 JCS, vol.1, no.1, pp.11-17, 2019

Secure Model of Medical Data Sharing for Complex Scenarios

Wei She[1], Yue Hu[2], Zhao Tian[1], Guoning Liu[3], Bo Wang[4] and Wei Liu[1, 2, *]

Abstract: In order to secure the massive heterogeneous medical data for the complex scenarios and improve the information sharing efficiency in healthcare system, a distributed medical data ledger model (DMDL) is proposed in this paper. This DMDL model has adopted the blockchain technology including the function decoupling, the distributed consensus, smart contract as well as multi-channel communication structure of consortium blockchain. The DMDL model not only has high adaptability, but also meets the requirements of the medical treatment processes which generally involve multi-entities, highly private information and secure transaction. The steps for processing the medical data are also introduced. Additionally, the methods for the definition and application of the DMDL model are presented for three specific medical scenarios, i.e., the management of the heterogeneous data, copyright protection for medical data and the secure utilization of sensitive data. The advantage of the proposed DMDL model is demonstrated by comparing with the models which are being currently adopted in healthcare system.

Keywords: Medical data, sharing model, consortium blockchain, communication channel structure, DMDL.

1 Introduction

In recent years, the developed countries have invested heavily in the development of healthcare information system which has the features of the electronic medical records [Azaria, Ekblaw, Vieira et al. (2016); Dong and Quan (2016); Guo, Shi and Zhao (2018)], healthcare data sharing and so on. The purpose of the investment is to maximize the safety of the healthcare data, and improve the overall medical service quality and accessibility, and to reduce medical costs and the medical risk at the same time [Esposito, Santis, Tortora et al. (2018)]. However, in the current healthcare system, the factors related to the medical record check, data save and sync pose a big challenge for the patients and doctors to access to the shared data. What's more, the absence of a sound data transfer model or architecture causes problems for secure and efficient medical data processes and data sharing. Consequently, shortcomings of the current healthcare system

[1] The School of Software Technology, Zhengzhou University, Zhengzhou, 450001, China.

[2] The Cooperative Innovation Center of Internet Healthcare, Zhengzhou University, Zhengzhou, 450000, China.

[3] The School of Mechanical Engineering, Zhengzhou University, Zhengzhou, 450001, China.

[4] The State University of New York at Buffalo, Buffalo, 14200, America.

[*] Corresponding Author: Wei Liu. Email: wliu@zzu.edu.cn.

include slow response, data tampering, insecure data transfer [Guoyuan, Bowen, Pengcheng et al. (2018); Haux (2010); Liang, Zhao, Shetty et al. (2017)].

Blockchain is essentially a kind of decentralized distributed ledger [Lin, Yan, Huang et al. (2018); Mettler (2016)], which is a new computing paradigm based on distributed database, P2P transmission, distributed consensus mechanism, asymmetric encryption algorithm and other technologies [Mengwei, Rong, Tiangang et al. (2018)]. Due to the characteristics of decentralization, collective maintenance, traceability and tamper proof, blockchain technology is suitable for the construction of the programmable decentralized monetary system and has been widely researched on possible applications in the financial sector [Chatterjee and Chatterjee (2017)]. Now the applications of blockchain technology in other areas, such as the internet of things (IoT), the supply chain, the big data and cloud computing, have also attracted great attentions [Fang, Cai, Sun et al. (2018); Xia, Sifah, Asmoah et al. (2017)].

2 The medical data accounting model based on blockchain technology

2.1 Distributed Medical Data Ledger (DMDL)

DMDL has eight groups:

$$DMDL = (EN, CN, ON, MLMS, MDCA, \rho) \tag{1}$$

Where:

1) $EN=\{eni \mid i \in N+\}$ is Endorser nodes set, responsible for checking and endorsing transaction requests in DMDL and excluding illegal or malicious requests.

2) $CN=\{cni \mid i \in N+\}$ is Committer nodes set, responsible for the execution of the transaction and jointly maintain the state and structure of the ledger.

3) $ON=\{oni \mid i \in N+\}$ is Orderer nodes set, responsible for the global ordering of the received transactions in the network, so as to reach the consensus of the whole network.

4) *MLMS* is multi-channel data structure of data ledger, which is composed of multiple blockchains. Each blockchain is composed of different node code, data, contracts, to meet the requirements of different people to access to different data in different business scenarios. See definition 2 for details.

5) *MDCA* is a kind of BFT consensus algorithm running on distributed ledger, which is used to reach consensus on the entire network. See definition 4 for details.

6) $\rho: EN、CN、ON \rightarrow MC$ denotes the mapping of node set to multi-channel model.

2.2 Medical Ledger Smart Contract (MLSC)

Smart contract defines the protocol and interaction within participates in the medical scenarios with the script code [Zhang, Walker, White et al. (2017); Zheng, Xie, Dai et al. (2017)], describes the data sharing process in the digital form, and records the process information on the DMDL accurately. These provide the basis for developing variety of complex medical data sharing applications.

Smart contract is essentially a kind of codes running on the DMDL ledger node, which can be executed automatically according to the pre-set trigger condition. It reduces the

uncertainty caused by human operation. Smart contracts can be called in the specific medical processes like recording the entire medical data sharing processes in the blockchain which include data source, registration, storage, transfer and so on. The detailed smart contract model is shown in Fig. 1.

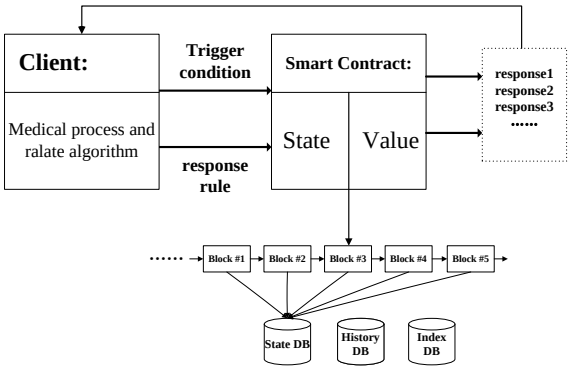

Figure 1: Smart contracts running process

2.3 Medical Data Consensus Algorithm (MDCA)

Medical service scenarios are numerous, and are always involved in large amount of data interaction. In order to guarantee the consistency of all distributed ledgers, we adopt the DBFT algorithm in the DMDL model. This algorithm strictly follows the principles of the byzantine fault tolerance. The algorithm may run in different channel structures depending on the application scenarios. The flow chart of the consensus process is shown in Fig. 2.

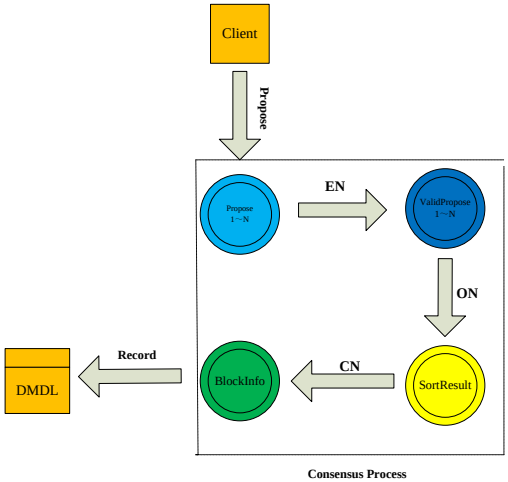

Figure 2: Consensus process

As shown in the Fig. 2, there are three steps in the whole process of consensus, i.e., auditing, execution and sorting. The three steps are independently completed by different nodes. Such a "decoupling" consensus method can greatly overcome the network performance bottleneck.

In the whole process of consensus, sorting is the most important part, which is the key to achieve consensus of the whole ledger. Therefore, the selection of the sorting node (Primary Node) is critical. In order to achieve the consensus and meet the requirements of the distributed and multi-channeled DMDL, the selection algorithm for order node is formulated in this paper. The algorithm is based on the evaluation of historical computation power of *EN, CN* and *ON* in this channel and credibility. All the nodes satisfying the requirements form node set MSP. Nodes in the set are responsible for sorting task and the forming of new block in every consensus processes by turns. The consensus process is as follows:

1) Collect the current and historical operation information in the channel.

2) The collected information is classified into categories according to the specific medical data sharing scenario as shown in Tab. 1.

Table 1: Medical information collection and assessment

	Endorsing	Committing	Ordering
Computing power	Endorsement rate	Committing rate	Ordering rate
	Failurefrequency.	Committing accuracy	
Trustrank evaluation	Malicious signature numbers	Malicious verification number	Ordering validity
	Malicious signature data volume	Malicious verification data volume	

The information of each subject is processed separately, and its information is fitted to the (-1, 1) interval I according to the weight value.

3) In order to consider the importance of a variety of information, the fitting results are multiplied by the influence factor e_i^j to calculate the comprehensive coefficient K_i^j (see Eq. (1)). $\Phi=\{M, N\}$. M and N are, respectively, values to be calculated and the credible value. The set $\Psi=\{EN, CN, ON\}$ represents the endorsement process, the validation process and the sorting process.

$$\begin{cases} k_j^i = e_i^j \times I_i^j, \forall i \in \phi, \forall j \in \psi \\ s.t. -1 \le I_i^j \le 1 \end{cases} \quad (2)$$

4) Integrating the channel nodes according to the comprehensive coefficient K_{ij}, the sorting values of EN, CN and ON. Using Kij, Rj is calculated using

$$R = \sum_{i \in \phi} K_i^j \quad (3)$$

As mentioned above, the consensus algorithm selects the master nodes according to the indicators of credit, scale and computing power, which satisfies the characteristics of distributed and diversified participants in medical data sharing processes.

Methodology here is to decouple the steps of the consensus process into various steps which include endorser, ordering, validation as well as the steps of choosing the main nodes and distributing these tasks to different nodes. The processes will not result in the performance bottleneck during the whole consensus process, and is conducive to the lateral extension of the network.

2.4 Application of the DMDL model

Blockchain has a brand new architecture which makes cross-subject business collaboration easy, efficient and safe. Different from the traditional protocol which is communication process oriented, DMDL is business process oriented. The running process of DMDL is shown in Fig. 3.

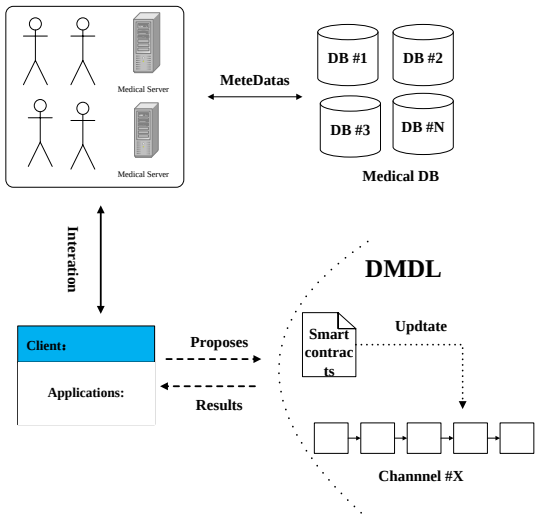

Figure 3: Running process *DMDL*

The practical application of *DMDL* model for medical data sharing is divided into three steps.

1) Confirm the identity information of each node participating in the medical business.

2) Write smart contracts to define business processes.

3) Define multi-channel structure according to the participating nodes and smart contracts.

According to the requirements of modern medical institutions featured by diversified service modes, we will write smart contracts respect to the protection and sharing of privacy data, copyright protection of medical data. The characteristics and advantages of *DMDL* will also be demonstrated.

3 Conclusion

In this paper, we use the Hyperledger Fabric as the experimental platform to develop the smart contract and demonstrate the operation performance. In addition, the advantages of DMDL data model are revealed by comparing with existing data models when identical problems in the process of medical data sharing are analyzed.

3.1 Performance analysis

The configuration of test environment is that, OS: Ubuntu 16.04, CPU:1*3.2 GHz, memory: 4 GB. The fabric network configuration can be seen from Tab. 2.

Table 2: Fabric network configuration

Fabric version	Orderer count	Peer count	OrdererType	BatchTimeout	MessageCount
1.0	1	2	solo	2s	10

There are two main operations in the test smart contract: query and invoke, and each of the two operations contains multiple execution function, such as “initialMedical(), transferMedical(), delete()” in invoke operation and “readMedical(), getHistory(), queryByOwner()” in query operation. We will call multiple times for each function and calculate the average through of the two operations. The result is showed in Tab 3.

Table 3: Throughput tests

Request	Query operation tps	Invoke operation tps
4/s	4/s	4/s
8/s	8/s	8/s
16/s	16/s	16/s
50/s	40/s	39/s

References

Azaria, A.; Ekblaw, A.; Vieira, T.; Lippman, A. (2016): MedRec: using blockchain for medical data access and permission management. *2nd International Conference on Open and Big Data*.

Dong, W. X.; Quan, Y. M. (2016): Pervasive medical information management and services: Key techniques and challenges. *Chinese Journal of Computers*, vol. 35, no. 5, pp. 827-845.

Esposito, C.; Santis, A. D.; Tortora, G.; Chang, H.; Raymond, K. K et al. (2018):

Blockchain: a panacea for healthcare cloud-based data security and privacy. *IEEE Cloud Computing*, vol. 5, no. 1, pp. 31-37.

Guo, R.; Shi, H. X.; Zhao, Q. G. (2018): Secure attribute-based signature scheme with multiple authorities for blockchain in electronic health records systems. *IEEE Access*.

Guoyuan, L.; Bowen, L.; Pengcheng, X.; Min, L.; Wei, B. (2018): Phishing detection with image retrieval based on improved texton correlation descriptor. *Computers, Materials & Continua*, vol. 57, no. 3, pp. 533-547.

Haux, R. (2010): Medical informatics: past, present, future. *International Journal of Medical Informatics*, vol. 79, no. 9, pp. 599.

Liang, X.; Zhao, J.; Shetty, S.; Liu, J.; Li, D. (2017): Integrating blockchain for data sharing and collaboration in mobile healthcare applications. *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*.

Lin, Q.; Yan, H.; Huang, Z.; Chen, W.; Shen, J. et al. (2018): An ID-based linearly homomorphic signature scheme and its application in blockchain. *IEEE Access*.

Mettler, M. (2016): Blockchain technology in healthcare: the revolution starts here. *IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom) IEEE*.

Mengwei, H.; Rong, W.; Tiangang, W; Yu, C.; Buyue, Q. (2018): Reliable medical recommendation based on privacy-preserving collaborative filtering. *Computers, Materials & Continua*, vol. 56, no. 1, pp. 137-149.

Chatterjee, R.; Chatterjee, R. (2017): An overview of the emerging technology: blockchain. *3rd International Conference on Computational Intelligence and Networks*.

Fang, S. Q.; Cai, Z. P.; Sun, W. C.; Liu, A. F.; Liu, F. (2018): Feature selection method based on class discriminative degree for intelligent medical diagnosis. *Computers, Materials & Continu*, vol. 55, no. 3, pp. 419-433.

Xia, Q.; Sifah, E. B.; Asmoah, K. O.; Gao, J.; Du, X. (2017): MeDShare: Trust-Less medical data sharing among cloud service providers via blockchain. *IEEE Access*, vol. 5, no. 99, pp. 14757-14767.

Zhang, P.; Walker, M. A.; White, J.; Schmidt, D. C.; Lenz, G. (2017): Metrics for assessing blockchain-based healthcare decentralized apps. *IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)*.

Zheng, Z.; Xie, S.; Dai, H.; Chen, X.; Wang, H. (2017): An overview of blockchain technology: architecture, consensus, and future trends. *IEEE International Congress on Big Data (BigData Congress)*.