Integrating Blockchain for Data Sharing and Collaboration in Mobile Healthcare Applications

Xueping Liang[1,2,3], Juan Zhao[3], Sachin Shetty[4], Jihong Liu[1,2], Danyi Li[1]†
[1]Institute of Information Engineering, Chinese Academy of Sciences, Beijing, 100093, China
[2]University of Chinese Academy of Sciences, Beijing, 100190, China
[3]College of Engineering, Tennessee State University, Nashville, TN 37209
[4]Virginia Modeling Analysis and Simulation Center, Old Dominion University, Norfolk, VA 23529

***Abstract*—Enabled by mobile and wearable technology, personal health data delivers immense and increasing value for healthcare, benefiting both care providers and medical research. The secure and convenient sharing of personal health data is crucial to the improvement of the interaction and collaboration of the healthcare industry. Faced with the potential privacy issues and vulnerabilities existing in current personal health data storage and sharing systems, as well as the concept of self-sovereign data ownership, we propose an innovative user-centric health data sharing solution by utilizing a decentralized and permissioned blockchain to protect privacy using channel formation scheme and enhance the identity management using the membership service supported by the blockchain. A mobile application is deployed to collect health data from personal wearable devices, manual input, and medical devices, and synchronize data to the cloud for data sharing with healthcare providers and health insurance companies. To preserve the integrity of health data, within each record, a proof of integrity and validation is permanently retrievable from cloud database and is anchored to the blockchain network. Moreover, for scalable and performance considerations, we adopt a tree-based data processing and batching method to handle large data sets of personal health data collected and uploaded by the mobile platform.**

***Keywords*-Healthcare, eHealth, Privacy, Permissioned Blockchain, Access Control, Scalability, Integrity, Wearable Devices, Mobile Platform**

I. Introduction

In recent years, the rise of wearable technology and the Internet-of-Things has brought great opportunities and challenges to the healthcare domain. Enabled by cloud computing and big data analytics, the data collected from individual devices contributes to big health data and valuable insights can be derived. Hospitals and medical institutions can use these data to link with other Electronic Health Record (EHR) data, such as clinical notes, to facilitate health monitoring, disease diagnoses and treatment. Health insurance companies can make detailed and strategic policies according to individual characteristics, benefiting customers to choose flexible insurance plans according to their needs.

To handle health data sharing between institutions, there is a need for a secure data sharing infrastructure. However, there are several challenges related to privacy, security, and interoperability. First, health data are highly privacy-sensitive, especially as more data are storing in a public cloud, raising the risks of data exposure. Second, current systems use centralized architecture, which requires centralized trust. Moreover, the effective integration of health data and the interoperability between healthcare systems remain a challenging task. Another challenge is that users have little control over their personal health data [1]. With the notion of self-sovereignty [2] concept and the increasing adoption of the mobile platform and wearable devices, there is an urgency to develop a new version of EHR systems with user-centric access control and privacy preservation.

Blockchain technology originated from Bitcoin [3], providing the robustness against failure and attacks, as well as functions for data provenance [4]. [5] discussed the existing vulnerabilities and propose measures to improve blockchain security. Blockchain relies on pseudoanonymity (replacing names with identifiers) and public key infrastructure (PKI), keeping the privacy of the users. The workshop [6] co-held by the Office of the National Coordinator for Health IT (ONC) and the National Institute for Standards and Technology (NIST) focused on the blockchain usage in healthcare and research, aiming to clarify the implications of blockchain as an infrastructure for healthcare use cases including privacy preservation for predictive modeling, increasing interoperability between institutions at a large scale, immutability of health records, health insurance claim process improvement, health information exchange, healthcare delivery models with artificial intelligence, identity management, monetization strategies and data provenance requirements.

In this paper, we propose a mobile user controlled,

†Corresponding author.

978-1-5386-3531-5/17/$31.00

blockchain-based system for personal health data sharing and collaboration. In the implementation, we build the system on Hyperledger Fabric [7], which is a permissioned blockchain requiring the network nodes to validate, and realizes a privacy preserving personal healthcare system with a broader coverage of the healthcare ecosystem from the end device to the cloud, as well as the emphasis of the user ownership for health data.

The rest of the paper is organized as follows. Section II introduces the overall system design, including the architecture, system entities, key establishment and system procedures. We describe the system implementation in Section III and give a performance evaluation and security analysis in Section IV. Section V presents some related work, concludes the paper and talks about the future work.

II. System Design

A. System Overview

Figure 1 is a general scenario for the user-centric personal health data sharing. Six entities are included, namely user, wearable devices, healthcare provider, insurance company, the cloud database and the blockchain network.

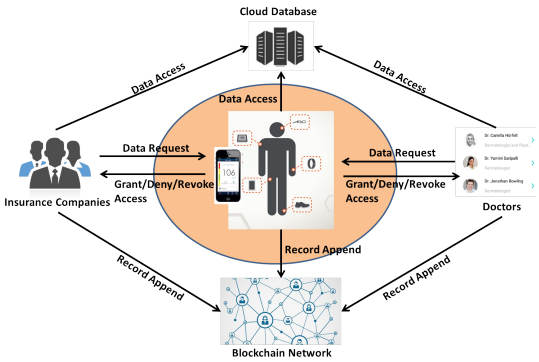

Fig. 1: User Centric Personal Health Data Sharing.

B. System Entities

- **User.** System users collect data from wearable devices which monitor users' health data such as walking distance, sleeping conditions, and heartbeat. Those data is then uploaded to the cloud database hosted on trusted platform via the mobile application. User is the owner of personal health data and is responsible for granting, denying and revoking data access from any other parties, such as healthcare providers and insurance companies. If the user is seeking medical treatment, the user would share the health data with the desired doctors. If the treatment is finished, the data access is revoked to deny further access from the doctors. Same scenario applies to user-insurance company relations. Besides, user can also record everyday activities according to a particular medical treatment such as medicine usage to share with the treatment provider for adjustment and better improvement.
- **Wearable Device.** Wearable Devices serve to transform original health information into human-readable format and then the data is synchronized by the user to their online account. Each account is associated with a set of wearable devices and possible medical devices. When a piece of health data generated, it will be uploaded to the blockchain network for record keeping and integrity protection.
- **Healthcare Provider.** Healthcare providers such as doctors are appointed by a certain user to perform medical test, give some suggestions or provide medical treatment. Meanwhile, the medical treatment data can be uploaded to the blockchain network for data sharing with other healthcare providers under the user's permission. And the current healthcare provider can request access to previous health data and medical treatment from the user. Every data request and the corresponding data access is recorded on the blockchain.
- **Health Insurance Company.** User may request a health insurance quote from health insurance companies or agents to choose a proper health insurance plan. To provide better insurance policies, insurance companies request data access from users including user health data from wearable devices and medical treatment history. Users with previous medical treatment(s) may need to pay a higher rate and the history cannot be denied by users to prevent insurance fraud. Users can choose not to share exercise information due to privacy issues but mostly they would desire to share because regular exercise can bring down the insurance pay rate. However, users cannot hide or modify medical treatment history data since those data is permanently recorded on the blockchain network and the integrity and trustworthiness is ensured. Moreover, the insurance claims can also be recorded on the blockchain.
- **Blockchain Network.** The blockchain network is used for three purposes. For health data collected from both wearable devices and healthcare providers, each of the hashed data entry is uploaded to the blockchain network for integrity protection. For personal health data access from healthcare provider and health insurance company, each of the data access request should be processed to get a permission from the data owner with a decentralized permission management protocol. The access control policies should be stored in a

distributed manner on the blockchain which ensures stability. Besides, each of the access request and access activity should be recorded on the blockchain for further auditing or investigation.

- **Cloud Database.** The cloud database stores user health related data, data requests from the healthcare provider and insurance companies, data access record and data access control policy. Data access is accountable and traceable. Once data leakage is detected, the malicious entity can be identified.

III. System Implementation

A. Personal Health Data Collection

Personal health data comes from wearable devices such as activity trackers or smart watches, and medical devices such as pacemakers or defibrillation, as well as manual user input for treatment tracking such as medicine usage and training. To synchronize the personal data to the cloud for convenient access and further process, the user first can register to the cloud service provider for an online account with enough storage capability. Figure 2 shows the data collection and synchronization architecture.

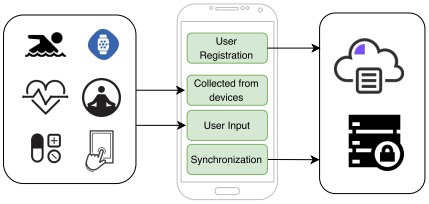

Fig. 2: Personal Health Data Collection.

B. Personal Health Data Integrity Protection and Validation

Figure 3 shows the basic data flow from the user device to the cloud server, finally anchored on the ledger with proof of integrity and validation. The health data comes from a variety of devices all day, resulting in a large number of data records. To facilitate scalable and efficient data processing and integrity protection, we develop a tree-based method for the integrity management of health data record. Some data records are batched to form a tree-based data structure and handle dynamic data enrollment. The adoption of Merkle tree [8] realizes the scalability requirement, and most importantly improves the efficiency to validate the data integrity. Merkle tree is a binary tree structure where the input is a list of hashed data records. These records are ordered by the time when they are generated. Every two records are grouped together and the hashes of the two data records become two leaf nodes of the Merkle tree and consequently constitute a high level group node with the group hash generated by concatenating two hashes. Two group nodes will follow the same way to generate a new higher level group node with a new hash. This step is repeated until there is a single hash which will become the tree root, that is, the Merkle root.

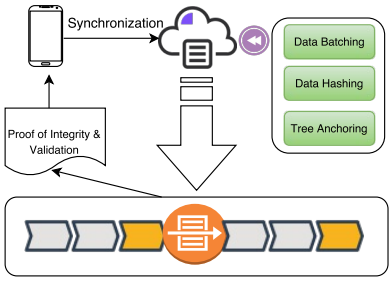

Fig. 3: Personal Health Data Integrity Protection.

Chainpoint [9] is an open standard for creating a timestamped proof of any data, file, or series of events, which proposes a scalable protocol for publishing data records on the blockchain and generating a Merkle proof for each data record. In our implementation, we anchor a list of data records to multiple Fabric channels by binding the Merkle root to a blockchain transaction and verify the integrity and existence of data without relying on a trusted third-party. The hash of data records brings two advantages. For one thing, each Merkle tree can host a large number of records since only the hash of the data record is stored. For another, the hash is an effective measure to detect changes so that once a piece of data is modified, the action can be detected easily by traversing the tree.

C. Data Sharing and Healthcare Collaboration

The user can share data with healthcare providers to seek healthcare services, and with insurance companies to get a quote for the insurance policy and to be insured. When data sharing is detected in the system, there will be an event generated to record the data access request. The event record can be described using a tuple as $\{recordhash, owner, receiver, time, location, expirydate, signature\}$. There are different types of operations from different parties, as listed in Table I.

TABLE I: Types of Operations in the Healthcare Collaboration System

Health Data	Operator	Operation
Personal Health Data	User	Update, Query
	Healthcare Provider	Query
	Insurance Company	Query
Medical History	Healthcare Provider	Update, Query
	User	Query
	Insurance Company	Query
Insurance Information	Insurance Company	Update, Query
	User	Query
	Healthcare Provider	Query

This record is then submitted to the blockchain network which is followed by several steps to transform a list of records into a transaction. A list of transactions will be used to form a block, and the block will be validated by nodes in the blockchain network. After a series of processes, the integrity of the record can be preserved, and future validation on the block and the transaction related to this record is available. Each time there is an operation on the personal health data, a record will be reflected to the blockchain. This ensures that every action on personal health data is accountable.

We implement an access control scheme by utilizing the Hyperledger Fabric membership service component and the channel scheme [7], as is shown in Figure 4. The CA, also known as the membership service provider, is responsible for membership enrollment by issuing enrollment certificates and transaction certificates for participating nodes in the Hyperledger Fabric blockchain network and participating Fabric client, and generating the access control list during channel establishment according to user settings and operations. Different access type can be specified in the certificate, such as query and update operations for chaincode execution in the channel. Chaincode is a piece of code that is deployed to Hyperledger Fabric for enabling interactions between peers and the shared ledger. There are three operations on the chaincode, including deploy, invoke and query. A chaincode can be installed on a blockchain by executing a deploy transaction while a chaincode execution is launched by invoke transactions. Channel is formed to isolate individual activities among authorized parties.

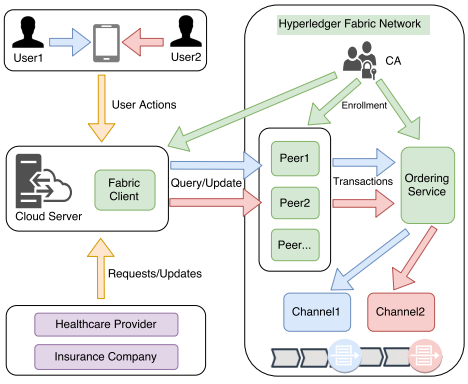

Fig. 4: Data Sharing and Collaboration Using Hyperledger Fabric and Channel for Mobile Users.

To provide isolation between different data sharing domain, the CA issues certificate to the Fabric client on the cloud server, blockchain network peers for transaction validation, and the orderers (for ordering service). We have two channels established for two users, respectively. In Figure 4, both user1 and user2 may perform data collection and synchronization on their individual mobile platforms, and the healthcare mobile application will send web requests to the cloud server to for data synchronization or query. Healthcare providers and insurance companies also communicate with the server to request or update health data and health insurance information. With the permission from users, these requests will be allowed to participate in a certain channel. The cloud server is configured with a Fabric client to communicate with the Fabric blockchain network peer. For different user activities, the data will be labeled with different channel ID to distinguish isolated domain. The query or update requests from the server will be forwarded to the Fabric network via Fabric client for transaction confirmation. Distributed peers will validate the incoming requests and propose transactions by executing chaincode. The ordering service is responsible for checking transaction signatures and order them with channel IDs. For each channel, there is a subledger, as part of the system ledger, to record all transactions in the form of blocks.

For privacy concerns, the user can selectively share health data with data requester, based on the necessity of how personal health data is required to assist the healthcare service. For example, a user's insurance history may not be important when the user is talking to a dentist. Similarly, the user's dental treatment is not necessary for skin testing or other treatment. To issue a specific certificate, the user can state clearly in the certificate what category of personal data is allowed access, whether read-only or read-write access is allowed. Moreover, in different channels, different grained information is shared. In this sense, our system provides a user-defined, fine-grained privacy protection and access control policy, enhancing the data ownership of individuals.

IV. System Evaluation

Our system adopts a user-centric model for processing personal health data using blockchain network, ensuring the data ownership of individuals, as well as data integrity. The operations on the data records are highly inter-operable and compatible with current systems. By enforcing access control policies, users can handle their personal data without worrying about the privacy issues. Meanwhile, each request and update from healthcare providers and health insurance companies are recorded and anchored to the blockchain network, making actions towards personal health data accountable.

With all the security objectives proposed in Section I achieved, it is crucial to evaluate the system performance, regarding to the scalability and efficiency of the data integrity proof generation and data validation process. We test different numbers of concurrent records with a range from 1 to 10,000. Figure 5 and 6 shows the average time cost, respectively.

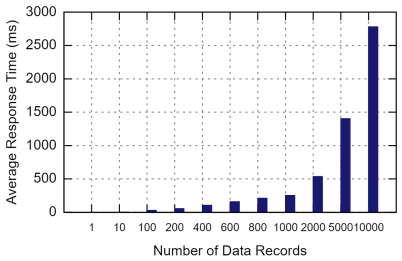

Fig. 5: Average Time for Integrity Proof Generation.

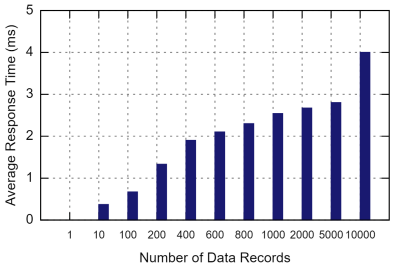

Fig. 6: Average Time for Integrity Proof Validation.

From these two figures, we can conclude that the system can handle a large dataset at low latency, which indicates the scalability and efficiency of the data process. By adopting Merkle tree method to batch data, we implement an algorithm with the computation complexity of $O(log_2 n)$. This is an important advantage when the data records are collected at a high frequencies.

V. Related Work and Conclusions

A mobile application is implemented in [10] for healthcare data sharing but is limited to patient and doctor. [11] proposes a proof of interoperability to avoid the computation cost but didn't mention the access control. [12] addresses the adoption of blockchain in social network domain but not fully explores the benefits of the blockchain. Patientory [13] is designed for healthcare storage network using Ethereum, but data privacy is highly dependent on the cryptography methods. [14] addresses the blockchain adoption in Internet of Things environment. MedRec [15] is a record management system focusing on EMRs using smart contract, but raises privacy concerns.

In this paper, we design and implement a mobile healthcare system for personal health data collection, sharing and collaboration between individuals and healthcare providers, as well as insurance companies. The system can also be extended to accommodate the usage of health data for research purposes. By adopting blockchain technology, the system is implemented in a distributed and trustless way. The algorithm to handle data records can preserve both integrity and privacy at the same time. Meanwhile, we adopt the concept of channel supported by Hyperledger Fabric to deal with the isolated communication required by specific scenarios. In the future, we will explore how to combine both personal health data and medical data together and cover a broader scenario.

VI. Acknowledgements

This work was supported by Office of the Assistant Secretary of Defense for Research and Engineering (OASD (R & E)) agreement FA8750-15-2-0120. The work was also supported by a grant from the National Natural Science Foundation of China (No.61402470) and the research project of Trusted Internet Identity Management (2016YFB0800505 and 2016YFB0800501).

References

[1] L. J. Kish and E. J. Topol, "Unpatients-why patients should own their medical data," *Nature biotechnology*, vol. 33, no. 9, pp. 921–924, 2015.

[2] J. H. Clippinger, "Why Self-Sovereignty Matters," https://idcubed.org/chapter-2-self-sovereignty-matters/, [Online; accessed 7-March-2017].

[3] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008.

[4] X. Liang, S. Shetty, D. Tosh, C. Kamhoua, K. Kwiat, and L. Njilla, "Provchain: A blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability," in *International Symposium on Cluster, Cloud and Grid Computing*. IEEE/ACM, 2017.

[5] D. K. Tosh, S. Shetty, X. Liang, C. A. Kamhoua, K. A. Kwiat, and L. Njilla, "Security implications of blockchain cloud with analysis of block withholding attack," in *Proceedings of the 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, ser. CCGrid '17. Piscataway, NJ, USA: IEEE Press, 2017, pp. 458–467. [Online]. Available: https://doi.org/10.1109/CCGRID.2017.111

[6] T. O. of the National Coordinator for Health IT (ONC), the National Institute for Standards, and T. (NIST), "Use of blockchain in healthcare and research workshop," 2016.

[7] C. Cachin, "Architecture of the hyperledger blockchain fabric," in *Workshop on Distributed Cryptocurrencies and Consensus Ledgers*, 2016.

[8] R. C. Merkle, "Protocols for public key cryptosystems," in *Security and Privacy, 1980 IEEE Symposium on*, April 1980, pp. 122–122.

[9] "Chainpoint: A scalable protocol for anchoring data in the blockchain and generating blockchain receipts," http://www.chainpoint.org/.

[10] H. Kim, H. Song, S. Lee, H. Kim, and I. Song, "A simple approach to share users' own healthcare data with a mobile phone," in *Ubiquitous and Future Networks (ICUFN), 2016 Eighth International Conference on*. IEEE, 2016, pp. 453–455.

[11] K. Peterson, R. Deeduvanu, P. Kanjamala, and K. Boles, "A blockchain-based approach to health information exchange networks," 2016.

[12] J. Zhang, N. Xue, and X. Huang, "A secure system for pervasive social network-based healthcare," *IEEE Access*, 2016.

[13] C. McFarlane, M. Beer, J. Brown, and N. Prendergast, "Patientory: A healthcare peer-to-peer emr storage network v1. 0," 2017.

[14] X. Liang, J. Zhao, S. Shetty, and D. Li, "Towards data assurance and resilience in iot using distributed ledger," in *Military Communications Conference, MILCOM 2017*. IEEE, 2017.

[15] A. Azaria, A. Ekblaw, T. Vieira, and A. Lippman, "Medrec: Using blockchain for medical data access and permission management," in *Open and Big Data (OBD), International Conference on*. IEEE, 2016, pp. 25–30.