Blockchain-Based Privacy-Preserving Healthcare Architecture

Koosha Mohammad Hossein*
Department of Electrical and Computer Engineering
University of Tehran
Tehran, Iran
kosha.hosseini@ut.ac.ir

Mohammad Esmaeil Esmaeili
Department of Electrical and Computer Engineering
University of Tehran
Tehran, Iran
me.esmaeili@ut.ac.ir

Tooska Dargahi
School of Computing, Science & Engineering
University of Salford
Manchester, UK
T.Dargahi@Salford.ac.uk

Ahmad khonsari
Department of Electrical and Computer Engineering
University of Tehran
School of Computer Science, Institute for Research in Fundamental Sciences (IPM)
Tehran, Iran
a_khonsari@ut.ac.ir

Abstract— **Since the introduction of Internet of Things (IoT), e-health has become one of the main research topics. Due to the sensitivity of patient data, preserving the privacy of patients appears to be challenging. In healthcare applications, patient data are usually stored in the cloud, which makes it difficult for the users to have enough control over their data. However, due to the General Data Protection Regulation (GDPR), it is the data subject's right to know where and how his data has been stored, who can access his data and to what extent. In this paper, we propose a blockchain-based architecture for e-health applications which provides an efficient privacy-preserving access control mechanism. We take advantage of Blockchain (BC) special features, i.e., immutability and anonymity of users, while modifying the classic blockchain structure in order to overcome its challenges in IoT applications (i.e., low throughput, high overhead and latency). To this end, we cluster the miners of BC, store and process data at the nearest cluster to the patient. While our proposal is a work in progress, we provide a security analysis of our proposed architecture.**

Keywords— Healthcare, IoT, Blockchain, Privacy

I. Introduction

Internet of Things (IoT) implies that any device can be connected to other devices and Internet at anytime and anywhere. Researchers estimated that over 75 billion devices will be connected to Internet by 2025 [1]. Aside from the advantages that the connectivity of devices in IoT has in several different scenarios, there are a variety of challenges especially in terms of security and privacy. One of the main applications of IoT is e-health, where different types of wearable sensors measure patient's blood glucose, heart rate, body temperature, blood pressure, etc. These sensors automatically collect data about users and transfer them to a central storage or cloud for further processing by physicians, nurses and medical staff [2].

Patients' data are privacy sensitive, usually stored on a server and processed remotely. This raises patients concern regarding the confidentiality and privacy of their data. This is due to the fact that several security attacks are possible in such scenarios, e.g., an attacker can intercept healthcare data on Internet; modify them and inject wrong data in healthcare data centers, or she can steal information from the remote servers [3].

The traditional privacy-preserving methods that are based on summarizing or creating noisy data [4] are not efficient in healthcare applications where users' original data are required for medical treatments. To address this issue, recently researchers proposed new privacy-preserving schemes based on Blockchain (BC) technology [5, 6]. BC is an immutable timestamp ledger of blocks that is used for storing and sharing data in a distributed manner by a peer to peer network [7]. Blocks in BC are shared across all participating nodes which eliminates the need for a central authority [8].

There are a number of challenges in applying BC to IoT scenarios: i) Network overhead, which is due to consensus operations especially in Proof-of-Works (POWs) for adding a new block and broadcasting transaction to all the participants; and ii) Low throughput, as the number of transactions that can be recorded in BC is low which is not acceptable considering the scale of IoT applications [8]. Nowadays, several research studies have adopted BC for storing users' healthcare data [9] [10]. However, to the best of our knowledge, the application of BC for improving data privacy remained uninvestigated in the literature. Researchers in [11] mentioned the irreversibility nature of BC (everybody has a copy of the ledger) as a possible reason, since this makes it hard to use BC for privacy purposes, particularly in data protection.

In this paper, we propose a BC-based access control architecture which preserves privacy of the patients. We store the hash of patient's healthcare data instead of the original data. We modify the general BC architecture to improve its efficiency in the healthcare domain as follows:

1) We cluster the miners to reduce data redundancy and prevent the involvement of all miners in the consensus operation. Moreover, in order to decrease the network overhead, we reduce the size of transactions to be light for transmitting over BC.

2) To address privacy and security challenges, we store and process data at the nearest location to the patient while each patient is assigned a pseudonym.

II. RELATED WORK

Several researchers have modified the BC architecture to overcome classic BC challenges in IoT scenarios. In [12], authors proposed optimal BC for IoT platform in the case study of smart home. They used a hierarchical structure to improve the scalability, throughput and the overhead in the BC network. They also analyzed privacy and security. In [13] authors highlighted the limitations of using common cryptographic and access control methods in cloud environment. They explored the possibility of adopting BC to protect patient's healthcare data that is stored in the cloud. The researcher in [14] propose a framework based on modified BC for IoT devices. In [10] a framework named *Ancile* is introduced that utilizes smart contract on Etherum-based BC for preserving user privacy and control access to the patients' sensitive information.

In the above-mentioned research studies, the data that is generated by IoT devices are distributed in the network and data owner does not have direct control over them. In this paper, we propose an architecture to enable the patients choose their own access control policy to define who and how can access their data.

III. PROPOSED BC-BASED ARCHITECTURE

In our proposed architecture, BC is used for storing hash of users' healthcare data and also users' access policy over their data. The policies are used for specifying who can access the users' data. Our system model is composed of the following main modules (Figure 1 shows our system model).

1) Sensors that are attached to the patient's body.
2) Patient's smartphone or PDA (Personal Digital Assistance).
3) A central server which manages and stores the patient's healthcare data, so-called IHM (IoT Health Manager).
4) Hospitals and health centers.
5) Blockchain network.
6) Miners.

A. Sensors connected to the patient's body

Each patient can carry a large number of sensors attached to his body which are responsible for gathering patient's information, such as blood pressure, and heart rate. These sensors are resource-constrained; having low energy, low storage and low processing power. By considering these limitations, sensors send the collected data using short-range communication, such as zigbee or Bluetooth, to a more powerful device, such as a smartphone or a PDA, which plays the role of a gateway to transfer data to the healthcare servers.

B. Patient's smartphone or PDA

PDA and smartphones have higher processing power and battery life compared to the sensors. They are able to carry out heavy tasks, such as cryptographic operations, and packet transmission via long-range communications, such as cellular networks to IHM (describe in Section 3.3).

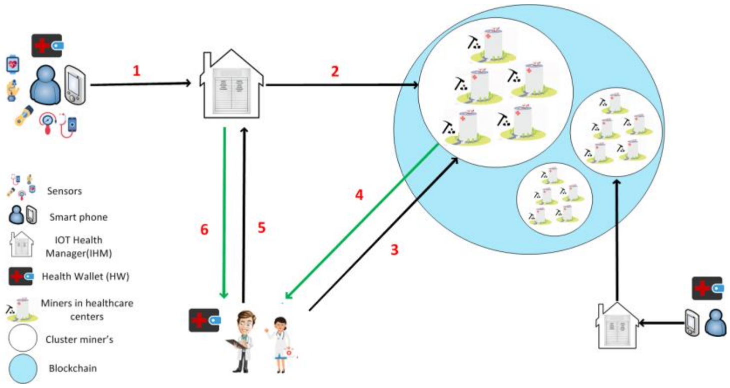

Fig. 1. Proposed system model

C. IHM (IoT Healthcare Manager)

IHM could be a simple PC for storing patient's dat, which is responsible for the following operations:

- Receiving information from smartphones and storing them.
- Performing hash and other cryptographic operations
- Transferring hash of the data and policies to the health center blockchain network.

D. Hospitals and health centers

The Hospitals and health centers can be enlarged to include all health centers of a country. They manage the users' data, BC network and miners. They receive and store hash of user's data and policies from IHMs. They are responsible for:

- Registering users (patient, medical staff, etc.) and assigning a health wallet (HW) (similar to the wallet in the bitcoin) to each of them which could be used to communicate with the BC network.
- Allocating a cluster miner to each patient (see Section IV*F. Mining and miners*).

E. Blockchain

BC is basically used for storing hash of data and access policies to improve integrity and availability of users' data. Storing such data in BC prevents against single point of failure and DoS attack. We use two separate ledgers for storing hash of users' data and policies. These two types of transactions have different structures and policy management will be easier (see Section IV. USERS' DATA POLICIES).

F. Mining and miners

In our architecture, there are a number of miners in each hospital and health center. The miners validate new transactions and record them on BC ledger. However, we could also use miners outside the health centers, though there are challenges such as how to convince and motivate miners outside the hospital to store data. Considering that we are dealing with a large number of devices in IoT, adopting POW is not affordable. Therefore, we use a new method called Practical Byzantine Fault Tolerance (PBFT) voting-based consensus [15], which involves multiple rounds of voting by all nodes of the network. This helps us to increase the network security and efficiency, as well as reducing network cost, i.e., bandwidth and processor usage.

In the traditional BC network, all nodes should store redundant data, which are the same copy of ledger, for providing data integrity. This is actually helpful in financial systems, such as cryptocurrency, but significantly reduces network throughput. This is due to the fact that each transaction must be distributed to all the nodes of the network. To avoid such an overhead, we cluster our miners and store each patient's healthcare data in one cluster. Each cluster operates independently, while all of them are simultaneous. This balances the load between different clusters.

IV. USERS' DATA POLICIES

A specific feature of our architecture is that users are able to specify access policies over their data. Users send policies in the form of a transaction to cluster miners in BC network. A policy could be defined as follows:

$$< ID_{owner}, ID_{req}, Type, T_e, < D_s, D_e >, Valid >$$

- ID_{owner}: The id of policy creator, i.e., the patient, e.g. Alice.
- ID_{req}: The id of the person who can access the data, e.g. Dr. Bob.
- $Type$: The kind of data which can be accessed by users, e.g. ECG information.
- T_e : The expiration date of policy.
- $< D_s, D_e >$: This attribute specifies time duration that the allowed users can access to data. For example, Dr. Bob is allowed to access the Alice's data from 2018/02/10 to 2018/03/10.
- $Valid$: This attribute gets a binary (0 or 1) value to determine the validity of the policy, 1 and 0 for valid and invalid policies, respectively. As BC ledger is immutable, the users can't change a specified policy before its expiration time. In case of emergency, users can insert a new policy with different $Valid$ attribute in BC. In that case, users' last policy will be considered by minors for checking access control of users' data.

V. ARCHITECTURE EXPLANATION

In this section, we describe the functioning of different modules in our architecture and their communication using a simple usecase scenario. Suppose that Alice has received an ID, a pair of private and public keys, as well as an HW from a hospital. She is wearing some sensors and has a smartphone (or a PDA) to receive data from the sensors. The following steps show how Alice's healthcare data will be registered and accessed by a medical staff (refer to Figure 1):

1) In the first step, Alice's smartphone receives data from the sensors. PDA classifies them according to the type of sensor (such as EEG and ECG) and sends them to the IHM.
2) Alice's IHM decrypts the data and stores them in a database. IHM computes the hash of the data and encrypts it using an asymmetric cryptographic method (e.g. ECC). Then it sends the encrypted data in the form of a transaction to Alice's predetermined cluster miner in BC. Each minor of the cluster receives Alice's transaction and stores it in BC.
3) If a medical staff wants to access Alice's healthcare data, he should create a transaction (i.e., request a transaction) specifying Alice's ID. This transaction is sent to the cluster in which the Alice's data is stored.
4) Alice's data policies will be checked by the miners. If Alice's policies contain the requesting medical staff's id, then the data and the location of the data is encrypted with medical staff's public key and will be forwarded to the medical staff.
5) After receiving the response about the access transaction, the medical staff can access to the hash of Alice's data, and decrypt it with his private key.
6) Finally, medical staff sends a message including hash of Alice's data to Alice's IHM. Alice's IHM decrypts the message and retrieves hash value. If this hash value is valid, it returns Alice's data, otherwise returns an access denied message.

VI. SECURITY AND PRIVACY ANALYSIS

In this section, we discuss the performance of the proposed architecture in terms of security and privacy. We analyse the CIA security triad (Confidentiality, Integrity, and Availability) of our architecture to show its resilience against several attacks. First, we explain the CIA triad aspects in our specific context.

- **Confidentiality:** Confidentiality means that the messages should be accessed only by authorized users. To provide confidentiality, we encrypt the communication between modules, which guarantees the users' data (generated by the sensors) will be protected against sniffing by unauthorized users.
- **Integrity:** The data integrity insures no one can change the stored data without permission. BC is inherently resistant to modification of the data. BC ledgers are immutable, so the BC data cannot be updated.
- **Availability:** The availability of the data is due to thedistributed feature of the adopted BC, where data is stored in all miner nodes instead of a central server.

We consider six attack scenarios, and analyse the resilience of our architecture against each of them (refer to Table 1).

TABLE 1. Security Analysis of the Proposed Architecture

Attack	Definition	Defence	Resilience
Appending	Attacker compromises a miner and generates blocks with fake transactions.	Due to the usage of private BC, as well as a good number of miners in each cluster, users cannot generate fake blocks, whereas any transaction is only verified by miners in the clusters.	High
Denial of Service (DOS)	The attacker uses HW to generate large number of transactions to disrupt the BC network.	• Limited number of transactions that can be sent by HW. • Due to clustering, traction flooding will affect only a subset of clusters instead of all the nodes of the network. • After receiving a few messages from a specific user, the rest of the user's transactions will be rejected.	High
Distributed DOS (DDOS)	This is a distributed version of the above attack.	• Limited number of transactions can be sent by a valid HW. • The miners check that received transactions have been produced by a valid ID and HW.	Moderate
Modification Attack	Malicious modification or removal of the stored hash and policies of the patient's data	Due to the usage of BC immutable ledger	High
Public BC Modification	Attacker advertises a false ledger and makes it as the longest ledger.	We use private BC within the hospitals and health centers, so the miners are not from outside the organization to create malicious block.	High
51% attack	The attacker controls more than 51% of miners and tries to compromise the consensus algorithm and generate fake block	The probability that all the cluster miners from various health centers are compromised to change the data is very low based on PBFT and consensus methods [15].	High

VII. Conclusion

In this work, motivated by the privacy challenge of patients' healthcare data in e-health use case, we proposed a new architecture based on the blockchain technology. Our proposed architecture (which is a work in progress) enables users to have full control over their sensitive data that are collected by their wearable sensors. In this architecture, users can store access control policies over their data in BC to specify who can accessed their data and to what extent. Considering different attack models, we discussed the security of our architecture against those attacks. In the future we are going to implement the architecture and perform experimental analysis to evaluate the performance of the proposed architecture.

References

[1] STATISTA, "INTERNET OF THINGS (IOT) CONNECTED DEVICES INSTALLED BASE WORLDWIDE FROM 2015 TO 2025 " URL:HTTPS://WWW.STATISTA.COM/STATISTICS/471264/IOT-NUMBER-OFCONNECTED-DEVICES-WORLDWIDE/, ACCESSED: 2018-01-25.

[2] S. M. R. Islam, D. Kwak, M. H. Kabir, M. Hossain and K.-S. Kwak, "The Internet of Things for Health Care: A Comprehensive Survey," *IEEE Access*, vol. 3, pp. 678 - 708, June 2015.

[3] K. Abouelmehdi, A. Beni-Hessane and H. Khaloufi, "Big healthcare data: preserving security and privacy," *Journal of Big Data*, December 2018.

[4] Y.-A. d. Montjoye, E. Shmueli, S. S. Wang and A. S. Pentland, "Protecting the Privacy of Metadata through SafeAnswers," *PLOS ONE*, vol. 9, July 2014.

[5] X. Yue, H. Wang, D. Jin, M. Li and W. Jiang, "Healthcare Data Gateways: Found Healthcare Intelligence on Blockchain with Novel Privacy Risk Control," *Journal of Medical Systems*, 2016.

[6] T.-T. Kuo and L. Ohno-Machado, "ModelChain: Decentralized Privacy-Preserving Healthcare Predictive Modeling Framework on Private Blockchain Networks," *arXiv:1802.01746*, Feb 2018.

[7] A. Kosba, A. Miller, E. Shi, Z. Wen and C. Papamanthou, "Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts," in *2016 IEEE Symposium on Security and Privacy (SP)*, San Jose, CA, USA, August 2016.

[8] A. Dorri, S. S. Kanhere, R. Jurdak and P. Gauravaram, "LSB: A Lightweight Scalable BlockChain for IoT Security and Privacy," *arXiv:1712.02969* , Dec 2017.

[9] P. Zhang, J. White, D. C.Schmidt, G. Lenz and S. Rosenbloom, "FHIRChain: Applying Blockchain to Securely and Scalably Share Clinical Data," *Computational and Structural Biotechnology*, vol. 16, pp. 267-278, 2018.

[10] G. G.Dagher, J. Mohler, M. Milojkovic and P. BabuMarella, "Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology," *Sustainable Cities and Society*, vol. 39, pp. 283-297, May 2018.

[11] P. J.Taylor, T. Dargahi, A. Dehghantanha, R. M.Parizi and K.-K. R. Chood, "A systematic literature review of blockchain cyber security," *Digital Communications and Networks*, February 2019.

[12] A. Dorri, S. S. Kanhere and R. Jurdak, "Towards an Optimized BlockChain for IoT," *IoTDI '17 Proceedings of the Second International Conference on Internet-of-Things Design and Implementation* , pp. 173-178, April 2017.

[13] C. Esposito, A. D. Santis, G. Tortora, H. Chang and K.-K. R. Choo, "Blockchain: A Panacea for Healthcare Cloud-Based Data Security and Privacy?," *IEEE Cloud Computing*, vol. 5, no. 1, pp. 31-37, Jan./Feb. 2018.

[14] A. D. Dwivedi, G. Srivastava and S. D. a. R. Singh, "A Decentralized Privacy-Preserving Healthcare Blockchain for IoT," *Sensors — Open Access Journal* , 2019 Jan.

[15] M. S. Ali, M. Vecchio, M. Pincheira, K. Dolui, F. Antonelli and a. M. H. Rehmani, "Applications of Blockchains in the Internet of Things: A Comprehensive Survey," *IEEE Communications Society*, 2018.