Blockchain-Based Personal Medical Data Protection Solution

Yifan Wang, Zhangang Wang

Abstract— **At this stage, medical data leaks are more and more frequent, and the consequences are getting more and more serious. The existing client-medical data server architecture is based on third-party services that store medical data. Medical data will be stored without the patient's knowledge. Forward to other institutions or individuals. This article proposes a blockchain-based personal medical data protection scheme, which uses the characteristics of blockchain decentralization, tamper-proof, traceability, and transparency to enable medical staff, patients, and medical service providers to authorize and access patients Personal medical data can only be accessed after the patient's consent and authorization. The control of personal medical data is handed over to the patient and recorded on the blockchain. On this basis, the Ethereum platform is used to design a personal medical data management system, and smart contracts are written using the Solidity language to realize the protection of personal medical data. After the smart contract is deployed, the results of the call and the performance of the smart contract are tested 100 times and 1000 times. The experimental results verify its effectiveness and safety.**

Index Terms— **personal medical data; blockchain; smart contracts; third-party services**

I. INTRODUCTION

With the advent of the cloud era, big data has attracted more and more attention, and potential problems involve data leakage that can cause serious consequences. In August 2019, 737 million medical data leaked globally, affecting more than 20 million people in 52 countries. China involved nearly 280,000 patient records. These patient records are very detailed, including: name and surname; date of birth; date of review; investigation Type of range imaging procedure; attending physician; research institute/clinic; number of images generated. In October 2020, the CybelAngel analysis team used a tool to scan a large number of IP addresses and found more than 45 million medical images and related privacy data exposures. These images were stored in 2140 sites in 67 countries (including the United States, the United Kingdom, France and Germany). On a protected (NAS) server. The team pointed out that attackers can also access data to tamper with the patient's medical records. Through the above-mentioned cases of major patient personal medical data leakage, we are particularly worried about the safety of personal medical data.

Medical data requires high security and privacy. This requires consensus between medical data providers and regulatory agencies, and the creation of agreed policies and procedures. As the scale of medical data increases, security mechanisms are needed to protect the data.

Blockchain has remarkable characteristics: tamper-proof modification, traceability, transparency and pseudo-anonymity. In addition to the hot Bitcoin recently, blockchain has also been widely used in data protection. For example, the document [1] proposed the MeDShare system to solve the problem of medical data sharing between medical big data custodians in a lack of trust, and provide data for the cloud storage and sharing of medical data between big data entities. Audit and control, using smart contracts and access control mechanisms to effectively track data behavior, and revoke access to violating entities in the case of violations of data permissions. Literature [2] proposed a blockchain-based framework for interoperable and efficient access to medical records, using smart contracts in the Ethereum-based blockchain to strengthen access control and data obfuscation, and using advanced encryption technology to Further improve safety. In order to ensure the validity of the electronic medical records encapsulated in the blockchain.

Medical data providers increasingly rely on smart technology. Such technologies help manage and diagnose patients' diseases, and help disseminate, receive, and collect patient information for health record management. The usability of the data stored on the HER system is crucial to the success of medical data providers [3]. The security of medical records is becoming more and more important to protect data from security breaches and criminal activities. If unauthorized users have access to patient data, the data may be sold or leaked to the market, and the patient's personal information will be leaked to anyone with access.

II. RELATRD THRORRIES AND RESEARCH

A. *BLOCKCHAIN AND SMART CONTRACTS*

In 2008, Satoshi Nakamoto first proposed the concept of blockchain in "Bitcoin: A Peer-to-Peer Electronic Cash System". Blockchain technology serves as a shared

Manuscript received March 21, 2021

Yifan Wang, male, graduate student with a master's degree in the School of Computer Science and Technology, Tianjin Polytechnic University since 2019, his main research direction: data security, blockchain technology

Zhangang Wang, master supervisor, graduated from the computer department of China textile university (now donghua university) in 1998. He received his master's degree in computer application technology from tianjin technical university in 2005. At present, he is pursuing his doctorate degree on the job. He was promoted to associate professor in 2009. He has published more than 20 papers, and his main research directions include computer application detection, computer network security, and block chain technology

decentralized ledger to record transactions. It can regard the event as the experience of the product or subject, from the origin of the event to the current state, recorded in an unalterable log [4-6]. Blockchain 1.0 is a programmable currency, a cryptographic currency application related to money transfer. remittance, and digital payment. Large financial institutions such as the New York Stock Exchange, Goldman Sachs, Chicago Board of Trade, Citigroup, and Nasdaq have all entered the blockchain field in the past year. Later, it developed into a blockchain with smart contracts, such as Ethereum and Hyperledger, which was called blockchain 2.0. Blockchain 3.0 is the core of the Internet of Value, which is a global distributed accounting system constructed by blockchain. Due to the low efficiency of the current blockchain, the vision and goal of blockchain 3.0 is to achieve higher performance and higher throughput.

Ethereum (English Ethereum) is an open source public blockchain platform with smart contract functions. It provides a decentralized Ethereum Virtual Machine (Ethereum Virtual Machine) through its dedicated cryptocurrency, Ether ("ETH" for short). Handling point-to-point contracts. The program running on the EVM is a smart contract. After the smart contract is deployed, the chain is executed by the EVM. The difference between smart contract and ordinary code is that after the smart contract runs on the chain, it is a transaction, which will be permanently recorded in the block and cannot be tampered with at will.

B. *RELATED RESEARCH*

With the advancement of technology, data can be disseminated under any organization to inspire people to introduce smart technological innovations. However, in the context of the era of big data, there are many unavoidable problems, and some professional researchers have carried out research in the field of medical data.

In order to make better use of medical big data. Saglani, Vatsal J et al. [7] used a convolutional neural network (CNN) model to identify medical concept relationships in clinical cases and predict the results in a big data environment.

In the management of patients' personal medical data, researchers have also proposed many methods. For example, in order to solve the serious privacy risks brought by remote patient monitoring, as well as the security problems of data transmission and data transaction records, the literature [8] proposed A new platform that uses Hyperledger fabric to design and develop blockchain-based smart contracts to detect patient vital signs provides patients with some benefits of global access to medical information. Data analysis and processing technologies such as big data and artificial intelligence mainly focus on using new modeling methods to process data and obtain more ideal results from experiments, while ignoring the protection of patients' personal medical data and the security of data sources.

III. PERSONAL MEDICAL PROTECTION ARCHITECTURE BASED ON BLOCKCHAIN

A. *CURRENT MEDICAL STRUCTURE AND EXISTING PROBLEMS*

The current article gives a comprehensive framework for the personal medical data protection system. As shown in Figure 1, this framework is a general framework for personal medical data, and different schemes use different algorithms. The architecture consists of 4 entities, namely the sick person (SP), the medical staff (MS) who accesses the patient's personal medical data, and the medical service provider (Medical Service Provider, HP) who manages the access rights to the patient's personal medical data. And the Resource Server (MDRS) that manages the storage of medical data.

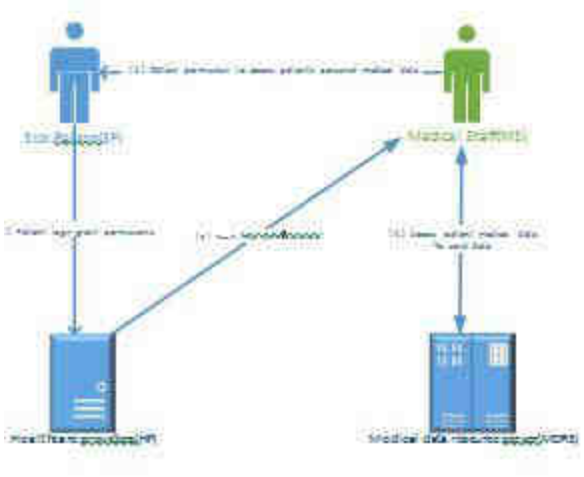

Figure 1 Existing client-server architecture

MS wants to access SP data, it needs to go through the following four steps:

1. When an MS wants to access the medical data of the SP, the MS obtains the permission to access the personal medical data of the SP, as shown in step ① in Figure 1.
2. SP grants a set of permissions to HP to access patient medical data, as shown in step ② in Figure 1.
3. HP verifies and authorizes the identities of SP and MS, and then provides an access token to MS, as shown in step ③ in Figure 1.
4. The MS uses the access token to access the HP. After the HP is verified, it processes the MS request and returns the data, as shown in step ④ in Figure 1.

In steps 1 and 2, the MS sends an access request to the SP. After the SP confirms the authorization, it submits the identity of itself and the MS and sends it to the HP to verify that the MS is authorized by the HP. If the HP verifies that the SP or the MS identity is illegal, the request is not passed, resulting in Authorization failed. The MS can only wait until the SP confirms the authorization and sends the MS and its own identity to the HP for verification. Currently, HP's method of meeting GDPR requirements is based on a client-server architecture, in which a central authority has full access to the system, which leads to limited transparency and trust. Because, most HPs follow the standard of access delegation is OAuth29. Because it is the only authorization and certification body that controls the access and source of data, it completely relies on the authenticity of the HP. The

centralization of the current method has caused serious concerns in [9]. In this case, lack of privacy or security flaws may cause system failures, which may result in network intruders gaining access to patient medical data.
From the perspective of SP, this will result in a lack of transparency and accountability in patient medical data management, and increase the risk of personal medical data leakage. Since all data management mechanisms are operated in HP, medical data can be transferred to other institutions or people without the SP's knowledge.

In response to the above problems, this article is a blockchain-based personal medical data protection program, which takes the patient as the center and hands the control of personal medical data to the patient. All activities are passed through the blockchain platform (BC) and recorded in the district. In order to realize the protection of personal medical data. The use of blockchain and smart contracts provides autonomous operations that are executed safely in a decentralized manner.

Off-chain data storage: Personal medical data is stored off-chain for better scalability and higher efficiency. In addition, storing personal data directly on BC, even in an encrypted manner, may cause potential privacy leaks, which does not comply with GDPR [10]. According to specific scenarios, traditional DBMS (such as Oracle or MongoDB), storage cloud services (such as S3, AWS, or Azure), or distributed storage systems (such as IPFS [11] or Storj[12]) can be used to store data. Only the data pointer to the data is stored on the chain, which can be hash10, connection string, absolute path, or identifier pointing to the data set; it depends on the specific off-chain storage system used by the platform.

B. DESIGN GOALS

The goal of this paper is similar to the solution in [13]. Blockchain is introduced to realize a confidential, fair, reasonable, and controllable personal medical data access program. The program has three goals, namely fairness, soundness, and confidentiality.

Fairness: The fairness defined in this article is similar to literature [13, 14]. The attribute of fairness guarantees that if the MS queries the medical data of the SP, the medical staff will get accurate results, and the query task will be performed by the miners, and the miners will be rewarded by running the protocol correctly. In addition, in order to access the medical data of the SP, the BC will send the access authority and provide the token to the MS at the same time. This means that if a dishonest entity does not execute the agreement in a pre-set way, it will be detected and no data will be obtained. In the previous scenario, this was achieved using verification methods.

Confidentiality. In the solution in this article, since the added document is independent of the previous document, there is no forwarding privacy issue. In addition, this solution is also applied to the index of personal medical data. The actual personal medical data can be stored in any other public storage system and can be protected by any other privacy protection method. Therefore, we only need to protect the confidentiality of the query expression from attackers.

C. Improved system architecture

The improved system architecture in this article is shown in Figure 2.

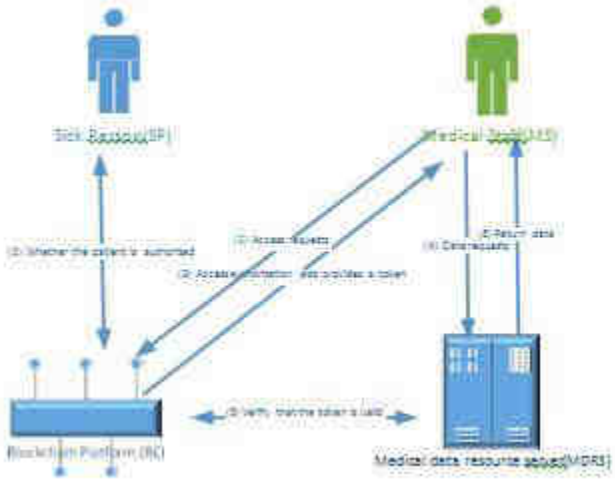

Figure 2 Improved system architecture

When the MS wants to search for the personal medical data of a certain SP, the MS must first request access from the BC and perform identity verification at the same time. After the SP is authorized, it obtains the access authorization and token. Then, the MS will use the token request to search for the MDRS. MDRS will verify the validity of the token through the blockchain platform and record it on the block, obtain the desired medical data and return it to MS.

D. SYSREM FLOW

1. As shown in Figure 2, an MS wants to access the personal medical data of the SP, and first submits an application to the BC for access to the personal medical data of the SP, step ①
2. After the BC verifies the MS identity, it sends the MS's access authorization request to the SP, and then whether the SP is authorized, step ②
3. The BC verifies the identity of the SP and processes the authorization. If the SP agrees to the authorization, the BC will send the MS access token and encrypted address, step ③
4. The MS uses the access token and the encrypted address as parameters and submits them to the MDRS, step ④
5. BC and MDRS synchronize access tokens, step⑤
6. MDRS verifies the token and uses the key to decrypt the address, retrieves the SP's medical data according to the address and returns it to the MS, step ⑥

IV. EXPERIMENTAL RESULTS AND ANALYSIS

The experiment in this article is based on the Ethereum platform, using the JavaScriptVM environment of the remix compiler to deploy the contract. The smart contract console is shown in Figure 3, where reqMange is the function for requesting the smart contract to manage personal medical data, and grant and revoc are functions for granting and

revoking MS access rights, respectively. The response after the SP requests the BC and after authorizing the designated MS access authority is shown in Figure 5 and Figure 6, respectively.

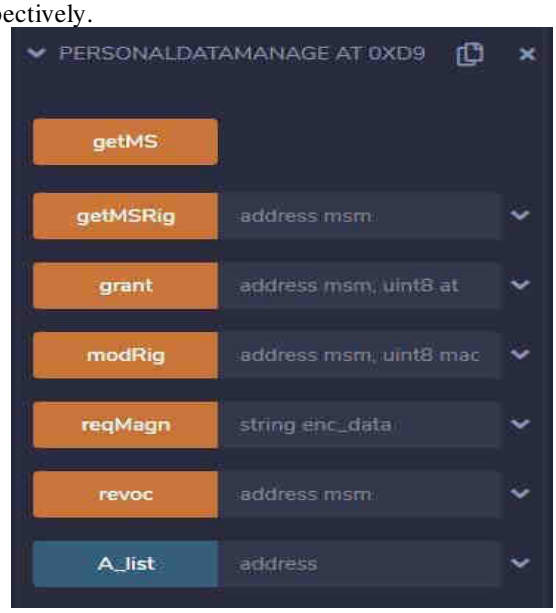

Figure 3 Smart contract console

This article uses the platform built by the remix Ethereum test network, and the average smart contract takes about 2s for each call. It should be that the test network has fewer nodes, so the difficulty of mining is lower, and the block time is shorter. In actual operation, as the number of nodes increases, the mining difficulty coefficient is lower, which will cause too many transactions to produce blocks at the same time, causing network congestion. In summary, the average time for the Ethernet main network to generate blocks has been adjusted to 15.5s. In addition, in actual operation, mining difficulty, gas price and network congestion must be considered.

In this paper, Caliper [15] is used to conduct smart contract performance testing. A total of 2 experiments are carried out. The first experiment consists of 2 rounds, each with 1000 times, and the second experiment consists of 3 rounds, with 100 times each. The experimental results are shown in Table 1 and Table 2 shows.

Table 1 Results of 1000 performance tests of smart contracts

Operating	Number of Successes	Number of Failures	Send Rate /TPS	Maximum Delay /s	Minimum Delay /s	Average Delay /s	Throughput Rate /TPS
ReqMagan	1	0	1.0	0.24	0.24	0.24	5.1
GetMS	1000	0	306.9	46.27	16.52	41.62	11.4
GetMS	1000	0	335.5	43.72	15.10	38.03	12.2

Table 2 Results of 100 performance tests of smart contracts

Operating	Number of Successes	Number of Failures	Send Rate /TPS	Maximum Delay /s	Minimum Delay /s	Average Delay /s	Throughput Rate /TPS
ReqMagan	1	0	1.0	0.24	0.24	0.24	5.1
GetMS	100	0	114.2	4.25	2.53	4.27	11.2
GetMS	100	0	216.3	4.41	2.34	4.39	11.5
GetMS	100	0	335.9	4.12	2.61	4.05	10.6

The experimental results show that the two tests were run successfully. In the case of 1000 runs, the delay performance is better than that of 100 runs, and the throughput is relatively stable, maintaining around 12 TPS. In this system, SP and MS are in a many-to-many relationship, but one MS address cannot correspond to the MS_policy of multiple SPs. In response to this problem, this article uses the Solidity language mapping to create a secondary mapping: mapping (address=>mapping (address=>MS)). The MS_policy of the MS corresponding to the SP address is used to uniquely identify the SP's authorization to the specified MS. For SP's identity authentication, use the solidity language's own feature msg.sender to ensure that authorization and other operations must be issued by the SP. If you use Hyperledger or other programming languages, you need to write the contract code for identity authentication. As the number of nodes in the system increases, the number of distributed ledgers increases accordingly. Only a small number of nodes' data cannot be tampered with to initiate 51% attacks, which can gradually increase the difficulty of tampering with blockchain data.

CONCLUDING REMARKS

Based on the existing client-medical data server architecture, third-party services that store medical data will forward medical data to other institutions or individuals without the patient's knowledge. This article proposes a blockchain-based personal medical data protection the program, centered on the data owner, transfers the control of personal medical data to the patient, and all activities are recorded in the block through the blockchain platform (BC), so as to realize the protection of personal medical data. In the future, the method proposed in this article will be improved in conjunction with the advancement of technology to further increase the block rate and throughput rate.

www.ijerm.com

REFERENCES

1. Xia Q I, Sifah E B, Asamoah K O, et al. MeDShare: Trust-less medical data sharing among cloud service providers via blockchain[J]. IEEE Access, 2017, 5: 14757-14767.
2. Dagher G G, Mohler J, Milojkovic M, et al. Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology[J]. Sustainable cities and society, 2018, 39: 283-297.
3. Hathaliya JJ , Tanwar S , Tyagi S , Kumar N . Securing electronics healthcare records in healthcare 4.0 : a biometric-based approach. Comput Electric Eng 2019;76:398–410 .
4. Griggs, K. N., Ossipova, O.,Kohlios, C. P., Baccarini, A. N.,Howson, E. A., and Hayajneh, T., Healthcare Blockchain System Using Smart Contracts for Secure Automated Remote Patient Monitoring. J. Med. Syst. 42:130, 2018.
5. Lin, C., He, D., Huang, X.,Choo, K.-K. R., and Vasilakos, A. V.,BSeIn: A blockchain-based secure mutual authentication with fine-grained access control system for industry 4.0. J. Netw.Comput. Appl. 116:42–52, 2018.
6. Lin, C., He, D., Huang, X., Khan, M., and Choo, K., A New Transitively Closed Undirected Graph Authentication Scheme for Blockchain-Based Identity Management Systems. IEEE Access 6:28203– 28212, 2018.
7. Saglani VJ, Rawal BS, Vijayakumar V, et al. Big data technology in healthcare: a survey. In: Proceedings of the 2019 10th IFIP international conference on new technologies, mobility and security (NTMS), Canary Islands, 24–26 June 2019, pp. 1–5. New York: IEEE.
8. Jamil F, Ahmad S, Iqbal N, et al. Towards a remote monitoring of patient vital signs based on IoT-based blockchain integrity management platforms in smart hospitals[J]. Sensors, 2020, 20(8): 2195.
9. T. Lodderstedt, M. McGloin, and P . Hunt, “Oauth 2.0 threat model and security considerations,” Tech. Rep., 2013.
10. M. Berberich and M. Steiner, “Blockchain technology and the gdpr-how to reconcile privacy and distributed ledgers,” European Data Protection Law Review, vol. 2, no. 422, 2016.
11. J. Benet, “Ipfs-content addressed, versioned, p2p file system,” arXiv preprint arXiv:1407.3561, 2014.
12. S. Wilkinson, T. Boshevski, J. Brandoff, and V. Buterin, “Storj a peer-to-peer cloud storage network,”2014.
13. S. Hu, C. Cai, Q. Wang, C. Wang, X. Luo, and Kui Ren, Searching an Encrypted Cloud Meets Blockchain: A Decentralized, Reliable and Fair Realization, in: Proceedings of IEEE INFOCOM, 2018.
14. G. Asharov, Towards characterizing complete fairness in secure twoparty computation, in: Proceedings of TCC, 2014, pp.291–316. Springer.
15. Hyperledger Performance and Scale Working Group. Hyperledger blockchain performance metrics [EB/OL] . [2019-12-10] . https : //www. hyperledger.org/wp-content/uploads/2018/10/HL_Whitepaper_Metrics_PDF_V1.01.pdf.

Yifan Wang, male, graduate student with a master's degree in the School of Computer Science and Technology, Tianjin Polytechnic University since 2019, his main research direction: data security, blockchain technology.

Zhangang Wang, master supervisor, graduated from the computer department of China textile university (now donghua university) in 1998. He received his master's degree in computer application technology from tianjin technical university in 2005. At present, he is pursuing his doctorate degree on the job. He was promoted to associate professor in 2009. He has published more than 20 papers, and his main research directions include computer application detection, computer network security, and block chain technology.