BLOCKCHAIN TECHNOLOGY AS A CATALYST FOR TRANSPARENT MEDICAL DATA SHARING

S. Sangeetha
Research Scholar, Department of Computer Science, School of Computing Sciences
Vels Institute of Science, Technology and Advanced Studies (VISTAS), Chennai, Tamil Nadu, India
sangeethasan284@gmail.com

Dr. T. Sree Kala
Professor, Department of Computer Science, School of Computing Sciences
Vels Institute of Science, Technology and Advanced Studies (VISTAS), Chennai, Tamil Nadu, India
sreekalatm@gmail.com

https://doi.org/10.34293/9789361639715.shanlax.ch.023

Abstract

Medical blockchain data-sharing uses blockchain technology to securely share electronic medical data. Blockchain is a decentralised digital ledger that uses consensus algorithms and cryptographic techniques to guarantee data security, transparency, and traceability. This approach has therefore drawn a lot of interest and study in the last few years.

Current approaches to medical blockchain data-sharing differ in how they store and transmit original data, which leads to variations in performance and privacy. To classify these methods, we divide them into on-chain and off-chain sharing, based on where the original data is stored. Off-chain sharing can be further broken down into on-cloud sharing and local sharing, depending on whether the data is moved.

In the final part of the analysis, we examine the basic processes and research topics related to each method. We also highlight the challenges that existing methods face and suggest potential directions for future research in the field.

Keywords: *Data-sharing; Electronic medical data, Federated learning.*

Introduction

Cloud computing, big data, and the Internet of Things (IoT) are making digital medical technology develop rapidly. The IoT sensors such as wearables are generating huge volumes of medical data. Analyzing such data with artificial intelligence can be used in intelligent medical systems such as disease detection and remote health measurement.This assists medical institutions to offer personalized treatment intervention to the patients.

Electronic medical information should be shared with protection of privacy. Encryption of the data and storing it in the cloud is one of the most common methods of guaranteeing privacy in order to share the information. The private clouds allow medical institutions to exchange data between other patients and staff, whereas the public one can be utilized to exchange data with other institutions. Cloud storage facilitates remote sharing of data and decongests the local storage.

Nevertheless, in the use of cloud storage medical institutions lose certain rights and control over the data as well. This can be resolved through restricting access to the data by applying access control methods such as role based or attribute based access. These controls will ensure that only the authorised users can see or modify the data.

These safeguards do not ensure that the system does not rely on centralised systems or trusted parties to perform supervision and control and therefore will lead to trust issues. Since consumers might not entirely trust cloud service providers to implement and enforce security controls to protect the privacy and integrity of data, cloud environments also face severe trust concerns. Blockchain technology refers to a digital registry that is transparent, auditable and decentralised. Electronic medical data sharing can solve issues such as trust and access control using blockchain technology. With blockchain, parties can communicate with each other without the middlemen involved and as a result, this assists in protecting and preserving the data. Smart contracts based on blockchain can improve and automatize the sharing of data without relying on third parties. It can also develop thorough access controls to control who can view or share data. Blockchain and its ability to remain undamaged and transparent ensures that every user can verify the integrity and validity of the information and address a lack of trust.

As medical blockchain expanded, numerous papers have been written regarding it in various viewpoints. Jin et al. categorized medical blockchain as permission-based and non-permission-based and discussed their advantages and disadvantages, Abu-Elezz et al. examined the advantages of medical blockchain to patients and organizations, the organizational, social, and technological risks of this technology. Chukwu et al. have given technical and architectural break down of privacy, security, cost, and performances of different medical blockchain systems.

The issues and uses of medical blockchain have also been examined in other studies. Attaran raised the most critical issues like access control, interoperability, data integrity and data sources. Haleem et al. studied the benefits and processes of medical blockchain and investigated fourteen possible applications. Rahman et al. concentrated on blockchain application to the Internet of Medical Things (IoMT) by describing the problems of privacy leakage, energy usage, and scalability.

We have made the following major contributions:

- Using blockchain and federated learning to enable sharing of electronic medical data securely, transparently, and traceably.
- The classification of medical blockchain data-sharing as on-chain and off-chain, the latter further divided into on-cloud or local sharing, depending on the storage location of the original data.
- Providing a detailed discussion of each approach and associated studies, determining the existing issues, and describing possible ways of how future studies might develop.

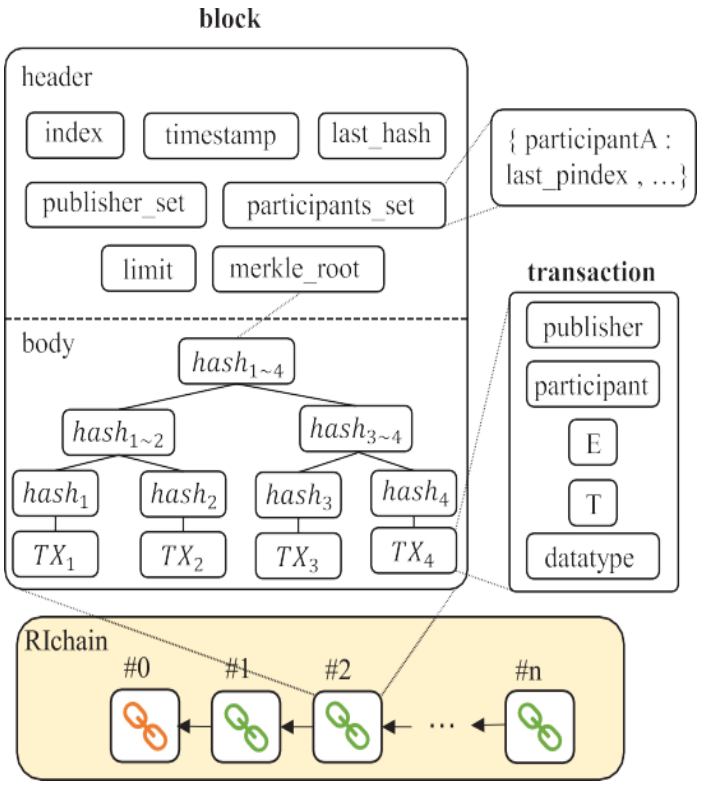

Figure 1: A blockchain's structure is made up of blocks.

The variety of blockchain has increased with technology, and more specific solutions have been developed to meet the individual needs of various industries. There are many types of blockchains, such as consortium, private, and public chains with distinct properties and usage. These versions provide a firm foundation on the wide use of blockchain technology. The blockchain is increasingly gaining importance in sectors such as real estate, healthcare, supply chains and banking. It enhances data safety and reliability, increases transparency and streamlines processes, advancing the digital revolution in sectors.

A brief overview of Federated Learning

Federated learning represents a state-of-the-art distributed machine learning that integrates training data of multiple devices or models to build a global model, all without compromising the privacy of user data. The flow of federated learning appears in Figure 2. Google originally conceived this idea in 2016 in order to deal with privacy and performance concerns whenever Android phones are upgraded to local models. The topic of data silos has been addressed by federated learning, which involves a combination of machine learning and privacy-preserving technology, as the problem with individual privacy and data safety gains momentum across the world.

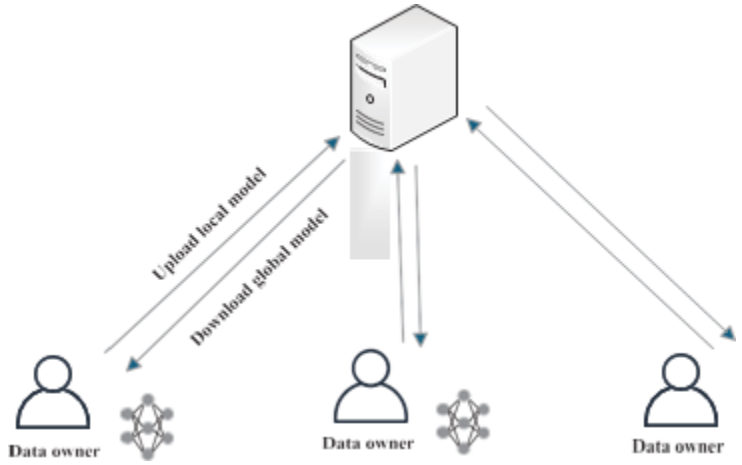

Figure 2: Federated learning framework

Table 1: Comparison of different review methods				
Research work	Introduction	Classification Comprehensive Bibliographic		Challenge
	to blockchain	of data-sharing analysis of comparison		discussion
		methods data-sharing		
		methods		
Jin et al. [25]	C	C × C		C
Xi et al. [40]	C	× × C		C
Młodawski	C	× × ×		C
et al. [41]				
Dubovitskaya	C	× × ×		C
et al. [42]				

Moreover, to make it easy to describe, based on the initial medical data storage and transmission in the medical blockchain, we divide the existing medical blockchain practices into three categories, namely, on-chain sharing, cloud sharing, and local sharing. Table 2 presents a comparison of these three methods of medical blockchain implementation.

Table 2: Comparison of different implementation methods of medical blockchain

Method	**Storage location**	**Metadata transfer**	**Size limit**	**Data security**	**Storage cost**	**Sharing performance**
On-chain sharing	On-chain	Yes	Yes	High	High	Normal
On-cloud sharing	Off-chain	Yes	No	Normal	Normal	High
Local sharing	Off-chain	No	No	High	Normal	High

On-Chain Sharing

On-chain sharing method employs smart contracts to encrypt the electronic medical data and then the data is stored on the blockchain to share safely. The nodes of the blockchain work together in order to maintain the data. Decentralised, anonymised and inaccessible properties of the blockchain ensure that the encrypted medical data is secure as long as it is being distributed.

Smart contracts are the avenue through which data requesters request access to data. The smart contract locates the relevant data on the blockchain, decrypts it and transmits the data to the requester after the data owner grants authority. This distributed storage method is more reliable by eliminating single points of failure. Moreover, the decentralised characteristics of the blockchain ensure that the information-sharing process will rely on the agreement among the involved nodes only, eliminating the role of middlemen and overcoming the issue of trust.

The solution has the capability of standardising multi-source data-sharing formats, and thus minimise the heterogeneity of data, by encrypting electronic medical data on the blockchain using smart contracts. This is an easy way of maintaining data integrity and trust and still provides a secure and efficient transfer of data.

Local Sharing

In order to reduce security threats during transmission and sharing of original electronic medical data, a local procedure of sharing, coupled with federated learning, has been established. This strategy uses federated learning to train local electronic medical data, whereby a new global model can be developed by sharing and combining model parameters trained locally by data owners through a central aggregator. This technique does not limit the size of data it uses. Federated learning is further improved by using blockchain as it offers a decentralized system of implementation and incentive. Local sharing compared to on-chain and on-cloud sharing is more effective in maintaining the privacy of original electronic medical data.

One type of local sharing implementation, as in Fig. 5, integrates blockchain and federated learning. The blockchain involves the use of smart contracts to aggregate parameters and store local and global model parameters. A requester of data displays the initial global model on the blockchain to initiate the federated learning process. The model parameters are formed at the data owner by loading this model at the latest block and then trained using his own local data. These parameters are then translated into the global model and moved to the blockchain in the form of transactions and aggregated with smart contracts.

The consensus mechanism of the blockchain has a mining node chosen to create a new block with the new model parameters in the world. The owners of the data read this block during the subsequent round of training and this process continues until the global model converges. This cycle guarantees safe, efficient and decentralised model training and protects data privacy.

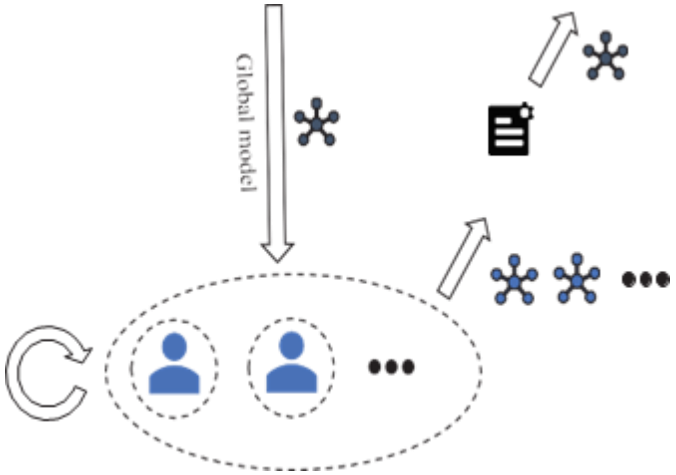

Figure 3: Operation process of the original electronic medical data local sharing methods implemented on-chain

Optimize Security Framework

There are three primary security risks to existing medical blockchain data-sharing approaches: the federated learning, the transmission of data and blockchain. Several security optimisation strategies may be put into consideration to address this. Routine verification of the consensus algorithm and application of external verification methods are significant in preventing blockchain-related threats such as collusion attacks. As an example, secret pseudorandom function seeds were distributed among authorities to prevent N -1 attempts at collusion (Guo et al. [62]). Monitoring behaviour, limiting the access of participants, and preventing collusion attacks can also be achieved by stringent participant authorisation and verification processes.

Privacy protection technologies such as anonymisation, de-identification and differential privacy could be deployed to transmit data safely. Wang et al. [102] described a periodic aggregation method that employed differential privacy to protect shared model parameters when Chen et al. [75] pre processed shared data with K- anonymity. Another important thing is to establish a legal and compliant data use authorisation framework with objectives on data use, data access control process and data use auditing.

To handle vulnerabilities such as poisoning and inference attacks in the local sharing approaches used in federated learning, a balance between data security and sharing efficacy must be achieved. Wang et al. [107] developed a smart contract to evaluate model parameters, identify assaults, and punish rogue nodes, but Rehman et al. [106] developed an intrusion detection system to locate attacks before transmission. It is possible to explore zero-knowledge-proofs along with encryption technology to prevent cheating of data owners who may end up replicating updates on other individuals and hence reduce the efficiency of federated learning. This approach secures the integrity of the federated learning processes and updates of parameters and ensures the privacy.

Conclusions

Medical blockchain to data-share is not a novel concept, yet a variety of different and complex solutions exist, which are not organised clearly. According to where the original medical information is stored, this study classifies these methods into three groups: local sharing, on-chain sharing and on-cloud sharing. It reminds the key research features of every method, explains its functioning, and analyzes the challenges that it faces. Future research directions are also mentioned in the study with a focus on new local sharing approach, which applies federated learning along with the established on-chain and on-cloud approaches. The paper presents valuable research and suggestions towards the advancement of the electronic medical data-sharing research field through the review and assessment of different approaches.

References

1. Uddin, Mueen, M. S. Memon, Irfana Memon, Imtiaz Ali, Jamshed Memon, Maha Abdelhaq, and Raed Alsaqour.,Hyperledger fabric blockchain: Secure and efficient solution for electronic health records., Computers, Materials , (2021), Vol. 68, 2377–2397. 102
2. Aggarwal, Shubhani, and Neeraj Kumar,Hyperledger, Advances in computers, (2021), Vol. 121, 323–343.
3. Cui, Yu, and Hiroki Idota, Improving Supply Chain Resilience with Establishing A Decentralized Information Sharing Mechanism, Proceedings of the 5th Multidisciplinary International Social Networks Conference, (2018), 1–7.
4. Yiu, Neo CK,Article toward blockchain-enabled supply chain anticounterfeiting and traceability., Future Internet, (2021), Vol. 13, 86.
5. Ivanov, D., and P. Pashkov,A blockchain-based approach to providing technically expressed trust in the supply chains of the fashion industry., Journal of Physics: Conference Series, (2021), Vol. 2032, 012086.
6. Rana, Sumit Kumar, Hee-Cheol Kim, Subhendu Kumar Pani, Sanjeev Kumar Rana, Moon-Il Joo, Arun Kumar Rana, and Satyabrata Aich, Blockchain-Based Model to Improve the Performance of the Next-Generation Digital Supply Chain., Sustainability, (2021), Vol. 13, 10008.
7. Azzi, Rita, Rima Kilany Chamoun, and Maria Sokhn,The power of a blockchain-based supply chain. Computers & industrial engineering, (2019), Vol. 135, 582–592.
8. Christidis, Konstantinos, and Michael Devetsikiotis,Blockchains and Smart Contracts for the Internet of Things, Ieee Access, (2016), Vol. 4, 2292–2303.
9. Dinh, Tien Tuan Anh and Wang, Ji and Chen, Gang and Liu, Rui and Ooi, Beng Chin and Tan, Kian-Lee,Framework for Analyzing Private Blockchains, Proceedings of the 2017 ACM international conference on management of data, (2017), 1085–1100.
10. Ramamurthy, S, Leveraging blockchain to improve food supply chain traceability, IBM Blockchain Blog, (2016).

11. Shahid, Affaf and Almogren, Ahmad and Javaid, Nadeem and Al-Zahrani, Fahad Ahmad and Zuair, Mansour and Alam, Masoom,Blockchain-Based Agri-Food Supply Chain: A Complete Solution, IEEE Access, (2020), Vol. 8, 69230–69243.
12. Kamilaris, Andreas, Agusti Fonts, and Francesc ,The rise of blockchain technology in agriculture and food supply chains, Trends in Food Science Technology, (2019), Vol. 91, 640–652.
13. Salah, Khaled and Nizamuddin, Nishara and Jayaraman, Raja and Omar, Mohammad,Blockchain-based soybean traceability in agricultural supply chain, Ieee Access, (2020), Vol. 7, 73295–73305.
14. Saberi, Sara and Kouhizadeh, Mahtab and Sarkis, Joseph and Shen, Lejia,Information asymmetry, blockchain and food safety, Res. China Mark. Superv, (2016), Vol. 11, 53–56.
15. Khanna, Abhirup and Jain, Sapna and Burgio, Alessandro and Bolshev, Vadim and Panchenko,Vladimir,Blockchain-Enabled Supply Chain platform for Indian Dairy Industry: Safety and Traceability, Foods, (2022), Vol. 11, 2716.