A Systematic Review of Privacy-Preserving Healthcare Data Sharing on Blockchain

Mustafa Tanrıverdi [1,*]

[1] Gazi University, Ankara, TURKEY; mustafatanriverdi@gmail.com

Abstract

Sharing the electronic health data helps to increase the accuracy of the diagnoses and to improve the quality of health services. This shared data can also be used in medical research and can reduce medical costs. However, health data are fragmented across decentralized hospitals, this prevents data sharing and puts patients' privacy at risks. In recent years, blockchain has revealed solutions that make life easier in many areas thanks to its distributed, safe and immutable structure. There are many blockchain-based studies in the literature on providing data privacy and sharing in different areas. In some studies, blockchain has been used with technologies such as cloud computing and cryptology. In the field of healthcare blockchain-based solutions are offered for the management and sharing of Electronic health records. In these solutions, private and consortium blockchain types are generally preferred and Public Key Infrastructure (PKI) and encryption are used for data privacy. Within the scope of this study, blockchain-based studies on the privacy preserving data sharing of health data were examined. In this paper, information about the studies in the literature and potential issues that can be studied in the future were discussed. In addition, information about current blockchain technologies such as smart contracts and PKI is also given.

Keywords: Blockchain, Data Sharing, Data Privacy, Healthcare

1. Introduction

Electronic health record (EHR) is the digital version of the patient's records such as physical examinations, treatment, etc. Sharing the electronic health records (EHRs) helps to increase the accuracy of the diagnoses and to improve the quality of health services. This shared data can also be used in medical research and can reduce medical costs. EHRs have a great potential to increase the quality of life in terms of health services. The EHRs have become much more important for the prevention, diagnosis and treatment of global pandemics such as coronavirus (Covid-19), which we have experienced recently. World Health Organization reports are already identified personal health record as an asset [1], [2]. Therefore, sharing of EHRs and data privacy should be considered an important issue other than medical use. When EHR is considered as a personal asset, patients should own and fully control their health records, While in reality EHRs are controlled by hospital information systems, cloud service providers and device manufacturers [3], [4]. Storing EHRs in different central data stores is risky in terms of privacy. Administrators of service provider companies can use EHRs commercially or database administrators can access this data. In addition, these data stores and service providers may be targeted against cyber attacks [5]. For these reasons, while there are already potential risks in terms of data privacy, it becomes even more difficult to ensure privacy-preserving data sharing between centralized hospital information systems.

Received: june 2, 2020 Revised: Augast 25, 2020 Accepted: September 2, 2020

DOI: 10.5281/zenodo.4014251 **Special issue on Research and Innovations in Data Protection and Privacy in Healthcare (RIDAPPH)**

Nowadays, in parallel with the rapid development of information technologies, attack methods that threaten privacy-preserving data sharing are also renewed. As a precaution against these current threats, new technologies should be followed and appropriate solutions should be developed. Blockchain technology, which has become popular recently and has been applied in many areas in different fields, will be very convenient to use for data management in the healthcare field. Blockchain provides data security, immutability and transparency thanks to its distributed structure. Thanks to blockchain features such as smart contracts and PKI, EHRs can be stored encrypted and protected against external attacks. At the same time, thanks to the transparent structure offered, the information of who has access to the EHRs of a patient can be accessed.

2. Data Sharing and Privacy

As a result of the development of internet and IoT technologies, the size of the personal data produced is rapidly increasing. Using cloud environments for storing and sharing these large-sized data provides many convenience and economic savings [6]. However, this method brings some difficulties in terms of data security and ensuring privacy [7]. The central structure of existing cloud providers, cyber attacks or technical failures can cause single point failure. There is also a trust issue. Users have concerns about where their data is stored and who accesses it [8]. The disclosure of millions of users' data leakage by Facebook Inc. [9] and the exposure of 500,000 users due to a bug in Google Plus [10] show that these concerns are justified.

Any healthcare application needs to store, process and analyze users' EHRs. At this stage, while handling the EHRs of patients, the basic requirements of the European Union General Data Protection Regulation (EU GDPR) [11] must be complied [8]:

- Personal data should be processed only with the consent of the data owner.
- All technologies connected to user data must preserve user privacy.
- The owner of the data must have full authority over their data. The owner determines which third parties can access his data and how they will use it.
- Data owners have the right to delete their own data

Fortunately, a new and popular technology called blockchain offers opportunities for the solution of these problems with decentralized architecture.

3. Blockchain

Blockchain technology, which became known and popular for the first time thanks to crypto currencies, attracted great interest from various international organizations, industries and institutions. The number of blockchain-based solutions has increased rapidly in many areas in recent years. Thanks to its solutions and features, blockchain has even been described by some researchers as a more powerful technology than the Internet [12]. The report, published by Allied Market Research, stated that the blockchain market, which was $228 million in 2016, could reach $5.4 billion with a huge increase in 2023 [13]. According to Nakamoto, the blockchain is a distributed and transparent data structure in which every transaction information is recorded and shared encrypted by the participants on the network. [14]. Reyna et al. defined the blockchain as a distributed, transparent, immutable and secure data structure where the stakeholders in the network verify the reliability of transactions [15]. There are many similar blockchain definitions in the literature, from a technical point of view, it would be correct to define blockchain as a combination of decentralization mechanism, combination of a cryptographic algorithms and distributed database [16].

Received: june 2, 2020 Revised: Augast 25, 2020 Accepted: September 2, 2020
DOI: 10.5281/zenodo.4014251 **Special issue on Research and Innovations in Data Protection and Privacy in Healthcare (RIDAPPH)**

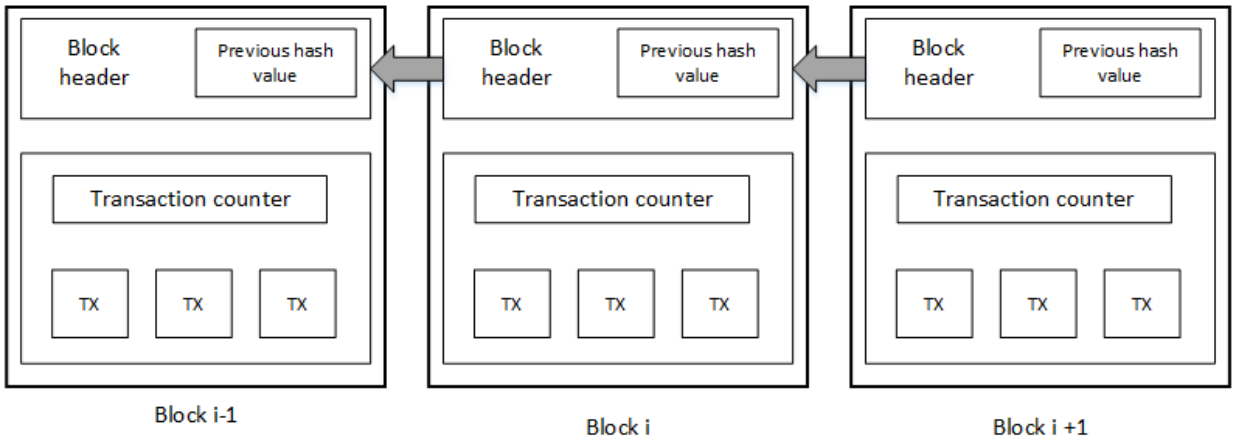

Figure 1. The general architecture of the blockchain

The general architecture of the blockchain is as in Figure 1. The hash value of the previous block is included in the header of each block in the chain. In this way, the integrity of the chain is ensured and malicious data change is prevented.

The advantages of the blockchain can be listed as follows [17].

- A copy of all the data in the blockchain is also saved by all stakeholders and anyone with permission can access this data. In this way, data security, transparency and immutability are ensured.
- Thanks to consensus algorithms and digital signatures, stakeholders can trust each other. This eliminates the need for third-party representatives.
- As a result of the removal of third-party representatives, dependency is eliminated, financial gain can be achieved and data security is increased.
- Advanced applications can be developed thanks to the smart contracts provided by blockchain.

In recent years, the number of blockchain-based studies conducted by researchers and technology companies in areas such as finance, public services, supply chain, health and education has been increasing rapidly. Blockchain-based solutions are also recommended to overcome problems arising from data sharing and privacy.

4. Privacy-Preserving Healthcare Data Sharing Based on Blockchain

With its features like decentralized control, immutability, cryptographic security, fault tolerance and capability to run smart contracts, the blockchain has an significant potential in privacy–preserving data sharing as well as in other areas. There are many studies in this area, the study by Zheng et al. is one of these studies [18]. In this study, it was recommended to keep personal data in the cloud encrypted and use encryption and sharing algorithms on blockchain to share and manage this data. In this way, the limitations of data storage on the blockchain were prevented and encryption transactions were made with the participation and compromise of all stakeholders using Paillier cryptosystem. According to the experimental results of the presented solution, it has been seen that distributed encrypted transactions can be performed with high performance with Paillier cryptosystem.

Considering its unique features and solutions, blockchain has an important potential for healthcare applications. The number of studies involving blockchain solutions is also increasing on the sharing and privacy of EHRs. Information about some of them is given below.

Received: june 2, 2020 Revised: Augast 25, 2020 Accepted: September 2, 2020
DOI: 10.5281/zenodo.4014251 **Special issue on Research and Innovations in Data Protection and Privacy in Healthcare (RIDAPPH)**

Liu et al. proposed a solution named BPDS [19]. In BPDS, EMRs are stored in cloud, and the indexes of these records are reserved in a consortium blockchain. In this solution, patients participate in the sharing of EMRs voluntarily and anonymously. Authorized healthcare institutions, patients and medical workers may request patients' EMRs for personal health plans, better clinical treatment, or medical research. To implement BPDS, all user groups should register unique accounts and create their key pairs (PK, SK) at the first. With these unique keys, patients will be able to approve or prohibit requests for access to EMRs received by smart contracts. In this way, information on which users can access which EMRs can be stored in the consortium blockchain. The structure of BPDS, which consists of three layers as data acquisition, data storage and data sharing, is shown in Figure 2. Blockchain network and cloud storage are located together in the data storage layer. Here, the data is encrypted with key data and timestamps and stored in the cloud, and patients can manage their own data with key pairs and smart contracts.

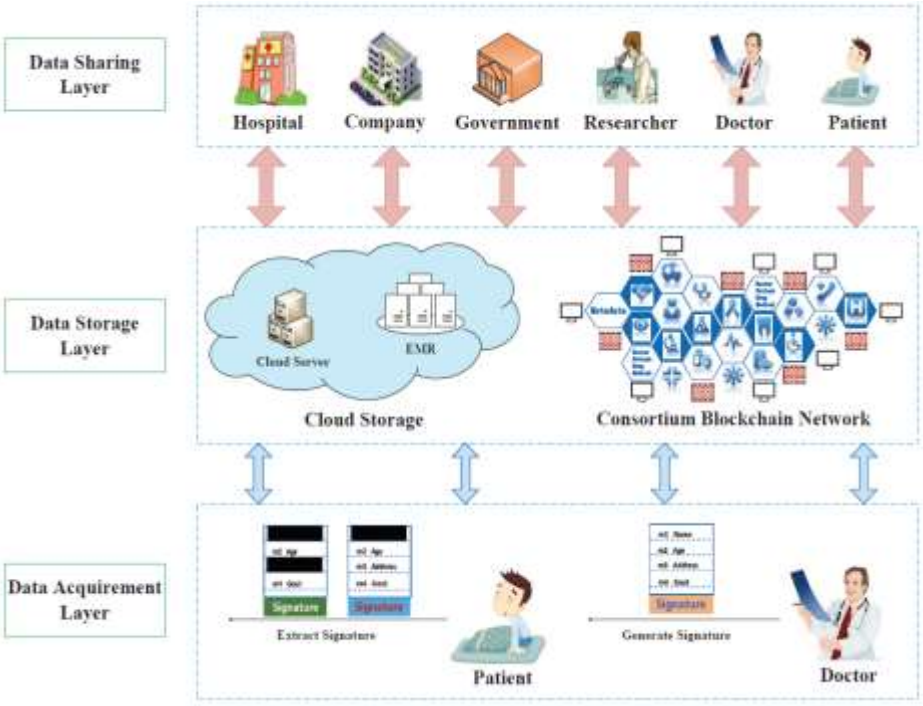

Figure 2. System architecture of BPDS [19]

Liu et al. proposed a blockchain-based medical data sharing and protection scheme [20]. In this study, the researchers offered a solution in which the electronic health system of hospitals are located in a private blockchain to improve the sharing and protection of the data. Proxy re-encryption technology and improved consensus mechanism methods have been applied to increase data security and privacy. A symptoms-matching mechanism is designed for data sharing and communication of patients who show same disease symptoms in different hospitals. Detailed information about general network model and architecture of the system where system management, patients and hospitals are actors are given. The general network model of the presented scheme is shown in Figure 3.

Received: june 2, 2020 Revised: Augast 25, 2020 Accepted: September 2, 2020
DOI: 10.5281/zenodo.4014251 **Special issue on Research and Innovations in Data Protection and Privacy in Healthcare (RIDAPPH)**

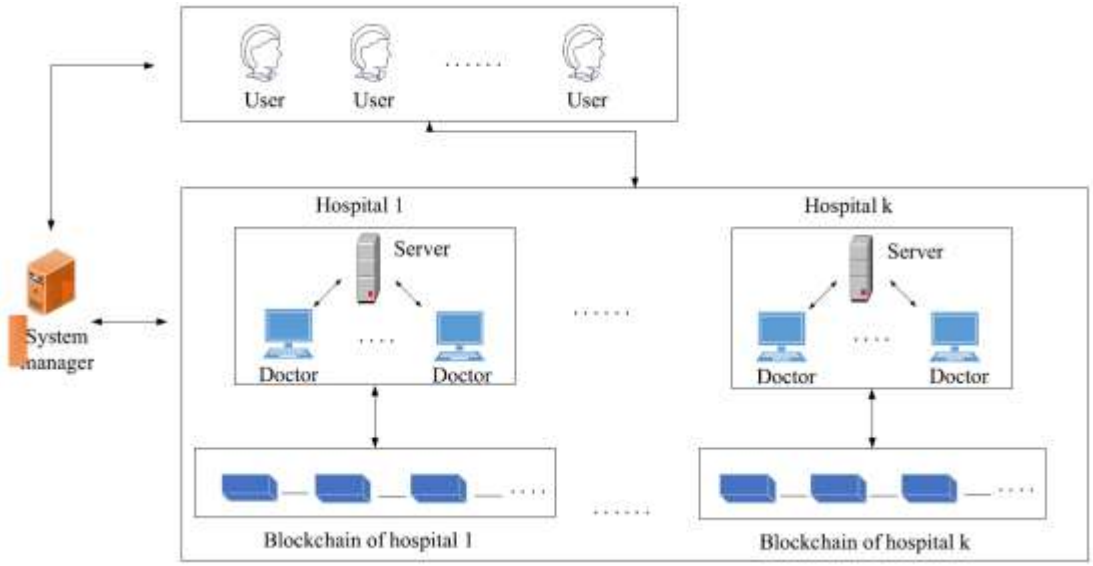

Figure 3. General network model of blockchain-based medical data sharing and protection scheme [20]

Nguyen et al. have conducted a study offering a blockchain solution for mobile cloud-based E-health systems, which has been frequently used recently [21]. In this study, an android application was developed in which EMRs recorded as distributed in Amazon cloud services, are managed with smart contracts on the ethereum blockchain. In the presented application, identity authentication is performed with smart contracts for patients through the mobile application. Data such as patient and location of the EHRs obtained from the wearable sensors is stored in the blockchain in encrypted form and this data are transferred to the cloud. The smart contracts required for the doctors' authentication to the EHRs are also checked. Thus, data such as who has access to which EHRs is kept in the blockchain. When the experimental results of the developed application were analyzed, it was observed that data management and sharing were provided in a mobile cloud environment with high performance and safely.

5. Conclusions

Summary information about several blockchain solutions on sharing EMRs and protecting privacy is given above. When these publications and other related studies are examined, it is seen that blockchain technology is generally supported by other technologies such as cloud, mobile and encryption. Due to the nature of the blockchain, a copy of each data is saved by all stakeholders. In this case, a large data store will be required in a blockchain network where many health institutions are involved as stakeholders. As a solution to this limitation of the blockchain, it was preferred to keep only the identity information of the EMRs on the blockchain and the original of the data to be stored in the cloud. Patients, doctors and other users use public addresses and private keys to authenticate on the blockchain. Manually entering and storing these texts causes difficulties for users. In some studies, mobile applications have been used to overcome this difficulty. The transparent structure of the public blockchain, which allows everyone to participate, is not suitable for storing personal data such as EMRs. Instead, private blockchain networks and consortium algorithms were preferred in the studies.

Some of the studies in this area include experimental results regarding the proposed solutions. According to the evaluations made with parameters such as number of users and file size, it was stated that the proposed solutions showed sufficient performance. Blockchain, which is a new technology, has real life applications in other areas and the number of these applications is gradually increasing. After the theoretical and experimental studies, it is thought that real life applications will be encountered in the field of privacy-preserving sharing of EMRs.

Received: june 2, 2020 Revised: Augast 25, 2020 Accepted: September 2, 2020

DOI: 10.5281/zenodo.4014251 **Special issue on Research and Innovations in Data Protection and Privacy in Healthcare (RIDAPPH)**

Funding: "This research received no external funding"

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] V. Blessing, A. Davé, and P. Varnai, "Evidence on mechanisms and tools for use of health information for decision-making," *Evid. Mech. tools use Heal. Inf. Decis.*, 2017.

[2] K. Michelsen, H. Brand, P. Achterberg, and J. Wilkinson, "Promoting better integration of health information systems: best practices and challenges," *Copenhagen WHO Reg. Off. Eur. Heal. Evid. Netw. Synth. Rep.*, p. 40, 2015.

[3] J. Zhang, N. Xue, and X. Huang, "A Secure System for Pervasive Social Network-Based Healthcare," *IEEE Access*, vol. 4, pp. 9239–9250, 2016.

[4] U. Varshney, "Pervasive healthcare and wireless health monitoring," *Mob. Networks Appl.*, vol. 12, no. 2–3, pp. 113–127, Jun. 2007.

[5] P. Esmaeilzadeh and T. Mirzaei, "The potential of blockchain technology for health information exchange: Experimental study from patients' perspectives," *Journal of Medical Internet Research*, vol. 21, no. 6. Journal of Medical Internet Research, 01-Jun-2019.

[6] Y. Yu, J. Ni, M. H. Au, Y. Mu, B. Wang, and H. Li, "Comments on a Public Auditing Mechanism for Shared Cloud Data Service," *IEEE Transactions on Services Computing*, vol. 8, no. 6. Institute of Electrical and Electronics Engineers, pp. 998–999, 01-Nov-2015.

[7] B. R. Kandukuri, P. V. Ramakrishna, and A. Rakshit, "Cloud security issues," in *SCC 2009 - 2009 IEEE International Conference on Services Computing*, 2009, pp. 517–520.

[8] I. Makhdoom, I. Zhou, M. Abolhasan, J. Lipman, and W. Ni, "PrivySharing: A Blockchain-based Framework for Integrity and Privacy-preserving Data Sharing in Smart Cities," pp. 363–371, Jun. 2020.

[9] "Facebook data breach: Hundreds of millions of records exposed on Amazon server, according to UpGuard cybersecurity research firm - CBS News." [Online]. Available: https://www.cbsnews.com/news/millions-facebook-user-records-exposed-amazon-cloud-server/. [Accessed: 17-Jun-2020].

[10] "Google bug exposed the information of up to 500,000 users." [Online]. Available: https://www.cnbc.com/2018/10/08/google-bug-exposed-the-information-of-up-to-500000-users.html. [Accessed: 17-Jun-2020].

[11] "General Data Protection Regulation (GDPR) – Official Legal Text." [Online]. Available: https://gdpr-info.eu/. [Accessed: 17-Jun-2020].

[12] K. Sultan, U. Ruhi, and R. Lakhani, "Conceptualizing Blockchains: Characteristics and Applications," in *11th IADIS International Conference on Information Systems*, 2018, pp. 49–57.

[13] "Blockchain Distributed Ledger Market Size by Type, End-User," *Allied Market Research Report*, 2017. [Online]. Available: https://www.alliedmarketresearch.com/blockchain-distributed-ledger-market. [Accessed: 14-Nov-2018].

[14] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System." [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.221.9986. [Accessed: 08-Nov-2018].

[15] A. Reyna, C. Martín, J. Chen, E. Soler, and M. Díaz, "On blockchain and its integration with IoT. Challenges and opportunities," *Futur. Gener. Comput. Syst.*, vol. 88, pp. 173–190, Nov. 2018.

[16] M. Tanriverdi and A. Tekerek, "Implementation of Blockchain Based Distributed Web Attack Detection Application," in *1st International Informatics and Software Engineering Conference: Innovative Technologies for Digital Transformation, IISEC 2019 - Proceedings*, 2019.

[17] V. Gatteschi, F. Lamberti, C. Demartini, C. Pranteda, and V. Santamaria, "To Blockchain or Not to Blockchain: That Is the Question," *IT Prof.*, vol. 20, no. 2, pp. 62–74, Mar. 2018.

[18] B. K. Zheng *et al.*, "Scalable and Privacy-Preserving Data Sharing Based on Blockchain," *J. Comput. Sci. Technol.*, vol. 33, no. 3, pp. 557–567, May 2018.

[19] J. Liu, X. Li, L. Ye, H. Zhang, X. Du, and M. Guizani, "BPDS: A Blockchain Based Privacy-Preserving

Received: june 2, 2020 Revised: Augast 25, 2020 Accepted: September 2, 2020

DOI: 10.5281/zenodo.4014251 **Special issue on Research and Innovations in Data Protection and Privacy in Healthcare (RIDAPPH)**

Data Sharing for Electronic Medical Records," in *2018 IEEE Global Communications Conference, GLOBECOM 2018 - Proceedings*, 2018.

[20] X. Liu, Z. Wang, C. Jin, F. Li, and G. Li, "A Blockchain-Based Medical Data Sharing and Protection Scheme," *IEEE Access*, vol. 7, pp. 118943–118953, Aug. 2019.

[21] D. C. Nguyen, P. N. Pathirana, M. Ding, and A. Seneviratne, "Blockchain for Secure EHRs Sharing of Mobile Cloud Based E-Health Systems," *IEEE Access*, vol. 7, pp. 66792–66806, 2019.

Received: june 2, 2020 Revised: Augast 25, 2020 Accepted: September 2, 2020

DOI: 10.5281/zenodo.4014251 **Special issue on Research and Innovations in Data Protection and Privacy in Healthcare (RIDAPPH)**