Hybrid blockchain–based privacy-preserving electronic medical records sharing scheme across medical information control system

Yu Cao1,2*, Yi Sun1,2* and Jiangsong Min3

Abstract

With the development of big data and medical information control system, electronic medical records sharing across organizations for better medical treatment and advancement has attracted much attention both from academic and industrial areas. However, the source of big data, personal privacy concern, inherent trust issues across organizations and complicated regulation hinder the great progress of healthcare intelligence. Blockchain, as a novel technique, has been used widely to resolve the privacy and security issues in electronic medical records sharing process. In this paper, we propose a hybrid blockchain–based electronic medical records sharing scheme to address the privacy and trust issues across the medical information control systems, rendering the electronic medical records sharing process secure, effective, relatively transparent, immutable, traceable and auditable. Considering the above confidential issues, we use different sharing methods for different parts of medical big data. We share privacy-sensitive couples on the consortium blockchain, while sharing the non-sensitive parts on the public blockchain. In this way, authorized medical information control systems within the consortium can access the data on it for precise medical diagnosis. Institutions such as universities and research institutes can get access to the non-sensitive parts of medical big data for scientific research on symptoms to evolve medical technologies. A working prototype is implemented to demonstrate how the hybrid blockchain facilitates the pharmaceutical operations in a healthcare information control ecosystem. A blockchain benchmark tool Hyperledger Caliper is used to evaluate the performance of hybrid blockchain–based electronic medical records sharing scheme on throughput and average latency which proves to be practicable and excellent.

Keywords

Medical information control system, medical big data sharing, blockchain, privacy and security, healthcare

Date received: 4 April 2020; accepted: 22 April 2020

Introduction

As the new modern information era is getting closer, digitization comes unavoidably and dives deeply into our daily life.1 In the medical information control system, electronic medical records (EMRs) are highly crucial and confidential to each person. They are generated by various medical institutions.2 As the essential digital form of healthcare data, EMRs can collaborate with other promising technologies in an amalgamative way, which brings both opportunities and challenges to medical information control systems, not only on improving the treatment process for patients but also on demanding more careful medical big data management.3 For example, big data and artificial intelligence (AI) have widely applied in the healthcare field for exploring, analyzing and utilizing the value of medical big data to invoke medical domain innovation.4 Meanwhile, blockchain technology has also gained interests from medical information control systems, which aims at solving the drawbacks of existing cross-organizational EMRs sharing process like exposing privacy information to the risk of leakage or abuse.5

1Network and Information Center, Institute of Network Technology, Beijing University of Posts and Telecommunications, Beijing, China
2National Engineering Laboratory for Mobile Network Security (No. [2013] 2685), Beijing, China
3National Engineering Research Center for Supporting Software for Enterprise Internet Services, Shenzhen, China

*Y.C. and Y.S. contributed equally to this work

Corresponding author:
Yu Cao, Network and Information Center, Institute of Network Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China.
Email: cao_yu@bupt.edu.cn

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
Several problems have already emerged in the current centralized pattern of medical big data sharing across different medical information control systems. To be first, single points of failure cannot be avoided entirely, which exposes the system to the unknown risks. Second, medical big data is often hosted in different custodians under their own unique management rules such as government agencies, communication companies or Internet giants. These organizations which should have collaborated are reluctant to share data just for the respective benefit. Third, the storage organizations instead of patients take charge of the medical big data, causing patients worried about the potential abuse of their EMRs. Fourth, multi-faceted surveillance and regulation are quite necessary when sharing these highly private and confidential EMRs among different medical information control systems due to the lack of trust. On one hand, the time-consuming and manpower engaged process hinders the continuous development of healthcare information control ecosystem. On the other hand, it is hard to find a trusted third party with absolute authority and credibility for the supervision and track of medical big data usage. Finally, the medical big data is important to the treatment and related research, but it cannot be guaranteed to be authentic when it is uploaded to information control systems. Therefore, an effective and efficient EMRs sharing scheme across medical information control systems which concentrates on privacy protection is necessary to enable the process easy and secure to unleash the power of big data. Considering the above concerns, we introduce blockchain technology into medical information control systems. It provides a novel solution to these problems with smart contract. Patients can take charge of their EMRs. The professionals like doctors can acquire the completed EMRs, and the others like researchers can get the non-sensitive parts of medical data which mainly focus on the symptoms instead of personal privacy information.

In this paper, we propose a hybrid blockchain–based privacy-preserving and patient-centered electronic medical records sharing scheme (HB-EMRS), which is privacy-preserving and patient-centered. Blockchain can meet the demand in medical big data sharing process by enabling consensus among distrustful parties, and it can guarantee the security and fine-grained data access. The data recorded on blockchain is tamper-resistant, persistent and traceable to achieve stability. In our proposal, sensitive parts of EMRs are recorded on permissioned blockchain. Only the organizations admitted into the consortium can get access to and patients take the actual control over their own EMRs. With the public blockchain, non-sensitive parts of EMRs are uploaded. Any participant in this network can get access to these data. In this way, patient privacy can be protected well from intruders, and data can be shared safely among different systems.

To verify our proposal, a working prototype is implemented to demonstrate important features. Various evaluation criteria are calculated to analyze the performance of HB-EMRS scheme. The key objectives of HB-EMRS are as follows:

1. Automated management of EMR request, approval and usage flow by digitalization and smart contract in the consortium blockchain.
2. Permissioned blockchain is employed to record the sensitive medical data and operations with the tamper-resistant transaction. It can ensure that only limited medical information control systems can join in, generate recordings and acquire data.
3. Permissionless blockchain is employed to record the non-sensitive parts and operations. It can ensure that these parts are available to the public to promote the development of medical information control ecosystems.

The rest of this paper is organized as follows. Section “Related works” introduces the big data and blockchain solutions briefly. In section “Public, consortium and private blockchain,” we discuss the classification of blockchain and the reason why we use hybrid blockchain, combining consortium blockchain with public blockchain. Section “HB-EMRS scheme design and implement” describes the details about our proposal. Section “Conclusion” concludes the study from the implemented prototype and executed experiments. Conflicting interests of authors, funding information and the ORCID of authors is given after. References are listed out in section “References.”

Related works

It is well recognized that cross-organizational data dissemination is able to motivate innovation, introduce intelligence and raise awareness in the certain domain. However, there are many unavoidable problems in big data circulation. Some professional teams have started research on this issue in healthcare field.

To make better use of medical big data, some attempts have been carried out. Saglani et al. made use of convolutional neural network (CNN) model to recognize medical concept relations in clinical records and predict consequences under big data environment. Chaudhary et al. and Yu et al. both proposed neural network–based cancer prediction models in their papers to improve diagnostic accuracy. Wulff et al. combined openEHR archetypes which are a semantically enriched clinical information model, terminology bindings and the Archetype Query Language (AQL) to design an interoperable concept, enabling integration of clinical decision support systems (CDSSs) easy across different institutions. Data analysis and process technologies like big data and AI mainly focus on using new modeling methods to process data and derive more desirable final results from experiments while ignoring the security of privacy protection and data source. EMR is entirely
private to patients. There would exactly be privacy issues such as information leakage or abuse if exposed to the public. According to the communiqué on 22 May 2019 from the National Health Commission of the People’s Republic of China, Chinese total medical treatment was over 8.31 billion person-times. Residents went to medical and health institutions of 6.0 visits on average in 2018.14 However, patients leave these data scattered across different information control systems which have their own principles, causing the phenomena of information island.15,16 It is worth noting that EMR is required to be shared among or accessed by different parties for different purposes frequently.17

To better achieve privacy protection in medical big data sharing across medical information control systems, some research studies have been implemented. Bandara et al.18 proposed a big data friendly blockchain system Mystiko, which supported high transaction throughput, scalability, availability and full-text search features. Chen and Xu19 proposed a blockchain-based ecosystem for big data exchange. Parties can cooperate in exchanging data with a decentralized way. Azaria et al.20 designed and implemented MedRec system, which was a novel decentralized records management system that handles EMR authentication, confidentiality, accountability and sharing issues based on Ethereum smart contracts. Yue et al.21 proposed an app called Healthcare Data Gateways (HDG) based on blockchain to enable data sharing process under risk control. Dubovitskaya et al.22 proposed and implemented a framework about managing and sharing EMRs for cancer care, which ensured privacy, security, availability and fine-grained access control. Al Omar et al.3 presented a patient-centered healthcare data management system named MediBchain, which was based on blockchain to attain privacy protection. Liu et al.23 proposed a blockchain-based privacy-preserving EMRs sharing scheme called Blockchain-based Privacy-preserving Data Sharing (BPDS). The original data were stored in the cloud environment, and the indexes of data were reserved in the consortium blockchain to reduce the risk of leakage. These applications unveil a prevalent phenomenon that current exploration works mainly focus on resolving the technical and methodological potential of blockchain in medical data management. All the works mentioned above organize the medical big data as a whole instead distinguishing the sensitive couples from the non-sensitive ones. Moreover, only one kind of blockchain was employed to provide services to ensure security, privacy and efficiency.

To better analyze the performance of deployed blockchain solutions, different evaluation criteria are derived from experiments and benchmark tools are used to set configuration environment. Sukhwani et al.24 run workload using the Hyperledger Caliper which is a blockchain benchmark tool to analyze the performance of different types of nodes in Hyperledger Fabric v1.0 + using Stochastic Reward Nets (SRNs), which is a formalism allowing convenient configuration of blockchain network systems. Baliga et al.25 characterized the performance and scalability features of Hyperledger Fabric v1.0. They conducted experiments on throughput, latency, chaincodes and transactions to analyze its performance. Kuzlu et al.26 calculated performance results of throughput, latency and scalability of a Hyperledger Fabric blockchain framework. Roehrs et al.27 proposed a blockchain-based network for storing personal health record (PHR) and implemented the corresponding prototype. It evaluated excellent on response time and availability. Taking these metrics into consideration, we hire Hyperledger Caliper to configure the test environment and do the experiments with workload of various parameters.

In our scheme, we exploit two kinds of blockchain to make the cross-organizational medical big data sharing process more secure and effective with sensitive parts of medical big data stored on the consortium blockchain and the non-sensitive parts on the public blockchain. Organizations which are authorized can join in the consortium after approval to launch requests for completed data and the others like universities and research institutions can access the non-sensitive parts from the public blockchain. The logics are implemented by chaincode, which is the smart contract in Hyperledger Fabric.

Public, consortium and private blockchain

Blockchain, which is famous as a distributed transaction ledger with an ordered list of successive records, has attracted tremendous attention from researches and industries since its advent in Nakamoto’s28 paper Bitcoin: A Peer-to-Peer Electronic Cash System in 2008. The fundamental structure of blockchain can be described as a chain of linked blocks, which are defined as individual components that contain details about transaction. The next block contains the hash of the previous block.29 Benefiting from such structure and the cryptographic algorithms used, blockchain provides developers with decentralized, immutable, open, traceable and fault-tolerant features to ensure accountable and auditable. Thus, it holds great potential to improve the current business models like supply chain, manufacturing, healthcare and so on. Blockchain can be divided into three categories: public, consortium and private blockchain.30,31 The overall comparison of these three is shown in Table 1.

Public blockchain means that any participant in the world can read the ledger, send and validate transactions, participate in the consensus process. There is neither a central authority nor a trusted third party in the network. Peers join in and leave at their own will. Nodes in the public blockchain network are interconnected in a flat topology. Bitcoin and Ethereum systems are the top two public blockchains considering the number of users and level of security. Bitcoin system is the initial implementation of public blockchain
based on the unspent transaction output (UTXO) transaction model, which records the internal transaction circulation and transfer of digital assets. Each output of transaction is used as the input of the next relevant transaction. The fundamental operations of UTXO model are depicted as Figure 1. The payment and balance both circulate as transaction. To be different, Ethereum bases on account model to complete the transaction. Peers in the Ethereum network launch transaction through smart contracts just as described in Figure 2.

Consortium blockchain means that the consensus process can only be completed by a limited and predefined set of participants. Write permissions are only in the hands of someone or some organization, and data access has strict rules which can be public or restricted. Therefore,

| Participation rules | Public blockchain | Consortium blockchain | Private blockchain |
|---------------------|-------------------|-----------------------|--------------------|
| Consensus algorithm | PoW/PoS/DPoS/Pos  | Authenticated parties within consortium | Individual/organization |
| Recording rules     | All the participants | Distributed consensus algorithm (e.g. PBFT) | Distributed consensus algorithm |
| Motivation mechanism| Need              | Optional              | Do not need |
| Features            | Self-constructed credit management | Efficiency and cost optimization | Transparent and traceable |
| Transaction ability | (3–30) tps        | (1k–1w) tps           | (1k–10w) tps |
| Typical application(s) | Digital concurrency | Payment/data sharing | Audit issuance |

PBFT: practical Byzantine fault tolerance; DPoS: delegated proof of stake; PoW: proof of work; PoS: proof of stake.

Figure 1. UTXO transaction model of Bitcoin system.
private blockchain is efficient, privacy-concerned, low cost and product-protected compared to the public blockchain and the consortium blockchain.

In our scheme, we utilize the public and consortium blockchain to complete the sharing logics design and programs implementation. Because of the confidentiality of EMRs, we build a secure sensitive parts sharing environment utilizing Hyperledger Fabric. Meanwhile, the public blockchain Ethereum is used to share those non-sensitive parts to separate medical treatment from research activities.

**HB-EMRS scheme design and implement**

**HB-EMRS scheme design**

Currently, EMR is scattered across different medical information control systems. For instance, patient examination information may be stored in some professional examination centers, and their diagnosis and treatment recordings may be managed by different hospitals just as illustrated in Figure 4 blow. However, interoperability across information control systems has not well established yet. Patients must carry their records around for seeking a better therapy solution. Meanwhile, other organizations may request these data for usage. Research institutions and universities may need these data to improve the curing approaches and develop novel pharmacy methods. The government may need these data to formulate corresponding policies. But the medical examination and treatment process are relevant to privacy information which can target at a certain patient in real life. The analysis and research development only need the final diagnosis results and statistics which are not related to the personal information. Moreover, data sharing process is driven by complicated long-time procedures which require much manpower engaged in to ensure the authorization and review operations in the trustless cross-organizational environment.

To improve the efficiency and safety in EMRs sharing across medical information control systems, a uniform and interoperable approach is expected to advance the data management. Therefore, we introduce an EMR sharing scheme based on hybrid blockchain architecture. As depicted in Figure 5, all the participants involved in the consortium are tied together with a set of rules and predefined smart contracts. Only the

---

**Table 2. Different nodes in Hyperledger Fabric network.**

| Nodes    | Functionality                                                                 |
|----------|-------------------------------------------------------------------------------|
| Endorser | Responsible for checking and endorsing the transaction proposal and calculating the execution result of the transactions; |
| Commiter | Responsible for checking the legality of transactions again before accepting the results. Then accept the modification of the legal transaction and write the final results into the blockchain; |
| Orderer  | Responsible for sorting all transactions which are sent to the network. Then package the sorted transactions into blocks according to the conventions in the configuration. Finally submit them to the committers for processing; |
| CA       | Responsible for managing all certificates in the network and providing standard public key infrastructure (PKI) services. |

CA: Certificate Authority.
authorized organizations can access the sensitive parts of EMRs to do the medical diagnosis and make treatment plans. Considering the large amount of EMRs, HB-EMRS is designed with on-chain and off-chain storage. The data will be encrypted and stored in the InterPlanetary File System (IPFS), which is a distributed storage system supporting the content-based index.\(^\text{33}\)

HB-EMRS scheme can be expressed as equation (1).

In equation (1), \(ch\) and \(EMR\) locate at the center of medical big data sharing process. IPFS plays the role of secure and efficient off-chain data storage.

\[
\text{HB} - \text{EMRS} = \{ch, EMR, IPFS\} \cup \{ch_{\text{sor}} \text{hash} \text{EMR}, ch_{\text{pub}} \text{hash} \text{EMR}_{\text{non-sor}}\} \cup \{IPFS(\text{EMR})\}
\]

The core of equation (1) can be described as equation (2) below. The symbol \(ch\) means the blockchain and the subscripts \(\text{sor}\) and \(\text{pub}\) depict the blockchain type that is used. EMR to be shared is divided into two different parts named as \(EMR_{\text{sor}}\) and \(EMR_{\text{non-sor}}\), which will be shared within different scopes and organizations.

\[
\text{HB} - \text{EMRS} = \{ch_{\text{sor}} \text{hash} \text{EMR} \cup ch_{\text{pub}} \text{hash} \text{EMR}_{\text{non-sor}}\} \cup \{IPFS(\text{EMR})\}
\]

\[
\text{EMR} = \sum_i \text{EMR}_{\text{sor}} \cup \sum_i \text{EMR}_{\text{non-sor}}
\]

\[
\text{EMR} = \sum_i \text{EMR}_{\text{sor}} \cup \sum_i \text{EMR}_{\text{non-sor}}
\]
The storage structure of HB-EMRS is described in equation (3) below. The symbol \( tx \) stands for the transaction which is packaged into blocks after the authentication by minors of blockchain network. \( Get(x, tx) \) operation means getting \( tx \) information from \( x \) blockchain by calling the API defined by Hyperledger Fabric and Ethereum platforms correspondingly.

\[
\text{Storage} = \delta \left\{ \begin{array}{l}
\{ \text{IPFS}(\text{EMR}) \cup \text{ch}_{\text{sen}}[\text{hash}(\text{EMR})], \text{IPFS}(\text{EMR}) \} \uparrow \theta_{\text{hash}} \\
\{ \text{Get}(\text{ch}, tx) \leftrightarrow \text{IPFS}(\text{EMR}) \}
\end{array} \right.
\]

\[
(3)
\]

Data storage requires integrity, security and efficiency. Currently, the centralized storage methods and structures which are commonly used in medical information control systems have security risks, although these problems can be resolved to some extent by designing and applying the viable backup and emergency treatment solutions. This may require coordination and concessions of various factors, such as the resource utilization and efficiency allocation. Combining the data storage and guarantee capabilities of IPFS which is indexed by content, HB-EMRS stores the complete medical big data on it. Then, the hashes of the completed and the non-sensitive correspondingly are stored on the consortium blockchain and the public blockchain as depicted in equation (3). So that individuals may have absolute control over the management of private EMRs, eliminating the middle trust barriers. In addition, this structure can also realize the function of data redundancy backup. Once the EMR data on the consortium blockchain are found to be tampered with maliciously, the complete data stored on the IPFS can be used for safe recovery and traceback, ensuring the security of HB-EMRS solution.

As illustrated in Figure 6, the predefined smart contracts manage the EMRs sharing logics and operations between medical information control systems and users. Patients are responsible for uploading the corresponding hash of original EMRs. Doctors or someone else who join in the consortium and get permission from the data owners can access the sensitive data. Researchers or else who do not need personal information can get hash from public blockchain. Then, they will retrieve data from off-chain storage.

In the following equation (4), HB-EMRS implements the privacy-preserving medical big data sharing and authentication by predefined smart contracts.

\[
\text{Smart contract} = \rho(bc, sm, auth)
\]

\[
= \rho \left\{ \text{Get}(\text{ch}_{\text{sen}}, \text{tx}) \leftrightarrow \mu(\text{EMR}), \\text{Get}(\text{ch}_{\text{pub}}, \text{tx}') \leftrightarrow \mu(\text{EMR}_{\text{non-sen}}) \equiv \text{IPFS}(\text{EMR}) \Rightarrow \text{EMR} = \text{EMR}_{\text{sen}} \cup \text{EMR}_{\text{non-sen}} \right\}
\]

\[
(4)
\]
The main process of the smart contracts is shown in Algorithm 1 below. The P2P network of blockchain ensures that the untrusted participants can work well under the effort of smart contracts without any central party.

Figure 7 illustrates the archetype of HB-EMRS scheme. The blockchain access control layer defines various fundamental operations on EMRs. The service implementation layer is the encapsulation of logical processing which will call the function and interface defined in the underlying layer. The top service call layer takes charge of receiving the request from users and triggering the corresponding services.

In the consortium, each authorized participant has its own identification, such as hospital, government, insurance agency and so on. All the interactions among participants are recorded as transactions and are packaged into blocks on blockchain to be auditable and safe. EMRs generated by different medical control systems are stored in IPFS and the hash of it will be stored on blockchain to reduce the on-chain storing burden. The participants in this consortium who have the right to access can get the original data from IPFS, which will promise the integrity and safety of these data. Only the patients have the absolute control over their own EMR data.

**HB-EMRS scheme implementation and results analysis**

The implementation scheme focuses on realizing the workflow of operations which are depicted in Figure 8. Patients upload their data at first. The actual data are stored off-chain in IPFS and the corresponding separated hash is stored on hybrid blockchain. Then, doctors or insurance staff can acquire the data for their

![Figure 6. Agreement on EMRs sharing operations across medical information control systems which is related to the sensitive and the non-sensitive data.](image)

![Figure 7. The functional layers and components of HB-EMRS scheme.](image)
aims. Patients themselves can grant and also revoke permission. After receiving permission from patients,
doctors or else need to query blockchain to obtain the hash value and then the real data from IPFS. When the
granted permission is revoked, the following updated content will not be accessible unless they require and
get permission from owners again.

An open source blockchain performance testing tool Hyperledger Caliper is used to analyze the performance
of our HB-EMRS scheme. This blockchain benchmark tool can represent many clients and deploy various
configurations easily. Figure 9 depicts the architecture of Hyperledger Caliper for performance
evaluation. It defines the parameters by configuration files, including transaction rates, transaction numbers
and transaction types. Transaction rates are counted by tps (transaction per second) and conducted in fixed
rate, in which transactions are sent out by fixed intervals. Transaction numbers define the numbers of trans-
action generated in a round. Transaction type is specified as open and query operations.

We build the system with Hyperledger Fabric and Ethereum on Ubuntu 16.04LTS and test the fundamen-
tal functions under different configurations. The components of Hyperledger Fabric are launched as docker
containers. We deploy the environment of caliper @v0.3.0, docker @18.06.1-ce, docker-compose
@1.22.0, grpc @1.10.1, fabric-ca-client @1.1.0 and fabric-client @1.1.0. We create almost 500 users includ-
ing doctors, patients, insurance agencies and so on to test application scenarios.

The comparable performance results of two and four organizations (each with four peers) are depicted
in Figure 10. It is obvious that the response latency increases along with the request load. When reaching
four organizations, the response latency gets slower since more peers require validating and endorsing the
transactions. But it has to be noted that many factors like deployment environment, number of peers, used
databases and so on can affect the latency. These

---

**Algorithm 1.** Pseudocode process of HB-EMRS smart contracts.

**Initialization:**
- Patients EMR for sharing across medical information control systems: $\text{Pati} = \{p_1, \ldots, p_n\}$;
- Participants for acquiring sensitive parts of medical data: $\text{MedOrg} = \{\text{med}_1, \ldots, \text{med}_n\}$;
- Research institutions for acquiring non-sensitive parts of medical data: $\text{ReOrg} = \{\text{re}_1, \ldots, \text{re}_n\}$;
- The fundamental information for peers of HB-EMRS system: $\sum \text{Pati}, \sum \text{Org}$;

**Patients:**
1: for treatment process with $p_i \in \text{Pati}$ and $\text{med}_j \in \text{MedOrg}$ do:
2: $\text{shareList} + = \text{med}_j$ with deduplication;
3: $\text{end for}$
4: if revoke $\text{med}_k \in \text{MedOrg}$ do:
5: $\text{shareList} / C_0 = \text{med}_k$;
6: $\text{end if}$

**Medical Control Systems in consortium:**
1: for $p_i \in \text{Pati}$ and medical control system $\in \text{shareList}$ do:
2: $\text{hash} (\text{EMR}_i) : = \text{Get chsor, } \text{tx, patientID}$;
3: $\text{details} \rightarrow \text{IPFS} (\text{hash}_i)$;
4: $\text{EMR}_i : = \text{EMR}_0$;
5: $\text{end for}$

**Research Institutions interested in medical big data publicly:**
1: for $\text{rei} \in \text{ReOrg}$ do:
2: $\text{hash} (\text{EMR}_j) : = \text{Get chpub, tx} / C_0 / C_1$;
3: $\text{non-sensitivedetails} \rightarrow \text{IPFS} (\text{hash}_j)$;
4: $\text{end for}$
5: Start the following experiments.

P.S. All of the information will be recorded on blockchain

---

Figure 8. The workflow of HB-EMRS scheme operations.
factors are breakthrough to overcome the system performance barriers.

Several cases are defined by varying the configuration files of key parameters and conducted on Hyperledger Caliper. They are shown in Table 3.

In case I, to test the impact of transaction rates, parameters are set as 100, 150, 200, 250 and 300 tps with 1000 transactions per round. In case II, to evaluate the impact of transaction numbers, parameters are set as 1000, 10000, 100000 transactions with 200 tps per round. Each case is performed both with open and query operations.

Before our attempt HB-EMRS of evaluating the performance of blockchain in medical big data sharing, Kuzlu et al.26 had launched their experiments and got the final results. In this paper, we compare the running results with their paper. The throughput is defined as equation (5). Figure 11 shows the results of case I with open operation. Figure 12 shows the results of case I with query operation. In Figure 11, the throughput

![Figure 9. The model of Hyperledger Caliper.](image)

![Figure 10. HB-EMRS response latency comparison deployed with two and four organizations.](image)

| Table 3. Varied parameters for HB-EMRS performance evaluation. |
|-----------------|-----------------|-----------------|
| Tested cases    | Transaction rates | Transaction numbers |
| I. Impact of transaction rates | 100, 150, 200, 250, 300 tps | 1000            |
| II. Impact of transaction numbers | 200 tps         | 1000, 10000, 100000 |
decreases and the latency increases when the transaction rate reaches 250 tps, and this trend continues. Nodes in the blockchain network need to process more transactions because the open behavior requires one read and one write operations in one transaction.

\[
\text{Throughput} = \frac{\text{transaction numbers}}{\text{time period(s)}}
\] (5)

In Figure 12, the throughput almost equals to the transaction rates per round. The latency almost equals to zero. It is because that query operation simply read the state from the underlying database, that is, CouchDB.

It can be seen from Figures 11 and 12 that the throughput is highest and the latency is lowest when the transaction rate reaches 200 tps. Figures 13 and 14
show the result of case II with open and query operations correspondingly. They show that HB-EMRS works well under transaction rate of 200tps with open and query operations.

Concluded from the above performance results of Figures 11–14, HB-EMRS can achieve average performance in the test. It reaches its high quality of transaction process ability when the transaction rate sets as 200 tps. The transaction numbers varied from 1000 to 100000 make little impact on throughput and latency.

The advantages of our HB-EMRS scheme over the existing EMR management system are apparent from the view of overall system structure, distribution and sharing strategies, security and privacy and regulation overhead. First of all, EMR is scattered across organizations and managed in a centralized way in current EMR management system, causing the phenomenon of data island. However, HB-EMRS scheme is worked in decentralization and distinguishes EMRs from sensitive and non-sensitive to protect privacy. Second, there are
many agreement documents that need to be censored, and a sequence of procedures need to be launched in the centralized sharing pattern. In contrast, predefined smart contracts manage the sharing logics in HB-EMRS scheme and various policies are available. Next, the current EMR management system requires manpower for legal supervision. It is a long-time auditing process and can expose to the potential risks of privacy leakage. But in HB-EMRS scheme, all the operations are packaged into transaction and validated under the proof and consensus automatically; they cannot be tampered. Moreover, the approach of splitting the EMRs into parts can protect privacy better. Finally, in HB-EMRS scheme, the consensus is executed automatically, which is fast and accurate once smart contracts are deployed successfully.

Conclusion

In this paper, we mainly introduce and implement the HB-EMRS scheme, which utilizes the permissioned blockchain and permissionless blockchain to facilitate the safe, effective and efficient privacy-concern EMRs sharing. As a novel solution to the problems under the medical big data environment with smart contract, blockchain enables the process easy and secure to unleash the power of big data across medical information control systems. According to the performance evaluation results above, HB-EMRS can realize the functions predesigned well and achieve an average performance. As a reasonable expectation and exploration, this system has the chance to be further enhanced and improved according to the real-world requirements, which can better integrate with the existing on-duty systems to serve more patients and organizations that are associated with medical businesses.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: This work is supported by the National Natural Science Foundation of China (Grant no. 61601041) and the Fundamental Research Funds for the Central Universities (2019PTB-003).

ORCID iD

Yu Cao https://orcid.org/0000-0001-6196-7432

References

1. Raghupathi W and Raghupathi V. Big data analytics in healthcare: promise and potential. *Health Inf Sci Syst* 2014; 2(3): 1–10.

2. Krumholz HM and Waldstreicher J. The Yale Open Data Access (YODA) project—a mechanism for data sharing. *New Engl J Med* 2016; 375(5): 403–405.

3. Al Omar A, Rahman MS, Basu A, et al. MediBchain: a blockchain based privacy preserving platform for healthcare data. In: *Proceedings of the security, privacy, and anonymity in computation, communication, and storage, SpaCCS* 2017 (eds G Wang, M Atiquzzaman, Z Yan, et al.), Guangzhou, China, 12–15 December 2017, pp. 534–543. Cham: Springer.

4. He B, Guan Y and Dai R. Classifying medical relations in clinical text via convolutional neural networks. *Artif Intell Med* 2019; 93: 43–49.

5. Kish LJ and Topol EJ. Unpatients-why patients should own their medical data. *Nat Biotechnol* 2015; 33(9): 921–924.

6. Yue L, Huang J, Qin S, et al. Big data model of security sharing based on blockchain. In: *Proceedings of the 3rd international conference on big data computing and communications (BIGCOM)*, Chengdu, China, 10–11 August 2017, pp. 117–121. New York: IEEE.

7. Xiao Z, Li Z, Liu Y, et al. EMRSshare: a cross-organizational medical data sharing and management framework using permissioned blockchain. In: *Proceedings of the 2018 IEEE 24th international conference on parallel and distributed systems (ICPADS)*, Singapore, 11–13 December 2018, pp. 998–1003. New York: IEEE.

8. The Office of the National Coordinator for Health Information Technology (ONC). Report to congress: report on health information blocking. ONC report, The Office of the National Coordinator for Health Information Technology (ONC), Washington, DC, April 2015.

9. Li C, Cao Y, Hu Z, et al. Blockchain-based bidirectional updates on fine-grained medical data. In: *Proceedings of the 2019 IEEE 35th international conference on data engineering workshops (ICDEW)*, Macao, China, 8–12 April 2019, pp. 22–27. New York: IEEE.

10. 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.

11. Chaudhary K, Poirion OB, Lu L, et al. Deep learning-based multi-omics integration robustly predicts survival in liver cancer. *Clin Cancer Res* 2018; 24(6): 1248–1259.

12. Yu X, Yu G and Wang J. Clustering cancer gene expression data by projective clustering ensemble. *PLoS ONE* 2017; 12(2): e0171429.

13. Wulff A, Haarbrandt B, Tute E, et al. An interoperable clinical decision-support system for early detection of SIRS in pediatric intensive care using openEHR. *Artif Intell Med* 2018; 89: 10–23.

14. Planning, Development and Information Department. Statistical bulletin of China’s health care development in 2018, www.nhc.gov.cn/guihuaxxs/s10748/201905/9b8d52727cf346049de8acece25ffceb0.shtml (2019, accessed 10 March 2020).

15. Chowdhury MJM, Colman A, Kabir MA, et al. Blockchain as a notarization service for data sharing with personal data store. In: *Proceedings of the 2018 17th IEEE international conference on trust, security and privacy in computing and communications* 12th IEEE international conference on big data science and engineering (TrustCom)
16. Zhang J, Xue N and Huang X. A secure system for pervasive social network-based healthcare. *IEEE Access* 2016; 4: 9239–9250.

17. Xia Q, Sifah EB, Asamoah KO, et al. MedShare: trustless medical data sharing among cloud service providers via blockchain. *IEEE Access* 2017; 5: 14757–14767.

18. Bandara E, Ng WK, De Zoysa K, et al. Mystiko—blockchain meets big data. In: *Proceedings of the 2018 IEEE international conference on big data*, Seattle, WA, 10–13 December 2018, pp. 3024–3032. New York: IEEE.

19. Chen J and Xue Y. Bootstrapping a blockchain based ecosystem for big data exchange. In: *Proceedings of the 2017 IEEE international congress on big data (BigData Congress)*, Honolulu, HI, 25–30 June 2017, pp. 460–463. New York: IEEE.

20. Azaria A, Ekblaw A, Vieira T, et al. MedRec: using blockchain for medical data access and permission management. In: *Proceedings of the 2016 2nd international conference on open and big data*, Vienna, 22–24 August 2016, pp. 25–30. New York: IEEE.

21. Yue X, Wang H, Jin D, et al. Healthcare data gateways: found healthcare intelligence on blockchain with novel privacy risk control. *J Med Syst* 2016; 40(10): 1–8.

22. Dubovitskaya A, Xu Z, Ryu S, et al. Secure and trustable electronic medical records sharing using blockchain. In: *Proceedings of the AMIA 2017 annual symposium*, Washington, DC, 4–8 November 2017, pp. 650–659. Bethesda, MD: American Medical Informatics Association.

23. Liu J, Li X, Ye L, et al. BPDS: a blockchain based privacy-preserving data sharing for electronic medical records. In: *Proceedings of the 2018 IEEE global communications conference (GLOBECOM)*, Abu Dhabi, United Arab Emirates, 9–13 December 2018, pp. 1–6. New York: IEEE.

24. Sukhwani H, Wang N, Trivedi KS, et al. Performance modeling of Hyperledger Fabric (permissioned blockchain network). In: *Proceedings of the 2018 IEEE 17th international symposium on network computing and applications (NCA)*, Cambridge, MA, 1–3 November 2018, pp. 1–8. New York: IEEE.

25. Baliga A, Solanki N, Verekar S, et al. Performance characterization of Hyperledger Fabric. In: *Proceedings of the 2018 Crypto Valley conference on blockchain technology (CVCBT)*, Zug, 20–22 June 2018, pp. 65–74. New York: IEEE.

26. Kuzlu M, Pipattanasomporn M, Gurses L, et al. Performance analysis of a Hyperledger Fabric blockchain framework: throughput, latency and scalability. In: *Proceedings of the 2019 IEEE international conference on blockchain*, Atlanta, GA, 14–17 July 2019, pp. 536–540. New York: IEEE.

27. Roehrs A, Costa CA, Righi RR, et al. Analyzing the performance of a blockchain-based personal health record implementation. *J Biomed Inform* 2019; 92: 103140.

28. Nakamoto S. Bitcoin: a peer-to-peer electronic cash system, 2008, www.bitcoin.org/bitcoin.pdf (2008, accessed 10 March 2020).

29. Moinet A, Darties B and Baril JL. Blockchain based trust & authentication for decentralized sensor networks, 2017, https://arxiv.org/abs/1706.01730.

30. Abay N, Akcora C, Gel Y, et al. ChainNet: learning on blockchain graphs with topological features. In: *Proceedings of the 2019 IEEE international conference on data mining (ICDM)*, Beijing, China, 8–11 November 2019, pp. 946–951. New York: IEEE.

31. Sankar LS, Sindhu M and Sethumadhavan M. Survey of consensus protocols on blockchain applications. In: *Proceedings of the 2017 4th international conference on advanced computing and communication systems (ICACCS)*, Coimbatore, India, 6–7 January 2017, pp. 1–5. New York: IEEE.

32. Androulaki E, Barger A, Bortnikov V, et al. Hyperledger Fabric: a distributed operating system for permissioned blockchains. In: *Proceedings of the thirteenth EuroSys conference*, Porto, 23–26 April 2018, vol. 30, pp. 1–15. New York: ACM.

33. Zheng Q, Li Y, Chen P, et al. An innovative IPFS-based storage model for blockchain. In: *Proceedings of the 2018 IEEE/WIC/ACM international conference on web intelligence (WI)*, Santiago, Chile, 3–6 December 2018, pp. 704–708. New York: IEEE.

34. github.com/hyperledger/caliper (accessed 10 March 2020).

35. Ampel B, Patton M and Chen H. Performance modeling of Hyperledger Sawtooth blockchain. In: *Proceedings of the 2019 IEEE international conference on intelligence and security informatics (ISI)*, Shenzhen, China, 1–3 July 2019, pp. 59–61. New York: IEEE.