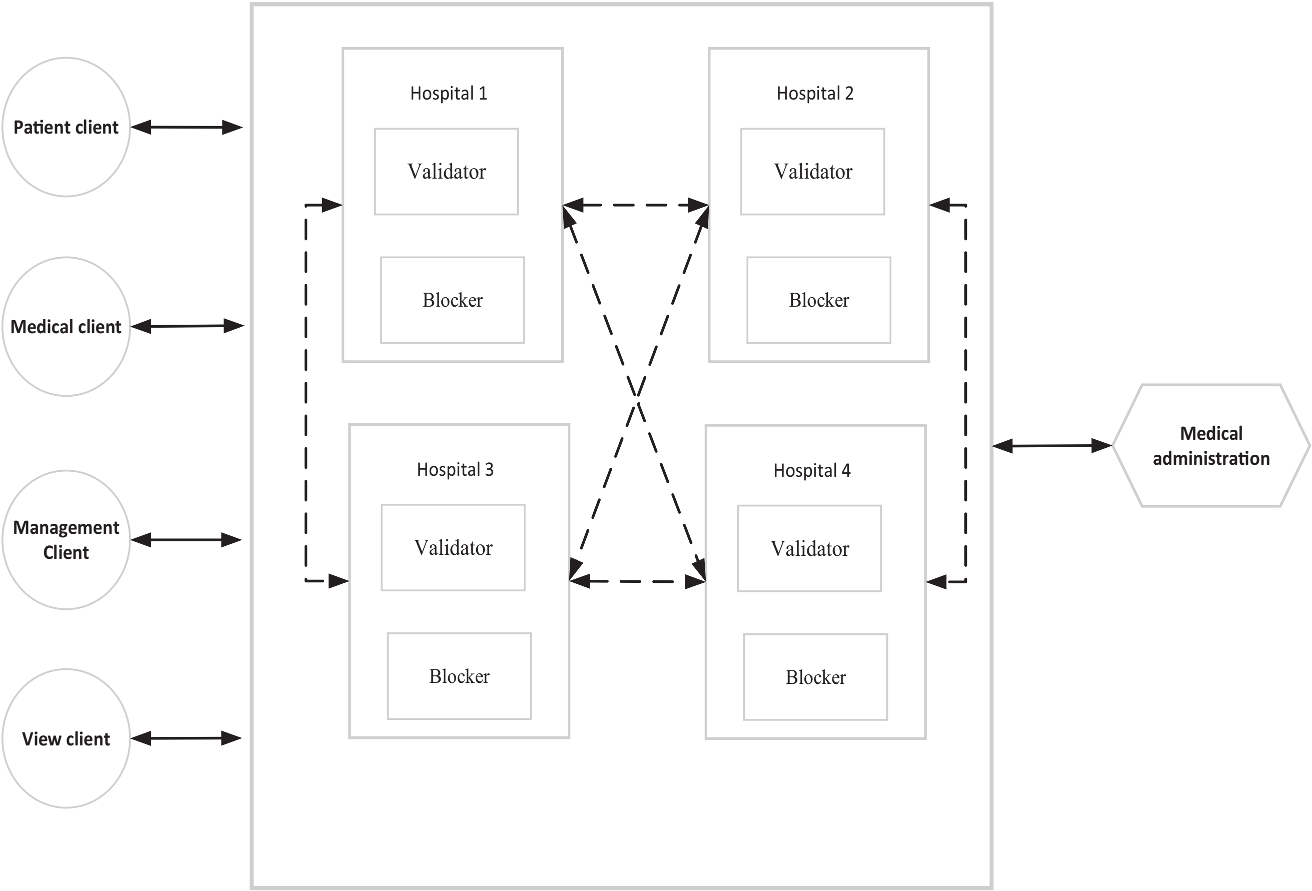


**Figure 5.** Medical chain node management.

mainly have 3 parts: client, Validator, and Blocker, as shown in Figure 5.

*2.2.1. Consensus algorithm.* The PBFT algorithm is used as the consensus algorithm in the medical blockchain because the PBFT algorithm is a consensus algorithm suitable for the alliance chain. Its advantages and advantages are:

1. The PBFT algorithm does not need to rely on a large amount of computing power to avoid the "51% attack" like the POW algorithm, nor does it need to rely on tokens as a standard to measure voting rights like the POS algorithm or DPOS algorithm. In the case of (n-1)/3 nodes error (data loss, nonoperation, etc.).
2. As a kind of Byzantine fault tolerance (BFT) algorithm, PBFT algorithm can guarantee the normal execution of a distributed consensus process when there are less than or equal to (n−1)/3 faults or malicious nodes in the system,[20] This requires that the nodes in the network using the PBFT algorithm have at least (2n + 1)/3 normal nodes in each consensus process, so the environment in which these nodes operate must be relatively safe and stable.
3. The medical blockchain is an alliance chain. The entities participating in the medical blockchain are endorsed by the government, have certain credibility, and are strictly supervised by the health management department. The occurrence of malicious behavior is far less than that in areas such as Bitcoin. At the same time, after years of information development, each hospital has a relatively complete network, server, and database system. Therefore, the existing medical system can provide a relatively safe and stable operating environment for the regular operation of the PBFT algorithm. At the same time, each node in the cluster running the PBFT algorithm is equal in status, there are no high or low voting rights, and it avoids the centralization of the medical blockchain system when verifying transaction orders or the blockchain. Therefore, the PBFT algorithm is very suitable for medical blockchain.

There is currently no medical blockchain system that uses the PBFT algorithm. This article uses this algorithm innovatively and proposes an alliance medical blockchain system that is very suitable for the medical field.

*2.2.2. Client.* The client is a component used to generate transaction orders, and its main function is to add, delete, modify, and check. The difference is from common databases, such as MySQL. The deletion and modification operations here are not directly deleting the corresponding transaction order from the blockchain, but regenerate a new transaction order, overwriting the original. This is because the blockchain uses a digital signature and a block to record the ID of a previous block to ensure that the existing content in the blockchain cannot be modified and deleted. And the content covered can be traced. Clients are mainly divided into 3 categories:

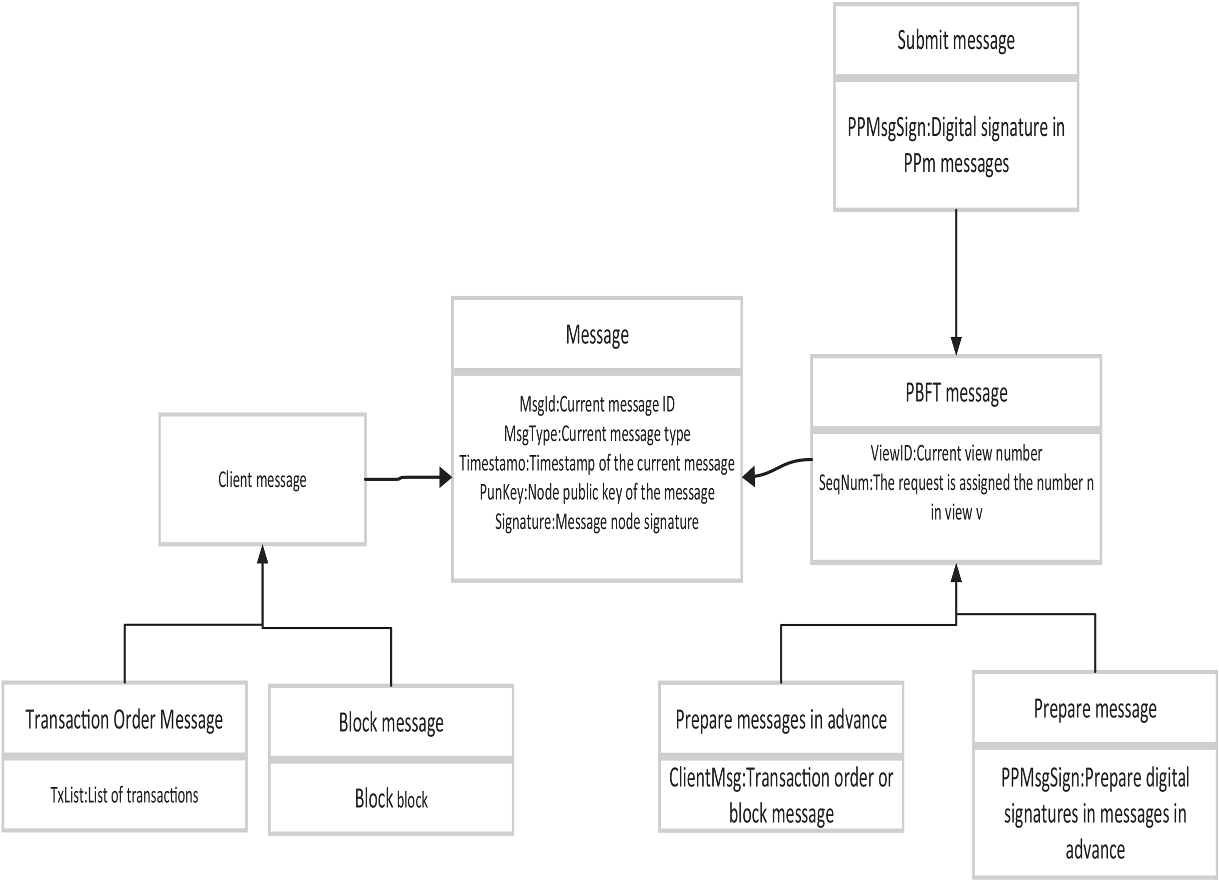

Submit message
PPMsgSign:Digital signature in PPm messages
Message
MsgId:Current message ID
MsgType:Current message type
Timestamo:Timestamp of the current message
PunKey:Node public key of the message
Signature:Message node signature
PBFT message
ViewID:Current view number
SeqNum:The request is assigned the number n in view v
Client message
Transaction Order Message
TxList:List of transactions
Block message
Block block
Prepare messages in advance
ClientMsg:Transaction order or block message
Prepare message
PPMsgSign:Prepare digital signatures in messages in advance

**Figure 6.** Medical chain message structure.

1. Patient client: The primary function of the patient-client is to add, delete, modify, and check. In terms of addition, patients can add authorized information, authorize doctors or others to add, delete, change, and check their medical information. In terms of deletion, patients can delete related authorizations information.
2. Doctor client: The primary function of the doctor-client is to add, delete, modify, and check. In terms of addition, doctors can add medical records to patients, such as outpatient records. In terms of deletion and modification, doctors can delete within the patient authorization period. There are incorrect medical records, or there are problems with modifying patient medical records.
3. Query client: The query client only has the query function. Some institutions may need to check the relevant information of patients.

***2.2.3. Validator.*** The validator is the replica node in the PBFT algorithm. The validator is mainly responsible for receiving the transaction order message from the client and the block message from the packer for verification. The 2 message structures are shown in Figure 6. After receiving the transaction order message or block message sent from the client, first verify that the digital signature of the transaction order message and the transaction order itself is correct, and then verify that the transaction order meets the requirements. For example, a patient cannot obtain the medical records of another patient without authorization from another patient. After the verification is completed, the master node validator generates a pre-prepared message and adds the transaction order message to the pre-prepared message. During the 3-stage process of PBFT, each node accepts the transaction order and stores the transaction order in its database. The verification node implements the 3-stage procedure of PBFT as follows:

1. The master node assigns a view editor v and a sequence number n to the client message based on the verified client message, generates a preread message, and broadcasts it to each backup node after generation.
2. After receiving the prepared message, each backup node verifies the digital signature of the prepared message, the view number v and the prepared message sequence number n, saves it after verification, and generates the prepared message based on the prepared message. The format of the prepared message is as follows: As shown in Figure 5, after each backup node generates a preparation message, it broadcasts the preparation message to all nodes except itself (that is, the master node and the backup node).
3. After receiving the preparation message, each node verifies the digital signature of the preparation message, the view number v and the number of the prepared message n, and saves it without error. When a replica node accepts $(2n+1)/3$ view number v, it is the same as the pre-prepared sequence number n, and generates a commit message according to the prepared message, and broadcasts the commit message to all replica

nodes except itself. The content of the commit message is shown in Figure 6.

4. After receiving the submission message, each node verifies the digital signature of the submission message, the view number v and the prepared message number n, and saves it without error. When a node accepts (2n + 1)/3 view number v, and if the pre-prepared serial number n is the same, the transaction order message or block message carried in the pre-prepared message serial number n is received. The operation after acceptance is to save the transaction order or block or perform the operation carried in the transaction.

***2.2.4. Blocker.*** Blocker is mainly used to collect transaction IDs, generate Merkle trees,[21] package them into blocks, and send them to the validator. After the 3-stage process of the PBFT algorithm, they are added to the blockchain. Blocker needs to obtain the current ID of the last Block of the blockchain, so only one block is generated in a period of time. In the Bitcoin blockchain, a block is generated every 10 minutes, and the Block is generated by a "miner." The work, the miner, does is mining. The so-called mining is that the mining software continuously generates a random number and performs an SHA-256 hash operation with the block header related content to get a hash value. If the hash value is less than a given threshold, then the miner successfully mines, generates a block broadcast to the entire network, and gets a reward, namely Bitcoin. In this blockchain system, every Block generated must be verified by a validator. If a malicious zone appears Block, the validator will be found during verification, and will not be accepted. In addition, the blockchain is deployed in various hospitals, and its server and network environment are relatively stable. There will not be a host at all times like Bitcoin. Join and exit, and service a dedicated administrator management, where the malicious operation is relatively small, considering the 2 cases above, the system generates a chain block by Block in the following manner:

1. Validator and Blocker are first deployed on a hospital node. A hospital node contains several servers and databases. The reason for deployment is that the verified transaction order information is saved in the database of each hospital node, reducing unnecessary network transmission. Hospital nodes are shown in Figure 5.
2. After each hospital node verifies and accepts several transaction orders within a period, determine whether the current block is generated by itself according to B=L% (N −1). Among them, B is the node that currently needs to generate a new block, L is the length of the current blockchain, and "%" is the remainder operation. Considering that the master node needs to receive messages from the client, the master node does not participate in the packaging process to achieve load balancing of the tasks of each node.
3. If the hospital node detects that it needs to generate a block by itself according to the formula in 2., it collects a certain amount of verified transaction order IDs, generates a Merkle tree, packs it into blocks, and sends it to the master node. After PBFT algorithm 3 after the staging process, the block is added to the end of the blockchain of each node. If the contents of the transaction order have tampered, the value of the transaction order ID after hashing will be different from the transaction order ID. Just store the transaction ID.

Through the above 3 components, it can be ensured that when a malicious or faulty node is less than or equal to (n−1)/3 in a certain consensus process, the consensus process can still be completed normally. The correctness verification process can be found in reference.[20]

### 2.3. User management

User management is mainly used to manage the accounts, keys, and permissions of users participating in Medical chain. It is a module that implements identity authentication and access control.

***2.3.1. Account management.*** Account management is mainly used to manage user login, logout, password retrieval, and public key binding. Authentication is bound to individuals and is used to manage users' information.

***2.3.2. Key and authentication architecture.*** In the Medical chain, to ensure the confidentiality, integrity, and validity of the blockchain system, cryptographic technologies such as asymmetric encryption, digital signatures, and public key infrastructure (PKI) authentication systems are used. Confidentiality refers to the process of data transmission. It cannot be seen by unauthorized persons. Integrity means that the data will not be tampered with during transmission. Validity means that the data generated by the participants of the blockchain system cannot be denied. The key and authentication architecture provides the previous functions such as the generation, backup, and authentication of public keys used in this section.

***2.3.3. Rights management.*** The rights management function is an essential part of medical blockchain. It is directly related to the security of the medical blockchain and whether it can adequately protect the privacy of patients. The participants of the medical blockchain are roughly divided into 3 categories, including health management departments, medical institutions, and medical service recipients. Health management departments in China are mainly divided into national, provincial, and municipal levels. Medical institutions are primarily composed of medical staff and managers, and medical staff is provided by medical institutions. Medical service personnel, such as doctors and nurses, and management personnel, are those who maintain the regular operation of the hospital, such as workforce and finance. Medical service recipients include patients and their families. Because of this feature, the medical chain uses role-based access control (RBAC) to implement rights management. RBAC maps users to roles, and users enjoy permissions through characters. The model defines dynamic roles or static relationships by defining different roles, inheritance relationships between roles, relations between tasks, and corresponding restrictions. Standardize user behaviors.[22] The roles in the medical chain are divided into 3 categories: health management departments, medical institutions, and medical service recipients shown in Figure 7.

Different from the general authorization system, the permission information in the medical chain is stored in the blockchain, which is extremely difficult to be tampered with. Besides, the smart contract method is used to perform addition, deletion, modification, and check operations. Intelligent contracts are provided by Szabo.[23] Proposed in 1995, his definition of a smart contract is: A smart contract is a series of digitally specified commitments, including an agreement for each party to fulfill these commitments. In simple terms, a smart contract is a piece of code posted on the blockchain. When the terms in the agreement are triggered at a particular time, the code will be automatically

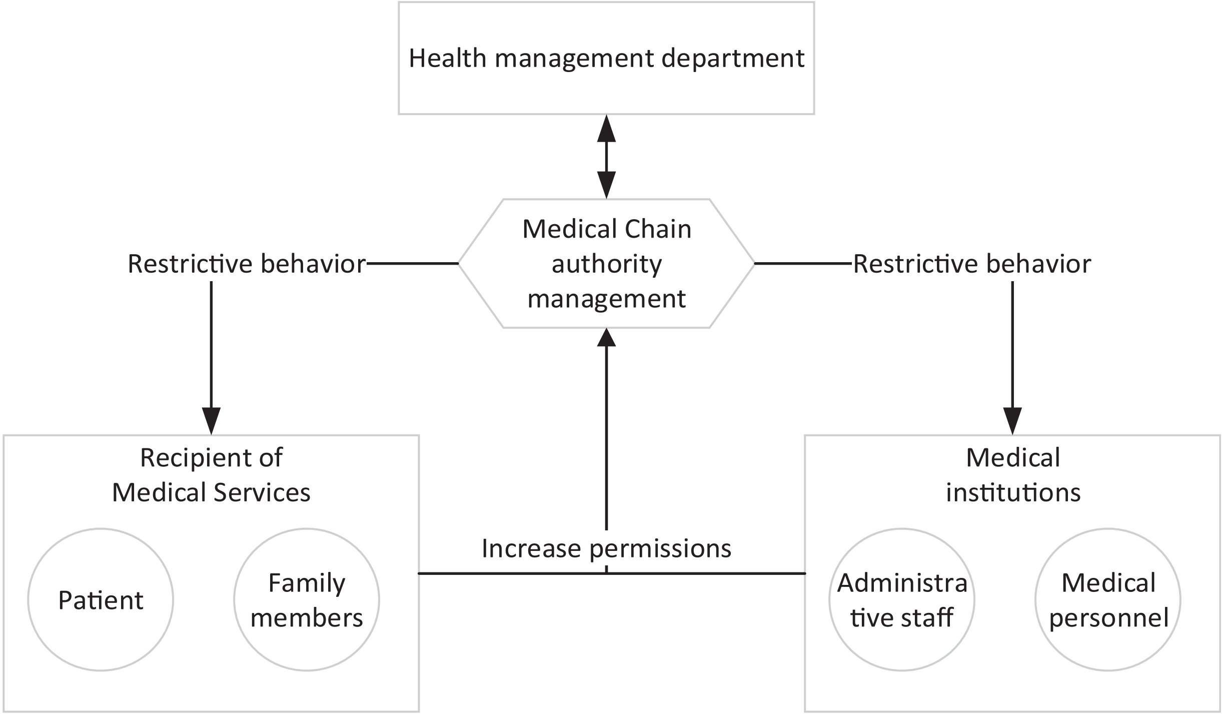


**Figure 7.** Medical chain role and authority.

executed. These codes are released by the health management department in the medical chain and made public on the entire network. In this way, transparent rights management is achieved.

*2.3.4. **Privacy and security.*** The medical chain mainly guarantees privacy and security protection through the system security module. The system security module mainly adopts:

1. Use "key and authentication architecture" to restrict the identity of participants.
2. Utilize the "authority management" module to ensure that all parties participating in the medical chain can generally use their required functions under the prescribed authority. Through the classification of the health management department, adequate supervision and distribution of authority are achieved.
3. Through the classified storage of electronic medical record (EMR), medical data can be stored safely, efficiently, and stably in the medical chain.
4. Through the Validator and Blocker components in medical chain, the PBFT consensus algorithm is used to ensure that the system can handle the malicious behaviors and operational failures of the nodes and that the medical data stored in the blockchain cannot be tampered with or denied.

### *2.4. Ethical approval*

Ethical approval is not necessary because no human subjects and patient information were collected and studied.

## 3. Analysis

### *3.1. Experimental environment*

This article uses the Hyperledger Sawtooth framework to implement the blockchain consensus algorithm PoW, DPOS, POI, and the PBFT consensus algorithm of this article for comparison experiments. Five computers with the same configuration in the laboratory were used as blockchain nodes to conduct investigations. The configuration information of the 5 nodes is shown in Table 1.

### *3.2. Consensus algorithm time-consuming experiment*

A total of 1000 nodes are generated in the experiment. In the experiment, the processing time of a single node after receiving

**Table 1**

**Node configuration information.**

| Node name | Experimental data size | CPU | Memory | Operating System |
|---|---|---|---|---|
| Node1 | 100M/Times | i-7 4700M | 16G | Window 10 |
| Node 2 | 100M/Times | i-7 4700M | 16G | Window 10 |
| Node 3 | 100M/Times | i-7 4700M | 16G | Window 10 |
| Node 4 | 100M/Times | i-7 4700M | 16G | Window 10 |
| Node 5 | 100M/Times | i-7 4700M | 16G | CentOS 7 |

**Table 2**

**Consensus algorithm time consuming.**

| Number of times | Consensus algorithm time consuming (s) | | | |
|---|---|---|---|---|
| | POW | DPOS | POI | PBFT |
| 1 | 50 s | 20 s | 44 s | 13 s |
| 2 | 48 s | 30 s | 35 s | 11 s |
| 3 | 59 s | 22 s | 36 s | 14 s |
| 4 | 55 s | 24 s | 40 s | 12 s |
| 5 | 52 s | 26 s | 39 s | 12 s |
| 6 | 56 s | 27 s | 36 s | 10 s |
| 7 | 49 s | 26 s | 33 s | 11 s |
| 8 | 47 s | 25 s | 37 s | 14 s |
| 9 | 51 s | 26 s | 36 s | 9 s |
| 10 | 53 s | 28 s | 35 s | 10 s |
| Average value | 52 s | 25.4 s | 37.1 s | 11.6 s |

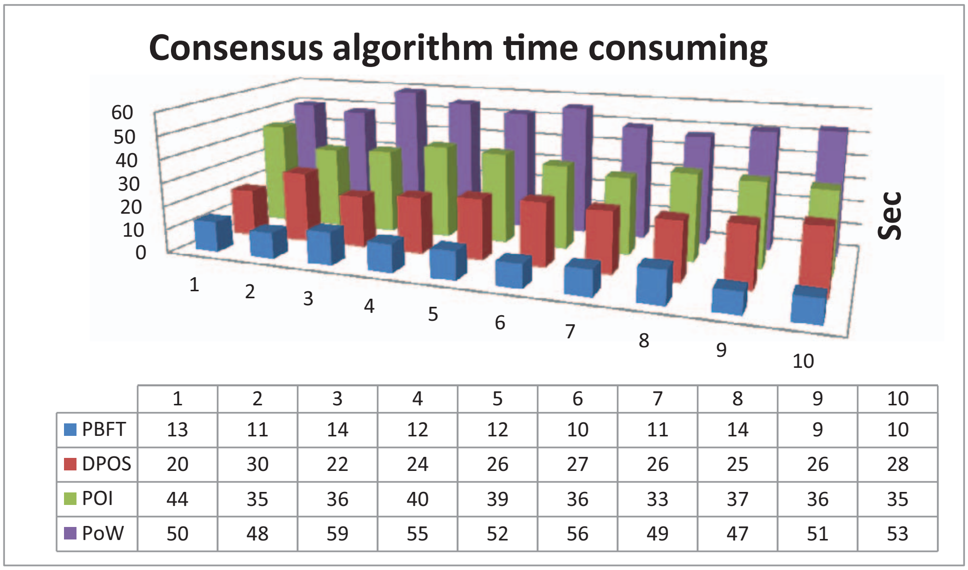
DPOS = Delegate Proof of Stack, PBFT = Possible Byzantine Fault Tolerance, POI = proof of information, POW = proof of work.

the information, the resource utilization of a single machine, CPU processing speed, disk read and write speed, network congestion, and other factors are not considered, and it is assumed that the single machine is at the same time. The reception and transmission of all messages above happen in parallel, and only the time from the start of the consensus process to the end of the consensus process is calculated. To compare and evaluate the time-consuming consensus of the 4 models, 10 consensus algorithm comparison experiments were carried out. In each test, the delay between nodes is randomly generated according to the above-mentioned simulation control conditions. Then the time consumption of a single consensus between the PoW, DPOS, and POI consensus algorithm and the PBFT consensus algorithm in this paper is recorded. The comparison results are shown in Table 1. The initial difficulty value of the PoW and POI consensus algorithms is set to 4, that is, the first 4 bits of the hash value obtained by calculating the hash value of the previous block of Nonce is 0. This can control the average calculation time within 1 minute.

From the experimental results in Table 2, it can be seen that the PoW consensus algorithm takes an average of 52 seconds to generate 1000 nodes, the DPOS consensus algorithm takes an average of 25.4 seconds, the POI consensus algorithm takes an average of 37.1 seconds. The PBFT consensus algorithm in this article receives an average of 11.6 seconds. Figure 8 is a comparison of the time-consuming consensus algorithm of 10 experiments.

## 4. Results

The comparative analysis method is used to compare the existing medical blockchain system with this medical blockchain system. At present, the chief medical blockchain systems are MDSM,[10] MedRec,[12] and ModelChain,[15] as shown in Table 3. The comparison results of various schemes show the medical chain using PBFT as the consensus algorithm:



| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 🟦 PBFT | 13 | 11 | 14 | 12 | 12 | 10 | 11 | 14 | 9 | 10 |
| 🟥 DPOS | 20 | 30 | 22 | 24 | 26 | 27 | 26 | 25 | 26 | 28 |
| 🟩 POI | 44 | 35 | 36 | 40 | 39 | 36 | 33 | 37 | 36 | 35 |
| 🟪 PoW | 50 | 48 | 59 | 55 | 52 | 56 | 49 | 47 | 51 | 53 |

**Figure 8.** Consensus algorithm time consuming.

**Table 3**

**Medical chain vs. existing medical blockchain.**

| System | Blockchain-based | Consensus mechanism | Algorithm type | Pay | Number of nodes required | Computing power requirements | Voting weight setting |
|---|---|---|---|---|---|---|---|
| MDSM | Yes | Improved DPOS | POX | No | 121 | Small | Yes |
| MedRec | Yes | POW | POX | Yes | Many | Big | No |
| ModelChain | Yes | POI | POX | Yes | Many | Big | No |
| Medical chain | Yes | Improved PBFT | BFT | No | Less, at least 4 | Small | No |

1. Compared with MDSM using the improved DPOS algorithm, the number of startup nodes required is far less than MDSM, and MDSM needs to artificially set whether each hospital has the right to vote and the proportion of voters in determining the final result.
2. Compared with MedRec using the POW algorithm, the number of nodes required to maintain the blockchain system is far less than MedRec, and there is no need to pay the blockchain system consensus participating node rewards. No significant amount of computing power is required to maintain the blockchain system.
3. ModelChain uses the form of a private blockchain, and the number of nodes required is uncertain. However, the proof of work consensus mechanism is vulnerable to "51% attacks," that is, nodes can master more than 51% of the computing power of the entire network successfully tampered with and forged the blockchain data,[3] so more nodes are needed to "average" the computing power to prevent this attack. So compared to ModelChain using the POI algorithm, there is no need to pay the consensus participating nodes for compensation, The number of nodes required is also small, and the POI algorithm is based on the POW algorithm, so the power of computing needed is too large.

Therefore, it can be seen that the PBFT consensus algorithm is more suitable for the medical blockchain system. It does not need to pay compensation, requires fewer startup and operating nodes, is scalable in the future, does not need to perform "mining" operations, and requires less computing power. It does not need to set the characteristics of the proportion of voting rights artificially and is fair to hospitals or other medical institutions, so it is consistent with the needs and features of the medical system.

The analysis of the medical blockchain system's computing power demand is based on the consensus mechanism adopted. The determination of computing power demand is based on the following aspects:

1. The block generation method of MDSM with improved DPOS consensus mechanism is that 101 nodes in MIFS take turns to generate blocks. Then the other 100 nodes in MIFS and 20 check nodes in AFS will check the blocks.
2. The blocks of MDSM using the improved DPOS consensus algorithm are generated by MIFS of 101 nodes in turn, so each generated block needs only 2 hashes to calculate the ID of the previous block and to digitally sign the newly generated block in addition to generating the Merkle tree.
3. Backup nodes generate the chunks in the Medical chain using the PBFT consensus algorithm in turn. Similar to MDSM, for each chunk generated, in addition to generating a Merkle tree, only 2 hashes are needed to calculate the ID of the block and digitally sign the newly generated block. However, after the chunk is generated, the MDSM needs to be verified by MIFS with 100 nodes and AFS with 20 nodes. However, the number of verification blocks in the medical chain is n-1, n is the total number of current nodes, and the number of nodes in the medical blockchain is usually not too large, so the medical chain is relatively more flexible than MDSM in computing requirements.

   Most of the blockchains that use the POW consensus mechanism are public chains. The motivation for their maintainers to participate in maintenance is to earn virtual currency to seek higher profits. The principle of the POW consensus mechanism is that "miners" are continually looking for a random number. A "miners" who finds a random number generates a block. This "miners" will receive Bitcoin rewards. Other "miners" continue to mine after this new block, so the POW mechanism will stimulate "miners" to improve their calculations. Find such a random number quickly.

   MedRec is based on Ethernet Fong, which is an open blockchain platform. Anyone can build and run blockchain applications in Ethernet Fong, but they need to pay a certain amount of "Ethernet currency." The maintenance mode of the ethernet platform is similar to that of Bitcoin. By using the consensus way of workload proof, anyone can join or withdraw from ethernet maintenance at any time, resulting in a lot of waste of computing resources. Therefore, the computation required for MedRec with workload proof consensus mechanism is huge.
4. ModelChain adopting the POI consensus mechanism integrates privacy protection, machine learning with a private blockchain network, uses privacy protection machine learning to predict the risk of readmission to patients, and uses proof of work (POW) as a consensus mechanism. Therefore, similar to MedRec, the required computing power is still enormous. And ModelChain also uses machine learning, which requires higher hardware.
5. The medical chain adopting the PBFT consensus mechanism is responsible for generating blocks in turn by the backup nodes. It does not need to perform a large number of useless "mining" operations, so the medical chain requires less computing power.

Compared with other consensus algorithms, the PBFT algorithm does not need to consume a lot of computing resources, and the consensus speed is faster. Still, it is only suitable for a situation where there are not many consensus nodes. When a large number of nodes join the blockchain system, all nodes need to jointly carry out a 3-phase consensus, which leads to a large increase in the number of communications and data transmission, which is likely to cause network congestion or network storms.

Because the PBFT algorithm also has apparent shortcomings in the application of medical systems, such as the efficiency of the system consensus algorithm continues to decrease with the increase of the number of nodes and poor scalability. In the future, by studying consensus algorithms such as POW, POS, and DPOS, analyzing their advantages and disadvantages, combining with the Hyperledger Sawtooth framework, a new PoET (Proof of Elapsed Time) consensus algorithm is proposed and applied in the medical system.

## 5. Discussion

At present, the blockchain technology is receiving more and more attention from researchers, and under their research, it is step-by-step toward perfection and maturity. Medical blockchain, as a type of blockchain technology, is used to realize the security of medical data. Sharing and storage have significant advantages, which is an important development direction in the application research of blockchain technology. The affiliated medical blockchain system proposed in this article uses the PBFT consensus mechanism, which can ensure that the system is implemented with a small computing power safe and stable operation. At the same time, the system is started with fewer nodes, which helps the application and promotion of blockchain technology in medical information. The alliance medical blockchain system medical chain is used for consistency confirmation (consensus) and block generation. There are still problems such as low efficiency, and future research will use a consensus method combining Byzantine fault tolerance algorithms and non-Byzantine fault tolerance algorithms to improve the operating efficiency of the system.

## Author contributions

**Data curation:** Jia Qu.
**Investigation:** Jia Qu.
**Methodology:** Jia Qu.
**Validation:** Jia Qu.
**Writing – original draft:** Jia Qu.

## References

[1] Tsai W-T, Yu L, Wang R, et al. Blockchain application development techniques. J Software 2017;28:1474–87.

[2] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system[R]. Manubot, 2019.

[3] Yuan Y, Wang F-Y. Blockchain: the state of the art and future trends. Acta Automatica Sin 2016;42:481–94.

[4] Lamport L, Shostak R, Pease M. The Byzantine generals problem. ACM Trans Programming Languages Syst 1982;4:382–401.

[5] Han X, Liu Y-M. Research on the consensus mechanisms of blockchain technology. Netinfo Security 2017;147–52.

[6] Yang F, Zhou W, Wu QQ, et al. Delegated proof of stake with downgrade: A secure and efficient blockchain consensus algorithm with downgrade mechanism. IEEE Access 2019;7:118541–55.

[7] Castro M, Liskov B. Practical Byzantine fault tolerance and proactive recovery. ACM Trans Comput Syst 2002;20:398–461.

[8] Lamport L, Shostak R, Pease M. The Byzantine generals problem[M]// Concurrency: the Works of Leslie Lamport 2019;203–26.

[9] Castro M, Liskov B. Practical Byzantine fault tolerance[C]//OSDI. 1999, 99: 173–186.

[10] Xue T-F, Fu Q-C, Wang C, et al. A medical data sharing model via blockchain. Acta Automatica Sin 2017;43:1555–62.

[11] Azaria A, Ekblaw A, Vieira T, et al. MedRec: using blockchain for medical data access and permission management. In: Proceedings of the 2nd International Conference on Open and Big Data (OBD). Vienna, Austria: IEEE, 2016. 25–30.

[12] Wood G. Ethereum: A secure decentralised generalised transaction ledger. Ethereum project yellow paper 2014;151:1–32.

[13] Ivan D. Moving toward a blockchain-based method for the secure storage of patient records[C]//ONC/NIST Use of Blockchain for Healthcare and Research Workshop. Gaithersburg, Maryland, United States: ONC/NIST. 2016: 1–11.

[14] Shrier AA, Chang A, Diakun-Thibault N, et al. Blockchain and health IT: algorithms, privacy, and data[C]//ONC/NIST Use of Blockchain for Healthcare and Research Workshop. Gaithersburg, Maryland, United States: ONC/NIST. 2016.

[15] Kuo TT, Ohno-Machado L. Modelchain: Decentralized privacy-preserving healthcare predictive modeling framework on private blockchain networks. arXiv preprint arXiv:1802.01746, 2018.

[16] Witchey N. Healthcare Transaction Validation Via Blockchain Proof-of-Work, Systems and Methods, WIPO Patent Application WO/2015/175722, November 2015.

[17] Eastlake D, Hansen T. US secure hash algorithms (SHA and HMAC-SHA)[J]. 2006.

[18] Josefsson S. The base16, base32, and base64 data encodings[R]. RFC 4648, October, 2006.

[19] Dong J-C. Analysis of status and causes of hospital information system in China. Chin J Hospital Administration 2003;19:228–30.

[20] Castro M, Liskov B. A Correctness Proof for a Practical Byzantine-Fault-Tolerant Replication Algorithm. Cambridge, MA: Massachusetts Institute of Technology; 1999.

[21] Merkle RC. A digital signature based on a conventional encryption function. In: Proceedings of the 1987 Conference on the Theory and Applications of Cryptographic Techniques. Heidelberg, Berlin, Germany: Springer, 1987. 369–378.

[22] Li F-H, Su M, Shi G-Z, et al. Research status and development trends of the access control model. Acta Electron Sin 2012;40:805–13.

[23] Szabo N. Formalizing and securing relationships on public networks. First Monday 1997;2:1–21.