A hybrid blockchain-based solution for secure sharing of electronic medical record data

Gang Han [1, 2, 3], **Yan Ma** [Corresp., 3], **Zhongliang Zhang** [1], **Yuxin Wang** [3]

[1] School of Management, Hangzhou Dianzi University, Hangzhou, Zhejiang, China

[2] The State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an, shaanxi, China

[3] The School of Cyberspace Security, Xi'an University of Posts and Telecommunications, Xi'an, Shaanxi, China

Corresponding Author: Yan Ma
Email address: vickiem528@gmail.com

Patient privacy data security is a pivotal area of research within the burgeoning field of smart healthcare. This study proposes an innovative hybrid blockchain-based framework for the secure sharing of electronic medical record (EMR) data. Unlike traditional privacy protection schemes, our approach employs a novel tripartite blockchain architecture that segregates healthcare data across distinct blockchains for patients and healthcare providers while introducing a separate social blockchain to enable privacy-preserving data sharing with authorized external entities. This structure enhances both security and transparency while fostering collaborative efforts across different stakeholders. To address the inherent complexity of managing multiple blockchains, a unique cross-chain signature algorithm is introduced, based on the Boneh-Lynn-Shacham (BLS) signature aggregation technique. This algorithm not only streamlines the signature process across chains but also strengthens system security and optimizes storage efficiency, addressing a key challenge in multi-chain systems. Additionally, our external sharing algorithm resolves the prevalent issue of medical data silos by facilitating better data categorization and enabling selective, secure external sharing through the social blockchain. Security analyses and experimental results demonstrate that the proposed scheme offers superior security, storage optimization, and flexibility compared to existing solutions, making it a robust choice for safeguarding patient data in smart healthcare environments.

A hybrid blockchain-based solution for secure sharing of electronic medical record data

Gang Han[1,2,3], Yan Ma[3], Zhongliang Zhang[1], and Yuxin Wang[3]

[1] School of Management, Hangzhou Dianzi University, Hangzhou, Zhejiang, China
[2] The State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an, Shaanxi, China
[3] The School of Cyberspace Security, Xi'an University of Posts and Telecommunications, Xi'an, Shaanxi, China

Corresponding Author:
Yan Ma [3]
Chang'an District, Xi'an, Shaanxi, China
Email address: vickiem528@gmail.com

Abstract

Patient privacy data security is a pivotal area of research within the burgeoning field of smart healthcare. This study proposes an innovative hybrid blockchain-based framework for the secure sharing of electronic medical record (EMR) data. Unlike traditional privacy protection schemes, our approach employs a novel tripartite blockchain architecture that segregates healthcare data across distinct blockchains for patients and healthcare providers while introducing a separate social blockchain to enable privacy-preserving data sharing with authorized external entities. This structure enhances both security and transparency while fostering collaborative efforts across different stakeholders. To address the inherent complexity of managing multiple blockchains, a unique cross-chain signature algorithm is introduced, based on the Boneh-Lynn-Shacham (BLS) signature aggregation technique. This algorithm not only streamlines the signature process across chains but also strengthens system security and optimizes storage efficiency, addressing a key challenge in multi-chain systems. Additionally, our external sharing algorithm resolves the prevalent issue of medical data silos by facilitating better data categorization and enabling selective, secure external sharing through the social blockchain. Security analyses and experimental results demonstrate that the proposed scheme offers superior security, storage optimization, and flexibility compared to existing solutions, making it a robust choice for safeguarding patient data in smart healthcare environments.

Introduction

Blockchain technology, recognized as a decentralized and secure distributed ledger system, has been significantly adopted in sectors such as finance and supply chains [1-3]. Its decentralized nature ensures that there is no single point of failure, making it resilient against attacks and system breakdowns, while its inherent immutability guarantees that once data is recorded, it cannot be tampered with. These characteristics are particularly valuable in the healthcare domain, where the integrity and trustworthiness of sensitive medical records are paramount. Moreover, blockchain's transparency allows authorized healthcare providers, patients, and other entities to securely verify and access data, which enhances trust in the system while safeguarding patient privacy through cryptographic techniques. Given these benefits, there has been burgeoning interest in applying blockchain technology to healthcare in recent years, where it has been envisioned as a novel means to manage patient records, diagnostic data, prescriptions, and other sensitive information [4-6].

Most previous studies on blockchain-based healthcare systems adopted an on-chain and off-chain storage model to achieve system decentralization and alleviate storage pressure [7-8]. For instance, vast amounts of encrypted data are stored on cloud servers, while single blockchains store the addresses. However, in practical scenarios, this storage model can lead to significant disarray in the storage spaces both on-chain and off-chain, consequently reducing system functionality and increasing the likelihood of system attacks.

Some current solutions improve data accessibility among healthcare providers by modifying access control policies or managing workflows [9-11]. These approaches typically use a single blockchain to store all medical-related data, which enhances system security and logical coherence to some extent but does not fundamentally solve the issue of chaotic storage models. In terms of system security, some methods alter the blockchain consensus mechanism [12-14] to avoid 51% attacks. However, due to the sensitivity and integrity requirements of medical data, more decentralized and verifiable consensus methods, which still carry certain risks of attacks, are necessary [15].

To address these issues, this paper proposes a novel method that introduces a hybrid blockchain-based solution for the secure sharing of electronic medical record (EMR) data. This approach employs three blockchains to identify different functions within the healthcare system and designs a unique cross-chain signature algorithm tailored for this system. This method enhances data security while optimizing storage space. Our approach makes the system resilient to 51% attacks, achieves data fitting, and improves the logical structure of data storage.

The contributions of this paper are as follows:

• A hybrid blockchain-based solution is proposed for secure sharing of EMR data, which rationally distributes healthcare system functions through a multi-chain storage model. This model reduces storage pressure and enhances the system's resistance to 51% attacks.

• A cross-chain signature algorithm is designed to improve data privacy protection, achieve data fitting, and alleviate the chaotic storage space issue in blockchain healthcare systems, making the storage model more organized.

- The introduction of a social chain enhances external data sharing capabilities and implements an efficient data layering strategy.

The remainder of this paper is organized as follows: Section II discusses related work, Section III introduces relevant background knowledge, Section IV presents the proposed system solution, Section V provides experiments and security analysis, and finally, Section VI analyzes the limitations of the solution and concludes the paper.

Related work

Blockchain technology has garnered significant attention in the design of EMR systems to address the challenges of fragmented health data, privacy protection, and secure data sharing. Previous studies typically adopt a hybrid on-chain and off-chain storage model to achieve decentralization and alleviate storage pressure. For instance, a distributed electronic health record (EHR) ecosystem was proposed that integrates EMR into a private and permissioned blockchain. This approach aims to unify fragmented patient records across various healthcare organizations, enhancing data consistency and security [16]. Similarly, Chelladurai et al. proposed a blockchain-based EHR system that offers a regulated solution for patients, physicians, and healthcare providers that addresses data fragmentation issues [17]. Kim et al. introduced a secure and efficient solution for managing EHRs using blockchain for data integrity, access control, and secure health data sharing, combined with cloud computing [18]. Fatokun et al. further expanded on this concept by proposing a patient-centric EHR system on the Ethereum blockchain platform that provides patients with greater control over their data and eliminates the need for third-party systems [19].

Other researchers have focused on enhancing data privacy and system scalability. Shuaib et al. proposed a blockchain-based healthcare data-sharing system that integrates a decentralized file system and a threshold signature to mitigate privacy-linking attacks and scalability challenges [20]. Liu et al. addressed secure storage and sharing of EMRs with a consortium blockchain-based solution that incorporates anonymous and traceable identity privacy protection, dual blockchain and cloud server storage, and an improved proxy re-encryption scheme [21]. In addition, Guo et al. developed a hybrid blockchain-edge architecture employing attribute-based cryptographic mechanisms for managing EHRs. This architecture features an innovative attribute-based signature aggregation (ABSA) scheme, multi-authority attribute-based encryption (MA-ABE), and Paillier homomorphic encryption (HE) for patient anonymity and EHR security [22]. Liu et al. suggested using proxy re-encryption and sequential multi-signature combined with cloud platform services to further protect patient privacy data on the blockchain [23]. Yuan et al. proposed a detailed, secure sharing scheme for medical data leveraging blockchain technology, addressing the issues of low throughput and instability in single-chain models while enhancing data confidentiality [24].

Recent studies have also explored the use of Byzantine consensus mechanisms in blockchain-based healthcare systems. For example, a blockchain-based healthcare platform with Byzantine fault tolerance (BFT) was proposed, ensuring data integrity, confidentiality, and availability, which is crucial for healthcare applications [25]. Another study introduced an efficient and secure health data sharing framework using blockchain with Byzantine consensus, which addressed issues like data tampering and unauthorized access [26]. Additionally, a novel approach for secure EHR using Hyperledger Fabric with BFT was explored that would enhance patient data security and accessibility while maintaining high performance and reliability [27].

As shown in Table 1, while these initiatives illustrate significant progress in integrating blockchain technology into EMR systems, they often rely on single blockchain models or hybrid storage solutions that may lead to chaotic storage management and reduced system functionality. Additionally, current methods for enhancing system security, such as modifying consensus mechanisms, still face challenges in completely mitigating the risks of attacks, especially given the sensitivity of medical data. To address these limitations, our research introduces a novel hybrid blockchain-based solution for the secure sharing of EMR data. By employing three blockchains for different functions within the healthcare system and designing a unique cross-chain signature algorithm, our approach optimizes storage space, enhances data security, and improves system resilience to 51% attacks. This method ensures organized data storage, provides a more robust solution for secure EMR data sharing, and advances the state of blockchain applications in healthcare.

Preliminaries

This section provides a brief review of relevant knowledge.

A. Blockchain-related theory

Blockchain represents a novel application paradigm of computer technology that integrates various cutting-edge technologies including distributed data storage, P2P transmission, consensus mechanism, and encryption algorithms. It serves as a decentralized and trustless infrastructure that operates on a distributed computing paradigm. The theoretical foundations of blockchain-primarily draw upon information asymmetry theory, free currency theory, and BFT theory, while the technical support is provided by P2P network technology, timestamp technology, asymmetric encryption, smart contracts, and database technology [28-30]. Generally, the infrastructure of blockchain is comprised of a data layer, network layer, consensus layer, incentive layer, contract layer, and application layer, as illustrated in Fig 1.

B. Attribute-based encryption

The basic idea of attribute-based encryption (ABE) is to integrate the access control of data into the decryption process of the cipher text, providing a new perspective on the access control of encrypted data [31-32]. The most important feature of this encryption method is that it does not rely on the user's identity information to encrypt and decrypt the data, but on a set of attributes of the user, and only when the user's attributes satisfy the access policy defined in the ciphertext can the user successfully decrypt the original text.

There are two main types of attribute-based encryption techniques, key-policy attribute-based encryption (KP-ABE) and ciphertext-policy attribute-based encryption (CP-ABE) [33]. In KP-ABE, the key is determined by an access structure and the ciphertext is marked by a set of attributes. A user can only decrypt a ciphertext if the access structure of his key matches the set of attributes of the ciphertext. In contrast, in CP-ABE, the access policy is specified by the ciphertext and the user's key is marked by a set of attributes. The user can only decrypt a ciphertext if the set of attributes of the key satisfies the access policy of the ciphertext.

C. BLS signature

Boneh-Lynn-Shacham (BLS) signature is a type of digital signature scheme that offers a short, computationally efficient signature and the ability to aggregate signatures [34]. BLS signature scheme is based on bilinear pairings on elliptic curves, which makes it possible to compress multiple signatures from multiple users into a single signature. This aggregation capability is particularly valuable for systems that need to manage a large number of signatures, like blockchain networks. This algorithm needs a bilinear pairing $\mathbb{G}_0 \times \mathbb{G}_1 \rightarrow \mathbb{G}_T$. The pairing is efficiently computable, non-degenerate, and all three groups have prime order q. This paper let g_0 and g_1 be generators of $\mathbb{G}_0$ and $\mathbb{G}_1$, respectively. It also needs a hash function H_0. The hash function will be treated as a random oracle in the security analysis.

BLS signature aggregation works as follows:

KeyGen (): choose a random $\alpha \xleftarrow{R} \mathbb{Z}_q$ and $h \leftarrow g_1^{\alpha} \in \mathbb{G}_1$. Output $pk:=(h)$ and $sk:=(\alpha)$.

Sign (sk,m): output $\sigma \leftarrow H_0(m)^{\alpha} \in \mathbb{G}_0$. The signature is a single group element.

Verify (pk,m,σ): if $e(g_1,\sigma) = e(pk,H_0(m))$ output "accept", otherwise output "reject".

Hybrid blockchain-based solution for secure sharing of EMRs

A. Notation table

To facilitate understanding of the proposed method, Table 2 summarizes the symbols used throughout this paper.

B. System architecture

The proposed scheme consists of three interconnected blockchains: the patient blockchain, healthcare provider blockchain, and social blockchain. A combination of on-chain and off-chain structures is employed to store medical records, providing the necessary flexibility and scalability to securely store and manage large volumes of sensitive healthcare information. The Blockchain-based healthcare architecture is shown in Fig 2. The data sharing process of the system is illustrated in Fig 3, which demonstrates the data flow between blockchains, the signature algorithm, and other steps.

(1) Patient blockchain

The patient blockchain stores the signature signed by the patient, along with hashed data that includes encrypted personal information $C_{pk}(PI_i)$. This information encompasses sensitive patient data such as name, age, gender, and other privacy-related information. To ensure the security and privacy of the data, it is uploaded to IPFS by the patient and then hashed by IPFS. Given the smaller size of personal data compared to medical data, CP-ABE is employed to encrypt personal data, allowing for a higher degree of privacy protection. Personal information and EHR are stored separately in distinct blockchains to enhance the security of EHR. This approach prevents adversaries from associating medical data with specific patients, thereby reducing the risk of data breaches and safeguarding patient privacy.

Patients and healthcare providers can access data from the patient blockchain and retrieve ciphertext from IPFS. This ciphertext can then be decrypted to reveal the actual personal information of the patient. In typical scenarios, healthcare providers require access to medical data from a healthcare provider blockchain that is connected to a patient's personal information, which can then be used to make a diagnosis.

(2) Healthcare provider blockchain

The healthcare provider blockchain primarily stores aggregate signatures S_A, which are composed of signatures generated by various institutions, along with dataset D_i, where $T_i = D_i || S_A$. The dataset is comprised of the medical data ciphertext $C_i = E_k(M_i)$, its hash value $H(C_i)$, and the symmetric key k encrypted with the patient's public key for encrypting medical data $E_{PK_u}(k)$. Specifically, D_i can be expressed as $D_i = C_i || H(C_i) || E_{PK_u}(k)$. This approach ensures that the medical data and associated signatures are securely stored, while also maintaining the privacy and confidentiality of patient information.

Similarly to the patient blockchain, patients and healthcare providers can access data from the healthcare provider blockchain and retrieve ciphertext from IPFS. Given the critical nature of medical data during physician diagnostic and data access procedures, additional security measures are necessary to enhance the protection of sensitive medical data.

(3) Social blockchain

The social blockchain serves as a crucial component in our scheme, facilitating connections to external blockchain networks. Transactions involving data sharing with other systems are uploaded to this blockchain. Each transaction includes the hashed ciphertext that has already been shared with others and the signature of the data's owner. To enable better differentiation of which parts of a patient's EHR are shared, a data processor is used to classify the medical records, dividing the data into finer-grained categories. When a patient transfers to another hospital, the social blockchain connects to the external blockchain network system, and the relevant patient data is transferred accordingly. The social blockchain does not require direct interaction with patients and healthcare providers. Our data classification scheme provides an effective solution for transferring different types of data, thereby reducing the workload for users.

C. Cross-chain signature algorithm

This paper proposes the cross-chain signature algorithm, which facilitates the execution of signature protocols among users on different chains, addressing two practical issues. First, it enhances the privacy of medical data by leveraging the immutability and decentralization of blockchain technology. Second, it enables the fitting of heterogeneous data in a distributed storage system within the blockchain. Our signature algorithm ultimately produces two types of aggregate signatures S_{A_s}, which is used to enhance data privacy, and S_{A_f}, which is used to achieve data fitting. The algorithm involves two main categories of users: patients in the patient blockchain and medical service providers in the medical service provider blockchain.

Regarding the aggregate signature S_{A_s}, it is assumed that when patients generate data, they transmit the ciphertext of the data along with other relevant information to all medical institutions. Subsequently, each user signs the ciphertext data. After obtaining the individual signatures from each user, an aggregate operation is performed to compute the aggregate signature S_{A_s}. Similarly, when new medical data is generated within medical institutions, the ciphertext is shared with other users, and the same signing and aggregation process is executed. For the aggregate signatureS_{A_f}, users participating in the system continuously generate new data. Any user is required to compute the aggregate signature S_{A_f} of all signatures S_i generated before time T_i. Finally, at time T_i, users upload both (S_{A_s},S_{A_f}) to the new block.

Since each user has access to the same data ciphertext and signatures, attackers cannot identify the true source of the data, thereby preventing targeted attacks and enhancing the privacy of medical data. Additionally, the presence of the aggregate signature S_{A_f} allows for the identification of all data signatures generated by a particular user, thereby achieving the fitting of heterogeneous data.

(1) Aggregate signature S_{A_s}

The process of obtaining the signature S_{A_s} is illustrated in Fig 4. As an example of patient-generated data, the details of the signature process are as follows:

Our scheme needs a bilinear pairing e: $\mathbb{G}_0 \times \mathbb{G}_1 \rightarrow \mathbb{G}_T$, the hash function H_0: $\mu \rightarrow \mathbb{G}_0$, and a second hash function H_1: $\mathbb{G}_1^n \rightarrow R^n$ where $R := \{1,2,\ldots,2^{128}\}$.

a) KeyGen ( )

The system assigns the public key PK_1 and secret key SK_1 to medical staff for signing.

b) Prepare data D_1

The patient's personal data, denoted as M_1, is self-generated by the patient, followed by encryption to derive the ciphertext $C_1 = E_k(M_1)$. Subsequently, C_1 is subjected to a hashing process to yield the hashed data, denoted as $H(C_1)$. The culmination of this process results in the final prepared data $D_1 = C_1 || H(C_1)$.

c) Sign (D_1,SK_1)

This algorithm takes the prepared data D_1 and the signing key SK_1 as inputs. Eventually, it returns the signature S_1 as a result.

$$S_1 \leftarrow H_0(D_1)^{SK_1} \in \mathbb{G}_0$$

d) Share data D_1 and sign

The patient shares the data D_1 with other healthcare providers, thus each provider has the same data D_1 and uses their signing key SK_i to output different signatures S_i.

$$S_i \leftarrow H_0(D_1)^{SK_i} \in \mathbb{G}_0$$

e) Signature aggregate $((PK_1,S_1),(PK_2,S_2),\ldots,(PK_n,S_n))$

This algorithm takes all the individual signatures related to different users, then computes t_1, $t_2,\ldots,t_n$ and outputs the aggregation signature S_A.

$$(t_1,t_2,\ldots,t_n) \leftarrow H_1(PK_1,PK_2,\ldots,PK_n) \in R^n$$

$$S_A \leftarrow S_1^{t_1} \ldots S_n^{t_n} \in \mathbb{G}_0$$

f) Public key aggregate

This process involves advanced preparation for verifying the signature. The algorithm incorporates all the relevant individual public keys associated with different healthcare providers, then computes $t_1,t_2,\ldots,t_n$ and outputs the aggregation of the public key PK_A.

$$(t_1,t_2,\ldots,t_n) \leftarrow H_1(PK_1,PK_2,\ldots,PK_n) \in R^n$$

$$PK_A \leftarrow PK_1^{t_1},PK_2^{t_2},\ldots,PK_n^{t_n} \in \mathbb{G}_1$$

g) Verify $(H(D_1),PK_A,S_A)$

This algorithm takes the hashed data $H(D_1)$, the aggregation of the public key PK_A, and the aggregation signature S_A to verify if $e(g_1,S_A) = e(PK_A,H_0(H(D_1)))$ the output "accepts", or otherwise output "rejects".

The verification process is proven as below:

$$
\begin{aligned}
& e\left(g_1, S_A\right) \\
&= e\left(g_1, S_1^{t_1} \ldots S_n^{t_n}\right) \\
&= e\left(g_1, S_1^{t_1}\right) \cdot e\left(g_1, S_2^{t_2}\right) \cdot \ldots \cdot e\left(g_1, S_n^{t_n}\right) \\
&= e\left(g_1^{t_1}, S_1\right) \cdot e\left(g_1^{t_2}, S_2\right) \cdot \ldots \cdot e\left(g_1^{t_n}, S_n\right) \\
&= e\left(g_1^{t_1}, H_0(D_i)^{\alpha_1}\right) \cdot e\left(g_1^{t_2}, H_0(D_i)^{\alpha_2}\right) \cdot \ldots \cdot e\left(g_1^{t_n}, H_0(D_i)^{\alpha_n}\right) \\
&= e\left(\left(g_1^{t_1}\right)^{\alpha_1}, H_0(D_i)\right) \cdot e\left(\left(g_1^{t_1}\right)^{\alpha_2}, H_0(D_i)\right) \cdot \ldots \cdot e\left(\left(g_1^{t_1}\right)^{\alpha_n}, H_0(D_i)\right) \\
&= e\left(\left(g_1^{\alpha_1}\right)^{t_1}, H_0(D_i)\right) \cdot e\left(\left(g_1^{\alpha_2}\right)^{t_1}, H_0(D_i)\right) \cdot \ldots \cdot e\left(\left(g_1^{\alpha_n}\right)^{t_1}, H_0(D_i)\right) \\
&= e\left(PK_1^{t_1}, H_0(D_i)\right) \cdot e\left(PK_2^{t_2}, H_0(D_i)\right) \cdot \ldots \cdot e\left(PK_n^{t_n}, H_0(D_i)\right) \\
&= e\left(PK_1^{t_1} \ldots PK_n^{t_n}, H_0(D_i)\right) \\
&= e\left(PK_A, H_0(D_i)\right)
\end{aligned}
$$

(2) Aggregate signature S_{A_f}

The process of obtaining the signature S_{A_f} is illustrated in Fig 5. This process is generally similar to the one described above, with the signing data being the privacy data generated by User 1 at different times. At time T_n, all signatures of User 1 are $S_{1T_1}, S_{1T_2}, \ldots, S_{1T_n}$. The aggregate signature $S_{A_{f_1}T_n}$ is then computed as $S_{A_{f_1}T_n} \leftarrow S_{1T_1}^{\ t_1} \ldots\ldots S_{1T_1}^{\ t_n}$.

The property of aggregate signatures, which allows for the verification of individual signature existence, is utilized to create an invisible chain formed by the aggregate signatures. This enables the identification and categorization of all data uploaded by a particular user from the mixed data, achieving the fitting of a specific type of data. This not only organizes the system's data more effectively but also enhances the efficiency and accuracy of data management.

Suppose User 1 updates data at time T_n and generates the aggregate signature $S_{A_{f_1}T_n}$. This data and signature are then uploaded to the off-chain IPFS distributed storage system. Additionally, at times $T_{n-1}, T_{n-2}, \ldots, T_1$, the system stores aggregate signatures of a large amount of user data, denoted as $S_{A_{f_i}T_j}$. All these signature collections are referred to as $\mathcal{H}$. For $\forall S_{A_{f_i}T_j} \in \mathcal{H}$, the system can search and verify all data generated by User 1 at times $T_j < T_n$ within the aggregate signatures.

The existence of $S_{A_{f_1}T_j}$ can be confirmed through effective aggregation. The verification formula is $\mathrm{e}\left(S_{A_{f_1}T_n},g_1\right) = e\left(H(D),PK_A\right) \cdot e(S_{A_{f_1}T_{n-1}},g_1)^{-1}$.

a) The aggregate signature $S_{A_{f_1}T_n}$ and the aggregate public key $S_{A_{f_1}T_n}$ are considered. Suppose $S_{A_{f_1}T_{n-1}}$ is a part of the aggregate signature $S_{A_{f_1}T_n}$.

$$\mathrm{e}\left(S_{A_{f_1}T_n},g_1\right)$$

$$= e\left(S_{1T_1} \cdot S_{1T_2} \cdot \ldots\ldots \cdot S_{1T_n},g_1\right)$$

$$= e\left(S_{A_{f_1}T_{n-1}} \cdot S_{1T_n},g_1\right)$$

b) Next, the following computation is performed:

$$\mathrm{e}\left(H(D),PK_A\right) \cdot e\left(S_{A_{f_1}T_{n-1}},g_1\right)^{-1}$$

$$= e\left(H(D),PK_1^n\right) \cdot e\left(S_{A_{f_1}T_{n-1}},g_1\right)^{-1}$$

$$= e\left(H(D),PK_1 \cdot PK_1 \cdot \ldots\ldots \cdot PK_1\right) \cdot e\left(S_{A_{f_1}T_{n-1}},g_1\right)^{-1}$$

$$= e\left(H(D),g_1\right)^{SK \cdot n} \cdot e\left(SK \cdot \mathrm{H}(D),g_1\right)^{-1}$$

$$= e\left(H(D),g_1\right)^{SK \cdot n} \cdot e\left(SK \cdot \mathrm{H}(D),g_1\right)^{-SK}$$

$$= e\left(H(D),g_1\right)^{SK \cdot (n-1)}$$

Here, $e\left(S_{A_{f_1}T_{n-1}},g_1\right)^{-1}$ is used to "cancel out" the contribution of $S_{A_{f_1}T_{n-1}}$ in the aggregate signature. If the equation holds, then it can be concluded that the signature $S_{A_{f_1}T_{n-1}}$ is indeed a part of the aggregate signature $S_{A_{f_1}T_n}$.

c) Finally, the expression $\mathrm{e}\left(S_{A_{f_1}T_n},g_1\right) = e(H(D),g_1)^{SK \cdot (n-1)}$ is obtained, which is equal to the left-hand side of the equation, indicating that the signature $S_{A_{f_1}T_{n-1}}$ is a part of the aggregate signature $S_{A_{f_1}T_n}$.

D. System operation details

(1) Patient blockchain: personal information addition

When a patient is initially registered in the system, they are required to provide basic personal information. To ensure privacy, the CP-ABE encryption method is utilized to encrypt the patient's private data. The encrypted data is uploaded into IPFS, where a hash value is generated and then transferred to the patient blockchain along with the signature S_A. Further details regarding this process are outlined below:

a) Setup $(\lambda) \rightarrow (pk, mk)$

This procedure takes the security parameter λ as input and produces the public parameter p and master key mk for the proposed CP-ABE mechanism.

b) KeyGen $(mk, A) \rightarrow sk$

This procedure takes the master key mk and the attribute set A as input and generates the user attribute secret key sk.

c) Encrypt $(p, T, M) \rightarrow C$

This procedure takes the public parameter p , accesses the structure T and the patient's personal information M, and encrypts the plaintext M into the ciphertext C.

d) Sign $(C) \rightarrow S_A$

In this process, the user generates shareable data D_i using the signature algorithm described above and signs D_i to obtain the corresponding signed data S_A.

e) Data upload to the blockchain

The signature S_A and the $hash$ value returned by IPFS are incorporated into a data transaction $T_i = D_i||S_A||hash$ and subsequently uploaded to the patient blockchain.

f) Decrypt $(p, C, sk) \rightarrow M$

This procedure takes the public parameter p, ciphertext C, and secret key sk. as input and generates the plaintext M.

(2) Healthcare provider blockchain: health record addition

Healthcare data is decentralized among a variety of healthcare providers, and encompasses entities such as hospitals, insurers, pharmacies, and governmental regulatory bodies. Different from the ciphertext present in the patient blockchain, this medical data is considerably more substantial in volume. Initially, the medical data is encrypted using symmetric encryption. Subsequently, the symmetric encryption key itself is encrypted via CP-ABE. The final encrypted data can be procured by concatenating these two ciphertexts. Further details regarding this process are outlined below:

a) Key generation

When a user affiliated with a healthcare provider partakes in the system, a symmetric encryption key k is allocated by the system. The key assignment for CP-ABE is the same as that in the patient blockchain and is not described in detail here.

b) Encrypt $(k, p, T, M) \rightarrow C$

Within this process, the healthcare provider generates the medical data M, which is encrypted using the key k, resulting in $C_s = Enc_k(M)$. Following this, attribute encryption is performed on the key k $Encrypt(P, T, k) \rightarrow C_a$. $(C_s||C_a)$ which is the final ciphertext data C.

c) Sign $(C) \rightarrow S_A$

In this process, the user generates shareable data D_i using the signature algorithm described above and signs D_i to obtain the corresponding signed data S_A.

d) Data upload to the blockchain

The signature S_A and the $hash$ value returned by IPFS are incorporated into a data transaction $T_i = D_i||S_A||hash$ and subsequently uploaded to the healthcare provider blockchain.

e) Decrypt $(p,C,sk,k) \rightarrow M$

This procedure takes the public parameter p, the ciphertext C_a, and the secret key sk. as input and generates the symmetric encryption key k. The key k is subsequently employed to decrypt the ciphertext C_s, facilitating the retrieval of the original data M.

(3) Social blockchain: data sharing externally

The social blockchain establishes a connection with the external blockchain system to facilitate the sharing of medical data among different healthcare institutions. To ensure proper data sharing, a data processor categorizes the medical data into five levels of sensitivity. Only the data that meets the sharing criteria are allowed to pass through the social blockchain for further dissemination among authorized entities within the network. The five levels of sensitivity are:

Level 1: Fully public data. This category includes information such as the hospital's name, address, and telephone number, which can be openly shared with the public on the Internet without any restrictions or privacy concerns.

Level 2: Data available for widespread access. This category is comprised of data that can be accessed on a large scale, typically after obtaining approval through a formal application process. These datasets are often made available for research and analysis purposes, facilitating scientific investigations and advancements in various domains.

Level 3: Data available for restricted access. This category includes data that can be accessed on a medium scale, typically limited to usage within the authorized project team operating under the purview of a specific institution. Access to this data is granted only to team members involved in the project, ensuring compliance with institutional policies and safeguarding the privacy and security of the data.

Level 4: Data available for limited access. This category pertains to data that can be accessed on a smaller scale, specifically restricted to individuals directly involved in the consultation process. Access to this data is confined to healthcare professionals and relevant stakeholders who require access to providing healthcare services and facilitating the consultation. Strict confidentiality measures are implemented to protect the privacy and sensitivity of the data.

Level 5: Data available for highly restricted access. This category encompasses data that can only be accessed on a very limited scale and under stringent restrictions. For instance, specific disease-related information, such as on AIDS or STDs, is strictly limited to access by primary care providers who require the data for clinical purposes. Comprehensive controls and protocols are in place to ensure the utmost confidentiality and privacy protection of this sensitive information, adhering to regulatory guidelines and ethical considerations.

The data-sharing process is shown in Fig 6. When an external user requires access to medical data, the system employs an automated process to determine the appropriate levels of data accessibility based on the user's attributes. Upon identification, the user can retrieve the corresponding hash from the social blockchain, which serves as a reference for obtaining the authorized data from the IPFS storage system. This dynamic approach ensures that users can securely retrieve and access the specific data they are permitted to view, maintaining the privacy and confidentiality of the overall healthcare ecosystem.

Two algorithms have been conceptualized, specifically tailored for the tasks of automatic data hierarchy creation and data dissemination. Algorithm 1 illustrates the mechanism of data categorization. This mechanism involves a detailed stratification of data according to sensitivity levels, thereby differentiating information that is permissible for open distribution from datasets that demand heightened sensitivity.

Algorithm 1: Medical data (MD) categorization based on sensitivity level

Input: MD
Output: Categorized data (CD) with sensitivity level (SL)

```
# Stage I: Determining Sensitivity Level
1: Determine the SL of RMD based on predefined criteria → SL

# Stage II: Categorizing data
2: CD = { 'data': MD, 'sensitivity': SL }

# Stage III: Returning categorized data
3: Return CD
```

Algorithm 2 elaborates the process of medical information distribution leveraging social blockchain technology. This process necessitates an evaluation of user permissions, only those users satisfying the specified criteria are granted data access. Furthermore, the data distribution procedure is inscribed in the framework of the social blockchain infrastructure, thereby establishing unequivocal transparency and traceability.

Algorithm 2: Medical data sharing via social blockchain

Input: Categorized Data CD, User User
Output: Hash value (HV) or none

```
# Stage I: Checking user permission
1: Check user permission for CD ['sensitivity'] → Permission
2: If permission is false → Return none

# Stage II: Retrieving HV from IPFS and uploading to social blockchain
3: Retrieve HV from IPFS for CD['data']  → HV
4: Upload HV to social blockchain → Record
5: Return HV
```

Experiments and analysis

A. Safety analysis

(1) Data privacy protection

This healthcare blockchain scheme employs a robust system of encryption and signature to safeguard both patient personal data and medical data, ensuring privacy and data integrity.

In the case of patient's personal data, it undergoes an encryption process based on specific attributes during the uplink phase. The decryption of this data is solely possible for users possessing the corresponding attributes $\mathbb{A}$. Unauthorized attackers lacking these decryption attributes fail to generate a valid key sk, and hence, cannot access the patient's information. This mechanism provides an essential layer of privacy protection for personal data.

The system handles medical data D_i emanating from different medical institutions in a unique way to ensure its security. During the signature process, each medical institution encrypts the updated medical data and distributes the resulting ciphertext to the other participating institutions. Each of these institutions subsequently signs the received ciphertext independently. Finally, the aggregated signature along with other associated data is uploaded to the blockchain.

Given that every medical institution holds the ciphertext and the signature of the data, and only the originating institution knows the actual plaintext, an attacker, even upon acquiring the ciphertext data and signature, cannot discern the real data generator. Further, without the necessary decryption tools, they cannot decrypt the data. This system effectively serves to protect the confidentiality and privacy of medical data.

(2) Anti-counterfeiting attacks

The BLS signature aggregation algorithm provides robust security measures, making it an ideal tool for user identity protection in blockchain structures utilized by healthcare organizations. This cryptographic method involves the creation of a pair of public and private keys (PK_i,SK_i). The private key SK_i is generated randomly, and the public key PK_i is a calculated product of this private key and a generator point g_1 on an elliptic curve, specifically, $PK_i = g_1^{SK_i}$. To establish a signature S_i, the user employs the private key SK_i to perform operations on the hash of the message, effectively creating the signature S_i. The verification process necessitates three inputs: the public key PK_i, the original message, and the produced signature S_i. The verifier employs the same hash function to create the hash of the message and subsequently utilizes the public key PK_i to authenticate the signature.

This procedure is fortified against forgery attempts by adversaries, chiefly due to the inherent mathematical complexity of the discrete logarithm problem. This problem renders it computationally unfeasible for attackers to deduce the private key SK_i from the public key PK_i. Thus, unless the attacker gains access to the private key SK_i, they cannot forge a valid signature. The degree of difficulty in obtaining the private key SK_i is safeguarded by the intrinsic complexity of the discrete logarithm problem.

(3) Man-in-the-middle attacks

In the proposed scheme each user is allocated a pair of public and private keys (PK_i,SK_i). The user utilizes the private key SK_i to digitally sign the message, followed by the transmission of this signed message coupled with the public key PK_i. The recipient, on their end, uses the sender's public key PK_i to authenticate the signature, ensuring the authenticity and integrity of the received message. An adversary would need access to the sender's private key SK_i to forge a legitimate signature, a highly improbable event given the stringent security practices around private key management.

Consequently, the feasibility of executing a man-in-the-middle attack becomes practically negligible unless an adversary successfully gains access to the sender's private key SK_i, which should ideally be concealed and securely stored. Furthermore, even if an adversary manages to intercept the communication and manipulate the message, their lack of the accurate private key SK_i will lead to signature verification failure at the recipient's end, revealing the attempted tampering.

(4) Resistance to replay attacks

In this paper, each transaction documented in the tripartite blockchain structure possesses a distinct identifier n, which is essentially a single-use numeric value, and is regenerated for every new transaction. This strategy maintains the singularity of each transaction within the system. Consequently, in instances where an adversary attempts to replay an already processed transaction n', the intrinsic checks within the system swiftly identify $n' = n$. As the system has recorded this transaction already, it instantly rejects the replayed transaction. This automated verification and rejection mechanism fortifies the system's resilience against replay attacks, thereby enhancing the comprehensive security framework of the blockchain.

Moreover, each block in the blockchain includes a timestamp T, which denotes the exact instance of block creation. This timestamp instills a chronological order within the blockchain, facilitating the tracking and verification of the transactional sequence. Thus, any replayed transaction with a timestamp $T' \neq T$ would be instantaneously flagged as having been transmitted at an incorrect time, leading to its rejection.

(5) Attack surface analysis

To systematically analyze the potential attacks and demonstrate our system's defenses, Table 3 provides a comprehensive overview of different attack types, their descriptions, and the corresponding defense mechanisms implemented in our system. Each attack type is associated with a specific defense mechanism designed to counteract the threat effectively.

B. Performance comparison

Based on the works referenced in [16-27], the proposed scheme is compared with those of Kim et al. [18] and Liu et al. [21]. The focus of the comparison is on analyzing the total computation time, complexity, and storage space required during the processes of authentication, encryption, and signing of medical data before upload, as shown in Table 4.

Assume that T_{ecenc} denotes the encryption time in the elliptic curve cryptography system, T_h represents the computation time of the hash function, $T_{ReKeyGen}$ refers to the re-encryption key generation time, T_{exp} is the time for an exponentiation operation, T_{bp} is the time for a bilinear pairing operation, T_S is the time for generating a signature, and n represents the number of attributes involved in the encryption process.

For storage space, let M_{meta} epresent the storage size of metadata, M_{enc} represent the storage size of the encrypted data, and M_S represent the storage size of the aggregated signature.

Table 4 provides a performance comparison of the different schemes in terms of data generation time, complexity, and storage space. The results show that the proposed scheme employs more efficient encryption and signature techniques, ensuring that the data generation time remains within a reasonable range while avoiding a significant increase in complexity. Most importantly, by utilizing an aggregated signature, the storage space required is considerably reduced, which further enhances storage efficiency.

C. Block security comparison

(1) Comparison with existing systems

Attacks on the blockchain are mainly based on the important concept of computing power. When an attacker has enough computing power, that is, when the attacker's computing power exceeds 50% of the blockchain consensus network, most blockchain systems can be compromised. However, in practical application scenarios, blockchain systems run with a certain amount of time accumulation, and it is almost impossible for attackers to achieve successful attacks through "51% computing power attacks." Therefore, this paper analyzes the security of blockchain based on the "gambling probability" method of attack.

This paper assumes that a malicious attack node (MAN) successfully joins the blockchain EMR system proposed in this paper and launches an attack on any one of the chains, causing a consensus node on that chain to fork. According to the Bitcoin white paper, blockchain nodes always follow the longest chain as the correct blockchain and extend it. Therefore, in order for MAN to succeed in the attack, it must produce a longer chain to replace the honest nodes' chain. This attack process can be viewed as a binomial random walk. The specific discussion of this process is as follows:

$$\begin{cases} 1 & ,k > z \\ \left(\frac{q}{p}\right)^{z-k} & ,k \leq z \end{cases}$$

As k can be any non-negative integer, the probability distribution of k follows a Poisson distribution, which is calculated as follows:

$$\sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!}$$

Therefore, the formula for calculating the probability of MAN successfully attacking the chain is as follows:

$$P = \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} \cdot \begin{cases} 1 & ,k > z \\ \left(\frac{q}{p}\right)^{z-k} & ,k \leq z \end{cases}$$

The formula after simplification is：

MAN uses a "gambling attack" to attack the EMR system in order to obtain all EMR data. The probability of a successful attack depends mainly on the probability of MAN generating the next block and the number of "chains" used in the blockchain system. This solution combines smart healthcare with a three-chain structure, while the solutions presented in references [20] and [22] are single-chain structures, and the solutions presented in references [21] and [23] are both double-chain structures.

Based on the above analysis, the specific experimental plan of this paper is to have MAN generate the next block with probabilities of 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, and 0.45 respectively. These probabilities were selected to represent a wide range of potential attack intensities, reflecting both realistic scenarios where attackers have limited computing power and more severe cases.

The selection of probabilities from 0.10 to 0.45 was based on the following considerations:

a) Representativeness: These probabilities cover a range from low to high attack success rates, fully reflecting the system security under different levels of attack intensity.

b) Practical application: In actual blockchain applications, attackers usually find it challenging to obtain extremely high computing power. Therefore, selecting probabilities that increment from low values is more in line with real-world scenarios.

c) Scientific basis: The commonly used blockchain attack models in the literature also adopt similar probability ranges, ensuring the comparability and scientific validity of our experiments.

The minimum secure block number that HN needs to generate is then calculated and counted for the EMR system of the blockchain to ensure the attack success probability is below 0.1, 0.01, and 0.001 respectively.

In the experimental design, several potential confounding variables were controlled to ensure the reliability of the results:

a) Node type and number: The type and number of nodes used in all experiments were kept consistent.

b) System load: The system load was controlled throughout the experiments to maintain uniform starting conditions.

c) Environmental settings: All experiments were conducted in the same hardware and software environment to avoid discrepancies due to equipment differences.

d) Repetition of experiments: Each set of experiments was repeated multiple times, and the average value was taken to reduce random errors.

To ensure the statistical significance of our results, a power analysis was conducted using MATLAB. The significance level was set at 0.05, with a target power of 0.80, indicating an 80% chance of detecting a true effect. Based on previous studies, a medium effect size was assumed. This analysis determined the minimum sample size needed to detect meaningful effects at different attack probabilities, ensuring the robustness of our findings.

The experimental tools and simulation environment used in this paper were a personal computer with an Intel(R) Core(TM) i5-1035G1 CPU @ 1.00GHz 1.19 GHz processor, 8.00GB memory, and a 64-bit Windows operating system. The experimental results are shown in Fig. 7, 8, and 9.

a) Probability of MAN attack below 0.1

Experimental results presented in Fig. 7 demonstrate that to maintain the probability (P) of a successful MAN attack on the blockchain system below 0.1, this scheme's increase in the minimum number of secure blocks required is smaller than that of other comparison schemes once the MAN's block generation probability reaches 0.2. Furthermore, elevating the MAN's block generation probability to 0.45 leads to a substantial enhancement in security. Specifically, the required minimum number of secure blocks is lowered by 73.5% in comparison to the single-chain scheme, and by 44.4% relative to the double-chain scheme.

b) Probability of MAN attack below 0.01

Experimental results depicted in Fig. 8 indicate that, to maintain the probability (P) of a successful MAN attack on the blockchain system below 0.01, increasing the MAN's block generation probability to 0.45 significantly enhances security. Specifically, this adjustment reduces the minimum number of secure blocks required by 71.3% compared to the single-chain scheme and by 38.2% when compared to the double-chain scheme.

c) Probability of MAN attack below 0.001

Experimental findings presented in Fig. 9 reveal that, to keep the probability (P) of a MAN successfully attacking the blockchain system under 0.001, the increase in the minimum number of secure blocks required at the genesis block by this scheme is less than that required by comparison schemes. Further, when the probability of MAN generating the next block is raised to 0.45, there is a significant improvement in security. Specifically, the minimum number of secure blocks needed by this scheme is reduced by 70.0% compared to the single-chain scheme, and by 36.2% in comparison to the double-chain scheme.

(2) Comparison with fortified chain systems

To highlight the novelty of our proposed model, a comparison was made with existing fortified chain-based EMR sharing systems [35-36]. Fortified chains enhance the security of blockchain systems through mechanisms such as enhanced consensus protocols and additional security layers. However, these systems often introduce significant redundancies, leading to increased complexity and resource consumption. In contrast, our proposed three-chain structure provides superior performance in terms of security and efficiency by eliminating unnecessary redundancies.

Table 5 summarizes the comparison of security and efficiency metrics between our proposed three-chain system and fortified chain systems. The key metrics include the required minimum secure blocks, throughput, latency, and redundancy.

Fortified chain systems typically incorporate multiple layers of security protocols to enhance resilience against attacks, which can lead to excessive redundancy. This redundancy not only increases computational overhead but also complicates the system architecture. Our proposed three-chain system addresses these issues by optimizing the blockchain structure and consensus mechanisms, thereby maintaining high security without unnecessary redundancy.

The following points illustrate how our system reduces redundancy:

a) Optimized consensus mechanism: By utilizing an efficient consensus algorithm, our system reduces the need for multiple security layers, streamlining the block validation process.

b) Three-chain structure: The separation of data into three distinct chains (Patient, Healthcare Provider, and Social) allows for targeted security measures, reducing the overall computational load and improving performance.

c) Focused security: Instead of applying broad, redundant security protocols, our system implements focused security measures that address specific threats, ensuring robust protection with minimal overhead.

In summary, while fortified chain-based systems have their own strengths, particularly in enhancing security through additional layers, our proposed three-chain structure offers a distinct advantage by balancing high security with greater efficiency. This makes our approach particularly practical and valuable for real-world applications in smart healthcare, where both performance and security are critical.

D. System performance analysis

To further validate the scheme, the Faker library was used to generate random medical data, and smart contracts were deployed on a local Ethereum network using Solidity. Transaction latency, throughput, and gas consumption were assessed for the proposed model.

A sample size of 1,000 transactions was selected to evaluate the performance of the system. This choice was based on the following considerations:

(1) Representativeness: A sample size of 1,000 transactions is large enough to capture the variability in transaction latency and throughput under typical operating conditions, providing a comprehensive assessment of the system's performance.

(2) Practical application: In real-world blockchain applications, it is common to handle a large number of transactions. Testing with 1,000 transactions ensures that the evaluation reflects realistic usage scenarios and can provide insights into how the system performs under substantial load.

(3) Scientific basis: Prior studies and benchmarks in blockchain performance testing often utilize similar or smaller sample sizes to evaluate system performance metrics, ensuring the comparability and scientific validity of our experiments.

By varying transaction send rates, the average latency per transaction was tested on the Patient blockchain, Healthcare provider blockchain, and Social blockchain. As shown in Fig. 10, latency increases with the number of transactions, remaining under 6 seconds at 50 transactions per second (tps). The Social blockchain exhibits slightly lower latency compared to the Patient and Healthcare provider blockchains, attributed to the pre-classification of data, allowing only low-sensitivity data to be shared.

Furthermore, the maximum and minimum throughput of the Patient blockchain, Healthcare provider blockchain, Social blockchain, and the entire system were evaluated at transaction send rates ranging from 10 tps to 40 tps. As shown in Fig. 11, the throughput across different transaction send rates concentrates around 35 tps.

In the performance testing, several potential confounding variables were controlled to ensure the validity of the results:

(1) Data generation: The Faker library was used to generate random but consistent medical data across different experiments.

(2) Smart contract deployment: All smart contracts were deployed using the same configuration and version of Solidity on the same local Ethereum network to maintain consistency.

(3) System environment: The tests were conducted on the same hardware and software environment used in the block security comparison tests, ensuring uniformity.

(4) Repetition and averaging: Each performance test was repeated multiple times, and the results were averaged to minimize random variations.

Our findings indicate that while the use of blockchain introduces some overhead, it does not significantly degrade the overall system performance. Specifically, the latency and throughput metrics remain within acceptable ranges, demonstrating that the integration of blockchain is feasible without compromising efficiency. More importantly, the blockchain-based system significantly enhances data security and integrity, offering robust protection against tampering and unauthorized access.

These results underscore the practicality and feasibility of using blockchain in our system. The enhanced security and data integrity provided by blockchain outweigh the modest increase in complexity and resource consumption, making it a valuable addition to our EMR system.

Discussion

Our findings highlight the effectiveness of a hybrid blockchain-based solution for secure EMR data sharing, optimizing storage, enhancing security, and improving resilience against 51% attacks.

(1) Comparison with existing literature:

Previous studies, such as those by Chelladurai et al. [17], Kim et al. [18], and Liu et al. [21], have focused on hybrid storage solutions and improving data privacy in blockchain-based EMR systems. While these solutions have achieved decentralization and scalability, they often face challenges related to disorganized storage management and limitations in security mechanisms. Additionally, even enhanced consensus protocols, such as those utilizing Byzantine fault tolerance (BFT) [25-27], struggle to fully protect against advanced attack vectors, especially in environments dealing with sensitive healthcare data. The proposed tripartite blockchain model, with its cross-chain signature algorithm, addresses these issues by offering organized data storage and a more robust defense against attacks.

(2) Significance to the research area:

Our research advances the field by providing a more secure and efficient method for EMR data sharing. This hybrid blockchain model enhances interoperability between healthcare providers, improving patient care and reducing administrative burdens. It sets a new benchmark for future studies on blockchain-based healthcare systems.

(3) Broader implications and future research:

The principles of our hybrid blockchain approach can be applied to other fields requiring secure data sharing, such as finance and supply chain management. Future research could explore the scalability of our method in larger datasets and real-world implementations, and develop more advanced consensus algorithms and cross-chain protocols.

Limitations

Despite the theoretical promise of the proposed solution, several practical limitations must be acknowledged. Implementing and maintaining the tripartite blockchain architecture and social blockchain in real-world scenarios is challenging, especially given the current performance and reliability issues of IPFS. Additionally, the high computational cost and low efficiency of the PoW consensus mechanism present further hurdles. While our experimental results support the feasibility of the approach, the controlled environment and specific dataset used may not fully capture real-world complexities. Future research should aim to address these challenges to enhance practical applicability.

Conclusion

In conclusion, the hybrid blockchain-based solution proposed in this study presents a robust approach to the pressing issue of patient privacy data security within the smart healthcare domain. The tripartite blockchain structure coupled with the application of IPFS technology ensures efficient data management and secure sharing of EMR data. Furthermore, the use of attribute encryption and BLS signature aggregation algorithms guarantees the safeguarding of patients and healthcare organizations' confidential data. Looking forward, it is anticipated that this system could be refined and optimized through continuous research, and its application could be extended beyond healthcare and potentially benefit other sectors requiring secure and efficient data management. Future studies could also explore the integration of advanced machine learning techniques for better categorization and analysis of data, ultimately improving the system's efficiency and security.

Acknowledgments

We would like to express our gratitude to the following authors for their valuable icon contributions used in this work: Icons for Patient blockchain, Healthcare provider blockchain, Social blockchain, Patient, Hospital, External users, and External network were created by Freepik; the External blockchain network and Level icons were created by Good Ware; the Data Processor and Data processor icons were created by Dewi Sari; the IPFS icons were created by Hilmy Abiyyu A.; the Insurance company icon was created by kerismaker; the Pharmacy icon was created by prettycons; the Government regulatory icon was created by nawicon; the PK&SK icon was created by Creative Stall Premium; the Di and Data icons were created by itim2101; and the External agency icon was created by geotatah. All icons are from www.flaticon.com.

References

[1] F. Casino, T. K. Dasaklis, C. Patsakis. “A systematic literature review of blockchain-based applications: Current status, classification and open issues.” *Telematics and informatics* 36: 55-81, 2019.

[2] P. Gonczol, P. Katsikouli, L. Herskind, N. Dragoni. “Blockchain implementations and use cases for supply chains-a survey.” *IEEE Access* 8: 11856-11871, 2020.

[3] W. Powell, S. Cao, T. Miller, M. Foth, X. Boyen, B. Earsman, S. D. Valle, C. Turner-Morris. "From premise to practice of social consensus: How to agree on common knowledge in blockchain-enabled supply chains." *Computer Networks* 200: 108536, 2021.
[4] T. F. Stafford, H. Treiblmaier. “Characteristics of a blockchain ecosystem for secure and sharable electronic medical records.” *IEEE Transactions on Engineering Management* 67.4: 1340-1362, 2020.
[5] L. Hang, E. Choi, DH. Kim. “A novel EMR integrity management based on a medical blockchain platform in hospital.” *Electronics* 8.4: 467, 2019.
[6] V. Chamola, A. Goyal, P. Sharma, V. Hassija, H. T. T.Binh, and V. Saxena. “Artificial intelligence-assisted blockchain-based framework for smart and secure EMR management.” *Neural Computing and Applications* 2022: 1-11.
[7] K. Miyachi, T. K. Mackey. “hOCBS: A privacy-preserving blockchain framework for healthcare data leveraging an on-chain and off-chain system design.” *Information processing & management*, 58(3): 102535, 2021.
[8] R. Kumar, N. Marchang, R. “Tripathi. Distributed off-chain storage of patient diagnostic reports in healthcare system using IPFS and blockchain.” *2020 International conference on communication systems & networks (COMSNETS)*. IEEE, 2020: 1-5.
[9] S. Tanwar, K. Parekh, R. Evans. “Blockchain-based electronic healthcare record system for healthcare 4.0 applications.” *Journal of Information Security and Applications*, 50: 102407, 2020.
[10] A. Khatoon. “A Blockchain-Based Smart Contract System for Healthcare Management.” *Electronics* 2020, 9, 94.
[11] S. David, K. Duraipandian, D. Chandrasekaran, D. Pandey, N. Sindhwani, B. K. Pandey. “Impact of blockchain in healthcare system[M]//Unleashing the Potentials of blockchain technology for healthcare industries.” *Academic Press*, 2023: 37-57.
[12] R. D. Garcia, G. Ramachandran, J. Ueyama. “Exploiting smart contracts in PBFT-based blockchains: A case study in medical prescription system.” *Computer Networks*, 2022, 211: 109003.
[13] S. E. Ali, N. Tariq, F. A. Khan, M. Ashraf, W. Abdul, and K. Saleem. “BFT-IoMT: a blockchain-based trust mechanism to mitigate sybil attack using fuzzy logic in the internet of medical things.” *Sensors*, 23(9): 4265, 2023.
[14] S. Gupta, J. Hellings, S. Rahnama, M. Sadoghi. “An in-depth look of BFT consensus in blockchain: Challenges and opportunities.” *Proceedings of the 20th international middleware conference tutorials*. 2019: 6-10.
[15] K. Košt’ál, T. Krupa, M. Gembec, I. Vereš, M. Ries, I. Kotuliak. "On transition between PoW and PoS." 2018 International Symposium ELMAR. IEEE, 2018.
[16] R. Cerchione; P. Centobelli, E. Riccio, S. Abbate, E. Oropallo, “Blockchain’s coming to hospital to digitalize healthcare services: Designing a distributed electronic health record ecosystem.” *Technovation*, 120, 102480, 2023.
[17] U. Chelladurai, S. Pandian, “A novel blockchain based electronic health record automation system for healthcare.” *J Ambient Intell Human Comput*, 13, 693–703, 2022.

[18] M. Kim, S. Yu, J. Lee, Y. Park, "Design of Secure Protocol for Cloud-Assisted Electronic Health Record System Using Blockchain." *Sensors*, 20, 2913, 2020.
[19] T. Fatokun, A. Nag, S. Sharma,"Towards a Blockchain Assisted Patient Owned System for Electronic Health Records." *Electronics*, 10, 580, 2021.
[20] K. Shuaib, J. Abdella, F. Sallabi, M. A. Serhani. "Secure decentralized electronic health records sharing system based on blockchains." *Journal of King Saud University-Computer and Information Sciences*, 34(8), 5045-5058, 2022.
[21] J. Liu, T. Liang, R. Sun, X. Du, M. Guizani, "A privacy-preserving medical data sharing scheme based on consortium blockchain." *In Proceedings of 2021 International Conference on Engineering Management of Communication and Technology*, Taipei, China, 07-11 December 2020.
[22] H. Guo, W. X. Li, M. Nejad, C. -C. Shen, "A Hybrid Blockchain-Edge Architecture for Electronic Health Record Management with Attribute-based Cryptographic Mechanisms." *IEEE Transactions on Network and Service Management*, 1-1, 2022.
[23] X. Liu, J. Yan, S. Shan, R. Wu, "A Blockchain-Assisted Electronic Medical Records by Using Proxy Reencryption and Multisignature." *Security and Communication Networks*, 13, 2022.
[24] J. Yuan, Y. Ma, W. Luo, G. Han, "B-SSMD: A Fine-Grained Secure Sharing Scheme of Medical Data Based on Blockchain." *Security and Communication Networks*, 2022.
[25] S. D. Okegbile, J. Cai, A. S. Alfa. "Practical Byzantine fault tolerance-enhanced blockchain-enabled data sharing system: Latency and age of data package analysis." *IEEE Transactions on Mobile Computing*, 2022.
[26] B. Zaabar, O. Cheikhrouhou, F. Jamil, M. Ammi, M. Abid. "HealthBlock: A secure blockchain-based healthcare data management system." *Computer Networks*, 2021, 200: 108500.
[27] P. Hegde, P. K. R. Maddikunta. "Secure PBFT consensus-based lightweight blockchain for healthcare application." *Applied Sciences*, 2023, 13(6): 3757.
[28] X. R. Zheng, Y. Lu, "Blockchain technology–recent research and future trend." *Enterprise Information Systems*, *16*(12), 1939895, 2022.
[29] H. Xiong, M. Chen, C. Wu, Y. Zhao, W. Yi, "Research on progress of blockchain consensus algorithm: a review on recent progress of blockchain consensus algorithms." *Future Internet*, *14*(2), 47, 2022.
[30] Nakamoto S. Bitcoin: a peer-to-peer electronic cash system [EB/OL].. https://bitcoin.org/bitcoin.pdf.
[31] Y. Zhang, R. H. Deng, S. Xu, J. Sun, Q. Li, D. Zheng, "Attribute-based encryption for cloud computing access control: A survey." *ACM Computing Surveys (CSUR)*, *53*(4), 1-41, 2020.
[32] M. Rasori, M. La Manna, P. Perazzo, G. Dini, "A survey on attribute-based encryption schemes suitable for the internet of things." *IEEE Internet of Things Journal*, *9*(11), 8269-8290, 2022.
[33] R. R. Al-Dahhan, Q. Shi, G. M. Lee, K. Kifayat, "Survey on revocation in ciphertext-policy attribute-based encryption." *Sensors*, *19*(7), 1695, 2019.

[34] R. Bacho, J. Loss, "On the adaptive security of the threshold BLS signature scheme." In *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*, 193-207, 2022.

[35] B. S. Egala, A. K. Pradhan, V. Badarla, S. P. Mohanty. "Fortified-chain: a blockchain-based framework for security and privacy-assured internet of medical things with effective access control." *IEEE Internet of Things Journal* 8.14 (2021): 11717-11731.

[36] B. S. Egala, A. K. Pradhan, P. Dey, V. Badarla, S. P. Mohanty. "Fortified-chain 2.0: Intelligent blockchain for decentralized smart healthcare system." *IEEE Internet of Things Journal* (2023).

Figure 1

Fig. 1. Blockchain infrastructure diagram

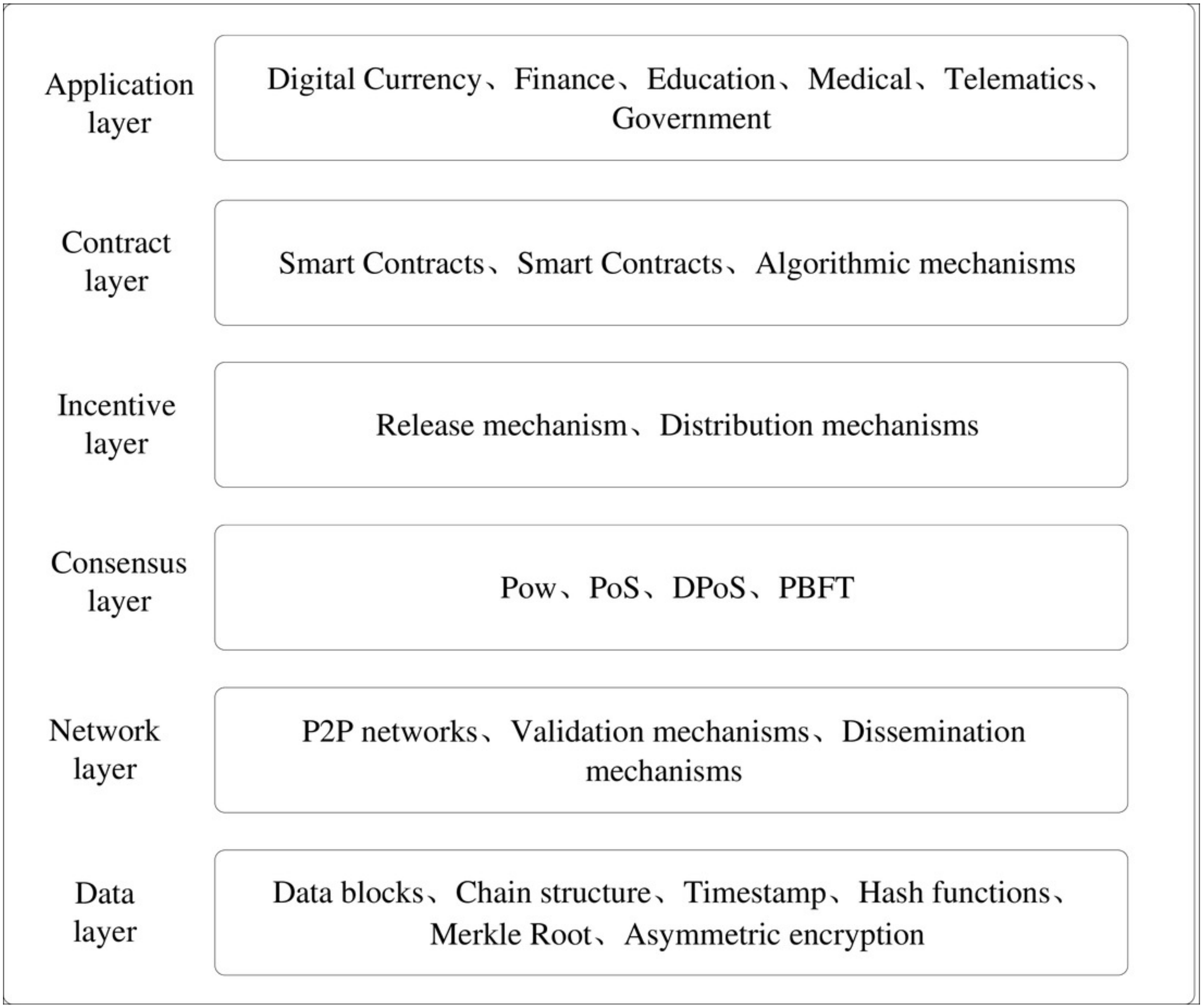

Figure 2

Fig.2. Blockchain-Based Healthcare Architecture

Patient blockchain, Healthcare provider blockchain, Social blockchian:
https://www.flaticon.com/free-icon/blockchain_10439415?term=blockchain&page=1&position=5&origin=search&related_id=10439415[p]

External blockchain network:
https://www.flaticon.com/free-icon/blockchain_2152539?term=blockchain&page=1&position=3&origin=search&related_id=2152539[p]

Data Processor:
https://www.flaticon.com/free-icon/machine_9857845?term=processor&page=3&position=66&origin=tag&related_id=9857845[p]

IPFS:
https://www.flaticon.com/free-icon/blockchain_11088312?term=blockchain&page=2&position=19&origin=search&related_id=11088312[p]

Patient:
https://www.flaticon.com/free-icon/patient_469444?term=patient&page=1&position=60&origin=tag&related_id=469444[p] **Healthcare provider:**
https://www.flaticon.com/free-icon/hospital_4005680?term=institution&page=1&position=93&origin=tag&related_id=4005680[p]

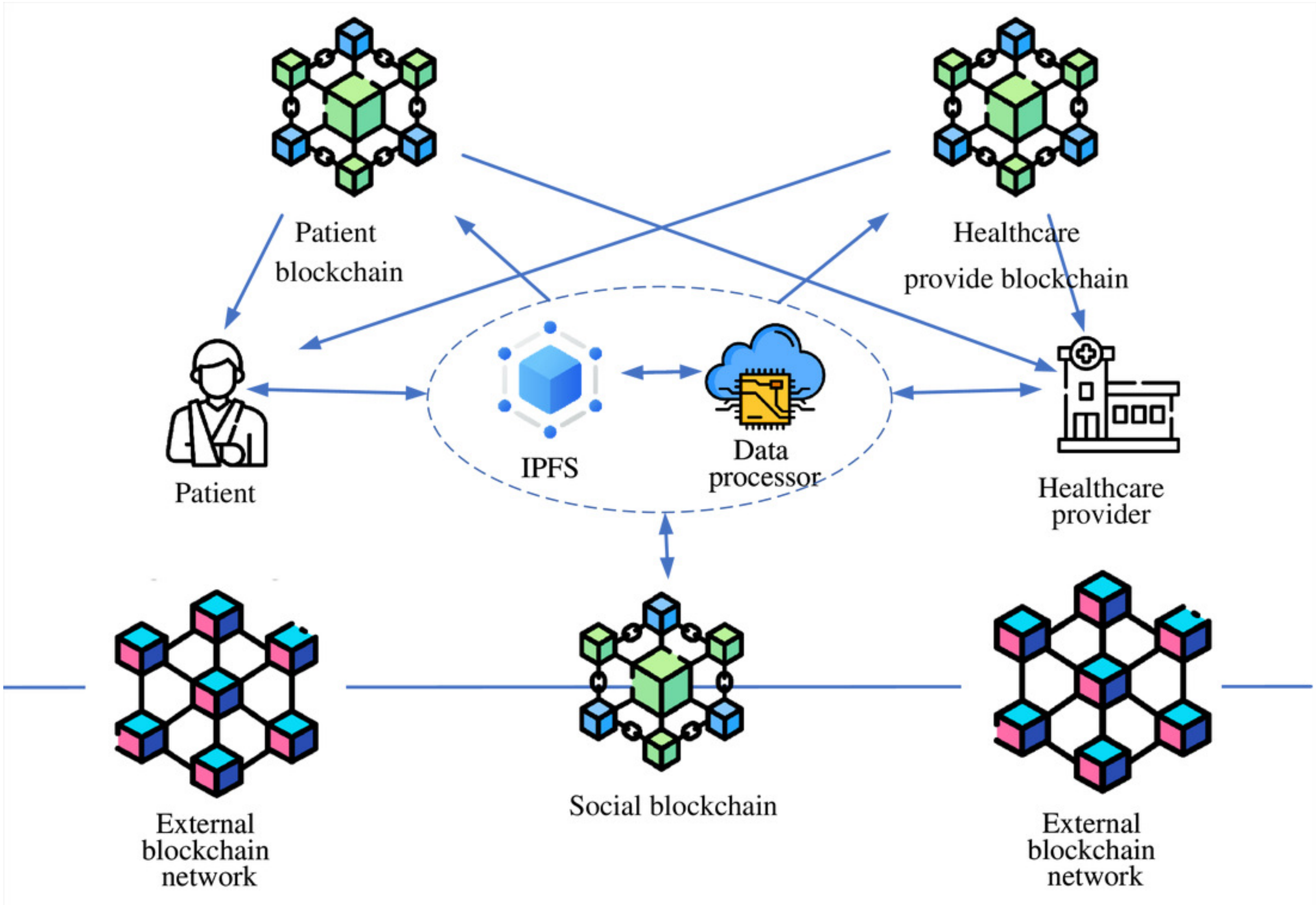
Patient
blockchain
Healthcare
provide blockchain
Patient
IPFS
Data
processor
Healthcare
provider
External
blockchain
network
Social blockchain
External
blockchain
network

Figure 3

Fig. 3 Hybrid Blockchain-Based EMR Data Sharing Flow

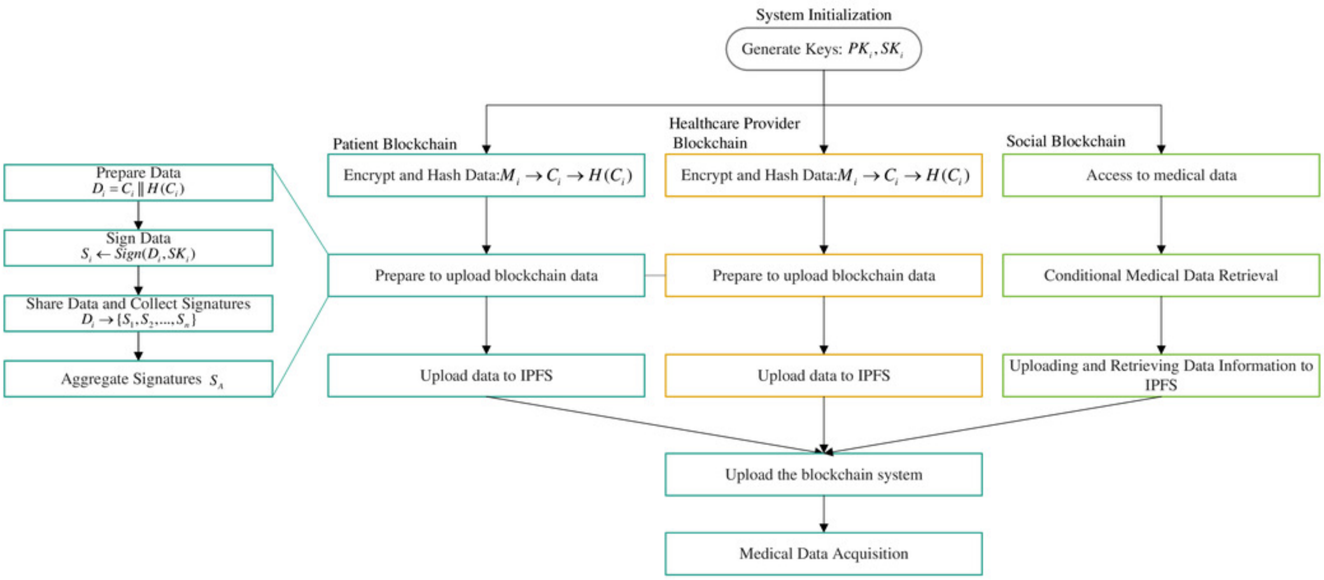

Figure 4

Fig. 4. Cross-chain signature SAs

Healthcare provider:
https://www.flaticon.com/free-icon/hospital_4005680?term=institution&page=1&position=93&origin=tag&related_id=4005680[p]

Hospital:
https://www.flaticon.com/free-icon/hospital_3063176?term=hospital&page=1&position=4&origin=tag&related_id=3063176[p]

Insurance company:
https://www.flaticon.com/free-icon/office_7902033?term=insurance+company&page=1&position=4&origin=search&related_id=7902033[p]

Pharmacy:
https://www.flaticon.com/free-icon/medicine_806962?term=pharmacy&page=2&position=65&origin=search&related_id=806962[p]

Government regulatory:
https://www.flaticon.com/free-icon/courthouse_2710012?term=government&page=1&position=50&origin=tag&related_id=2710012[p]

PK&SK:
https://www.flaticon.com/free-icon/key_2679978?term=smart+key&page=1&position=12&origin=tag&related_id=2679978[p]

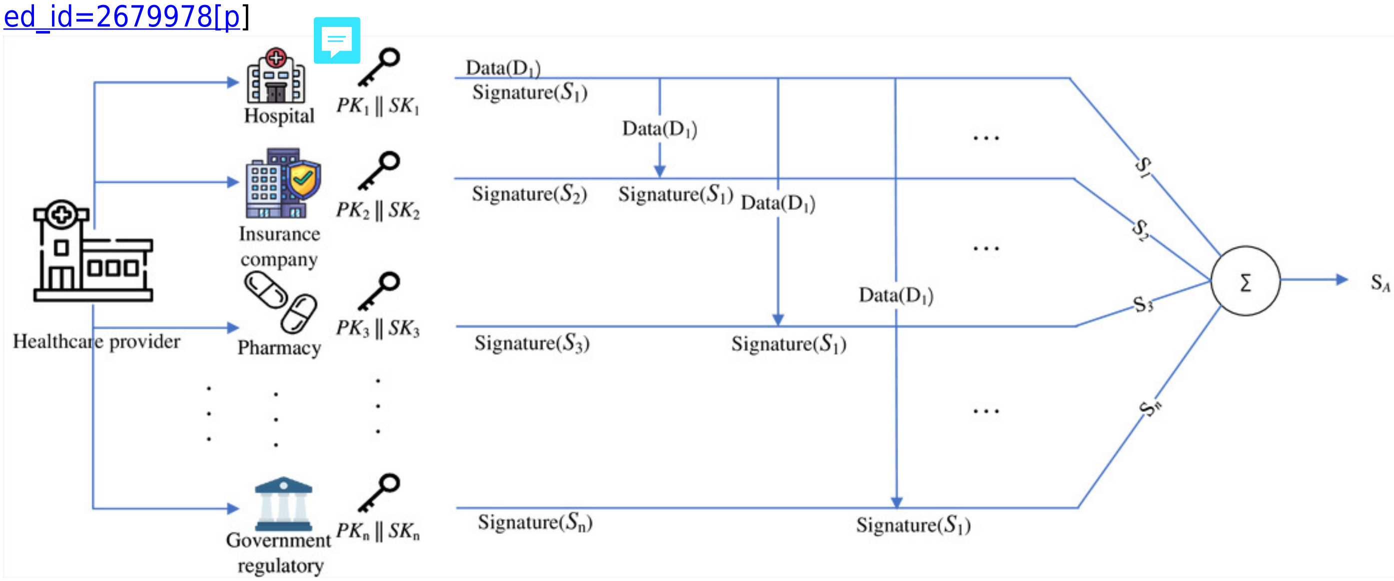

Figure 5

Fig. 5. Cross-chain signature SAf

User:
https://www.flaticon.com/free-icon/user_1077063?term=user&page=1&position=5&origin=tag&related_id=1077063[p]

D**$_i$:**
https://www.flaticon.com/free-icon/file_1573100?term=information&page=1&position=59&origin=search&related_id=1573100 [p]

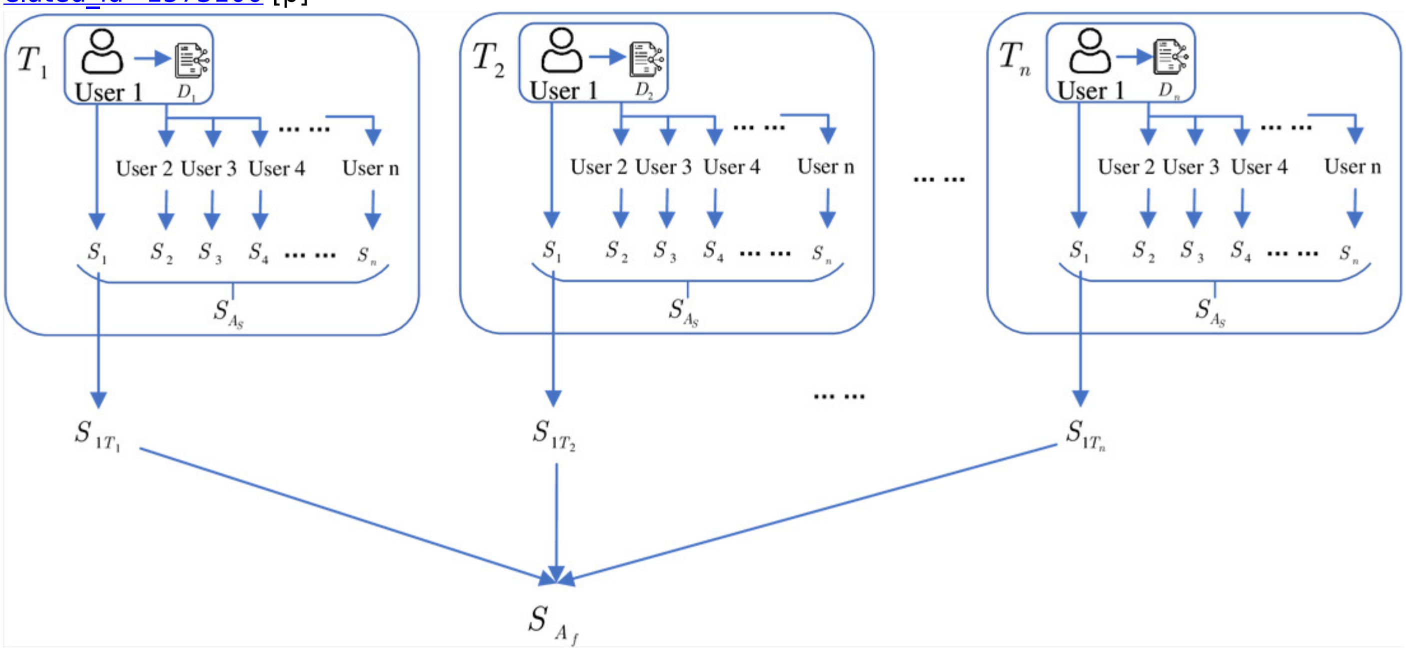

Figure 6

Fig. 6. Data-sharing process

Data:
https://www.flaticon.com/free-icon/file_1573100?term=information&page=1&position=59&origin=search&related_id=1573100

IPFS:
https://www.flaticon.com/free-icon/blockchain_11088312?term=blockchain&page=2&position=19&origin=search&related_id=11088312

Data Processor:
https://www.flaticon.com/free-icon/machine_9857845?term=processor&page=3&position=66&origin=tag&related_id=9857845

Social blockchian:
https://www.flaticon.com/free-icon/blockchain_10439415?term=blockchain&page=1&position=5&origin=search&related_id=10439415

Level:
https://www.flaticon.com/free-icon/volume-control_3871677?term=levels&page=1&position=2&origin=tag&related_id=3871677

External agency:
https://www.flaticon.com/free-icon/enterprise_993854?term=agency&page=1&position=6&origin=search&related_id=993854

External users:
https://www.flaticon.com/free-icon/user_1077063?term=user&page=1&position=5&origin=tag&related_id=1077063[p]

External network:
https://www.flaticon.com/free-icon/global-network_4207231?term=internet&page=1&position=46&origin=tag&related_id=4207231[p]

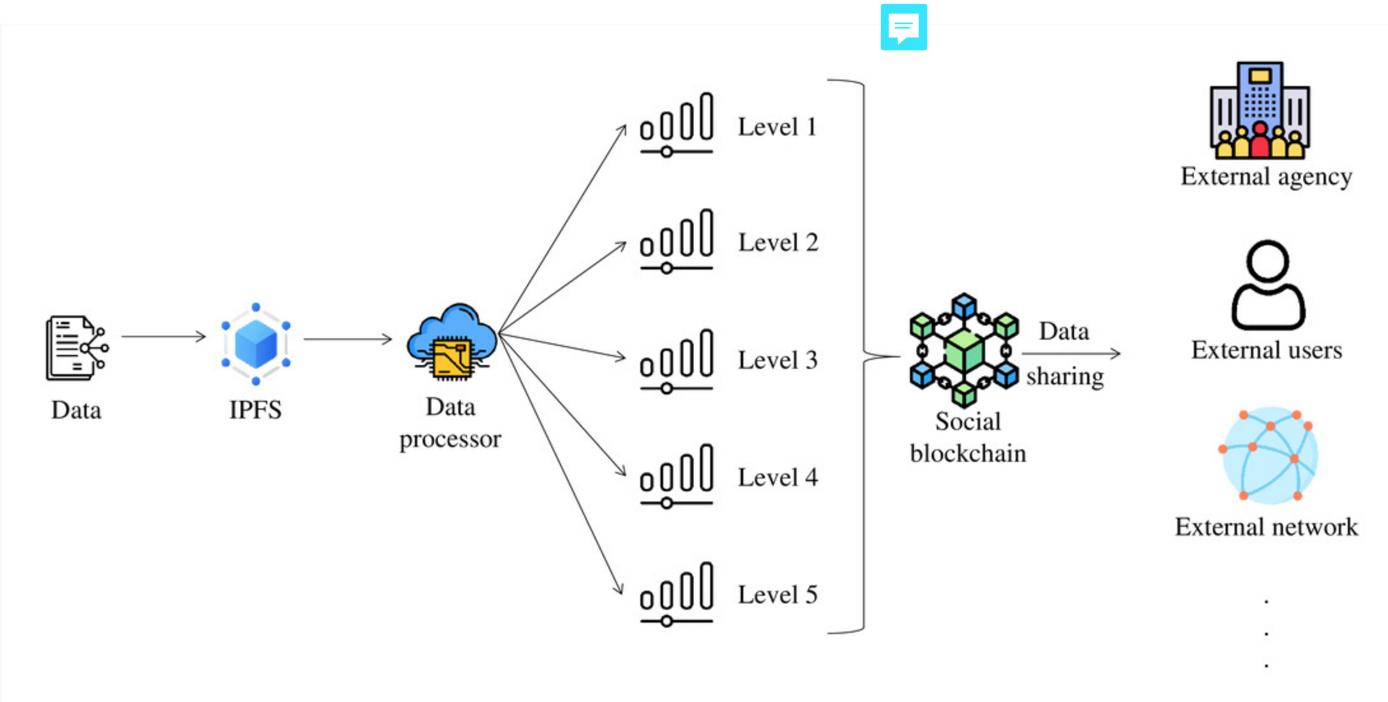
Data
IPFS
Data processor
Level 1
Level 2
Level 3
Level 4
Level 5
Social blockchain
Data sharing
External agency
External users
External network

Figure 7

Fig. 7. Comparison of the minimum number of safe blocks

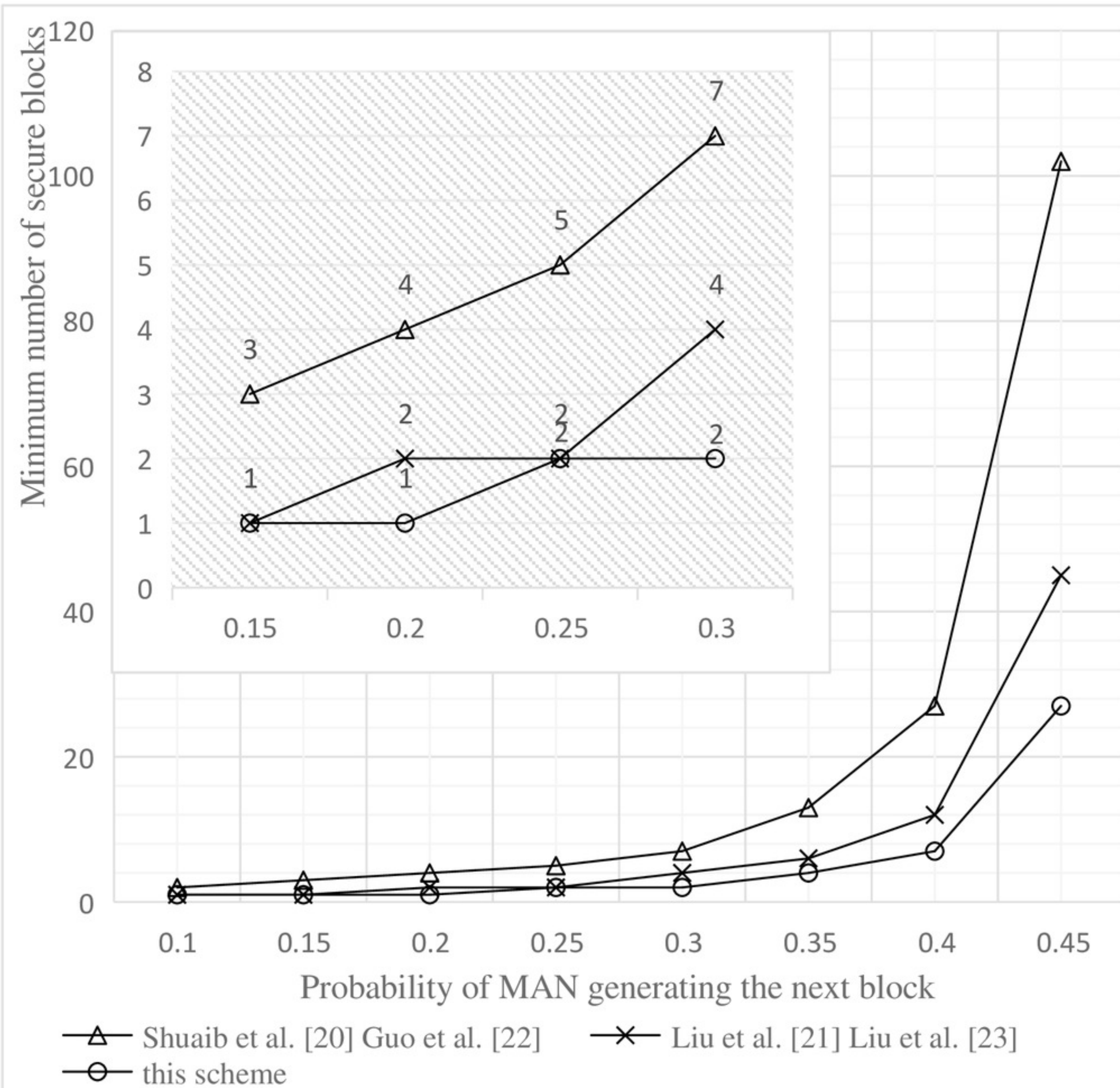

Figure 8

Fig. 8. Comparison of the minimum number of safe blocks at P0.01

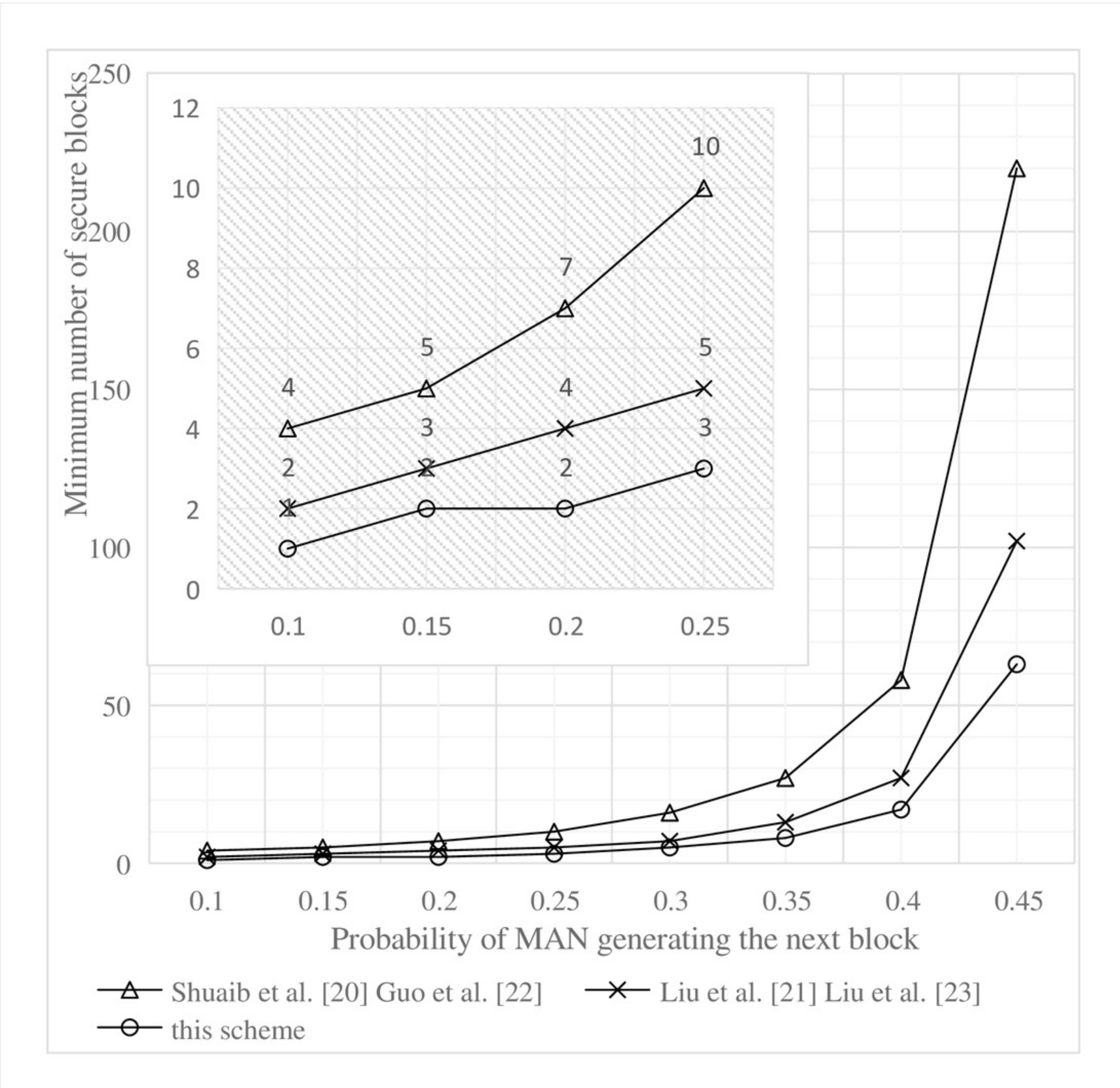

Figure 9

Fig. 9. Comparison of the minimum number of safe blocks at P0.001

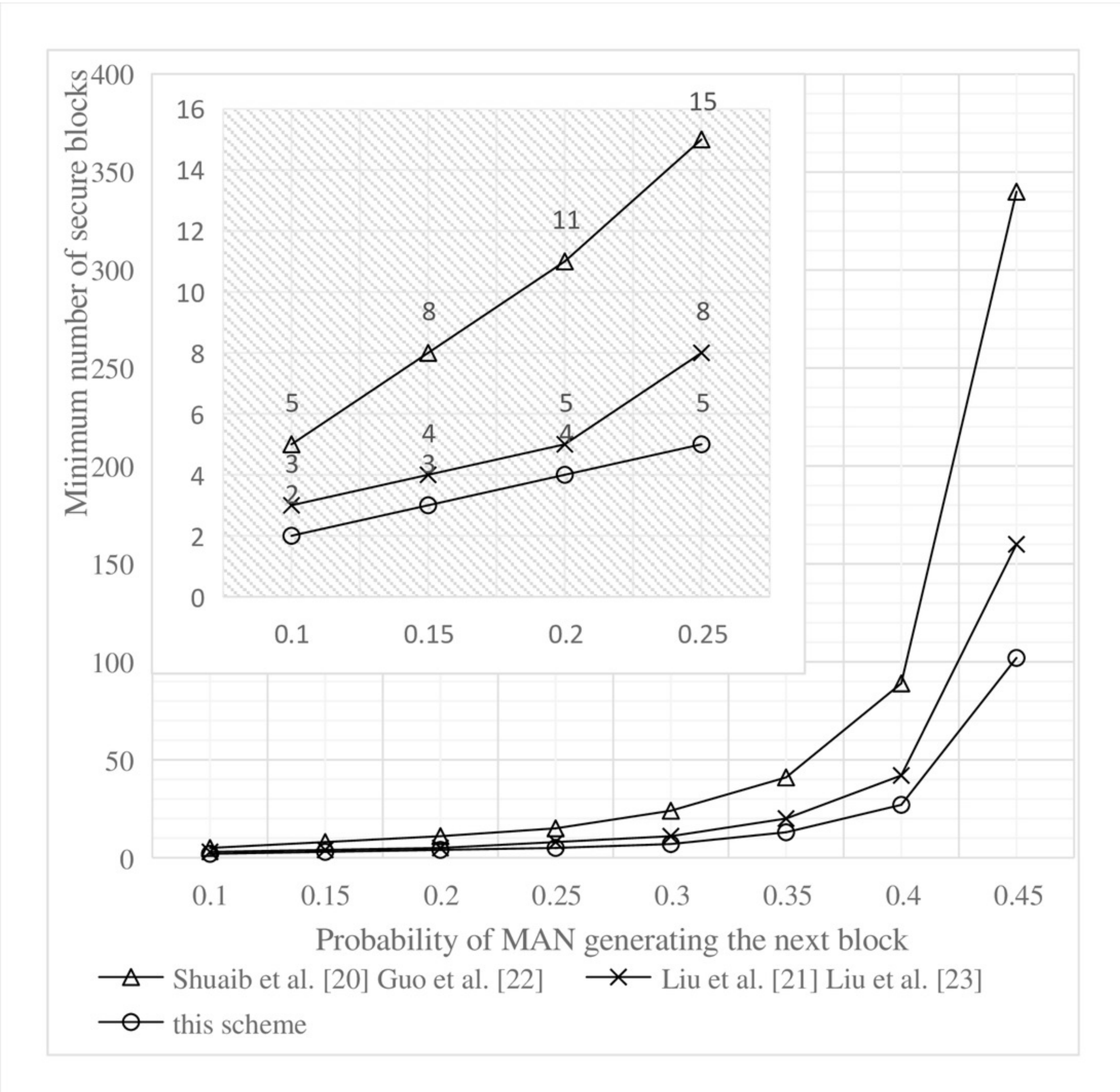

Figure 10

Fig. 10. Latency vs. Number of Transactions for Each Blockchain

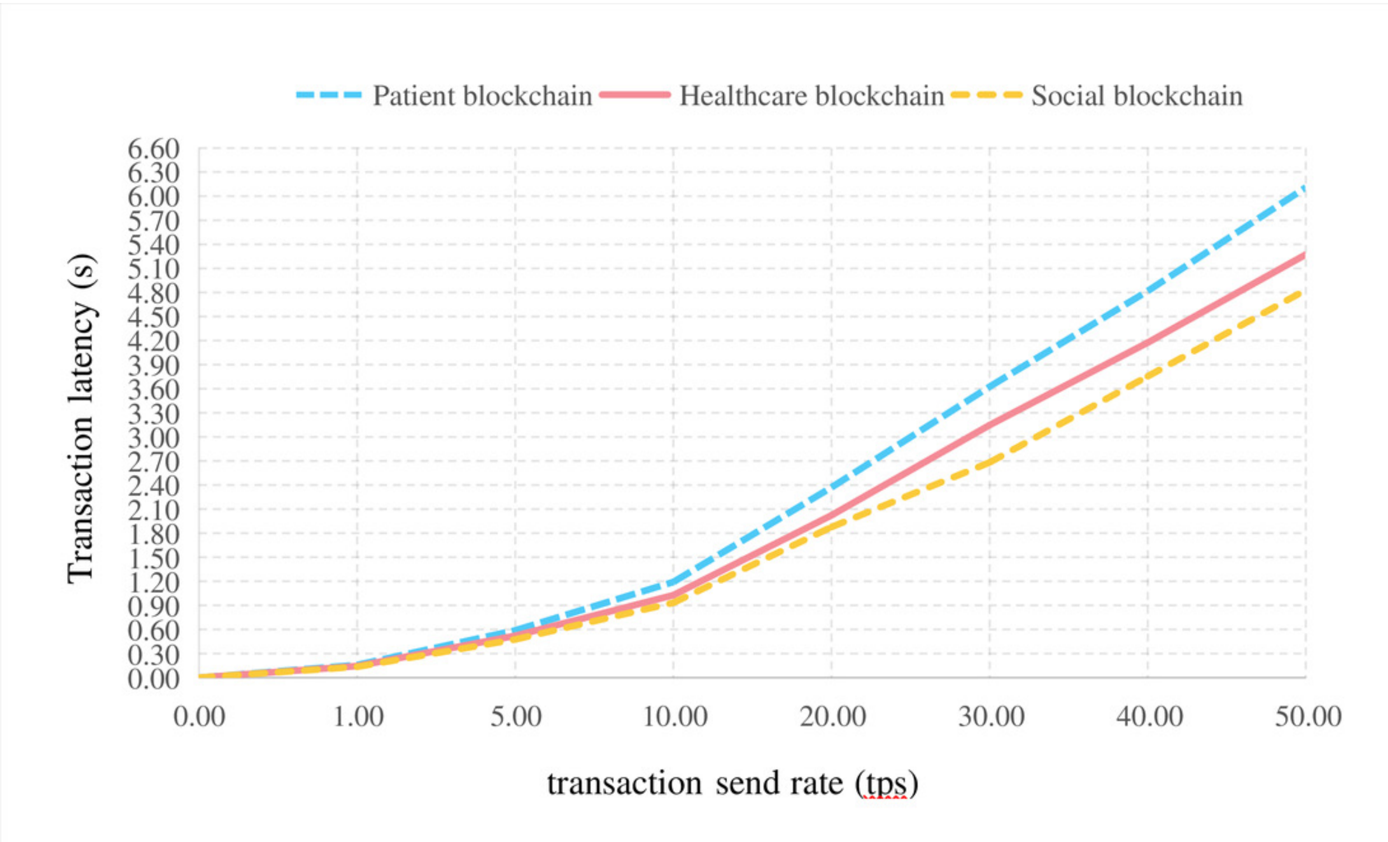

Figure 11

Fig. 11. Latency vs. Number of Transactions for Each Blockchain

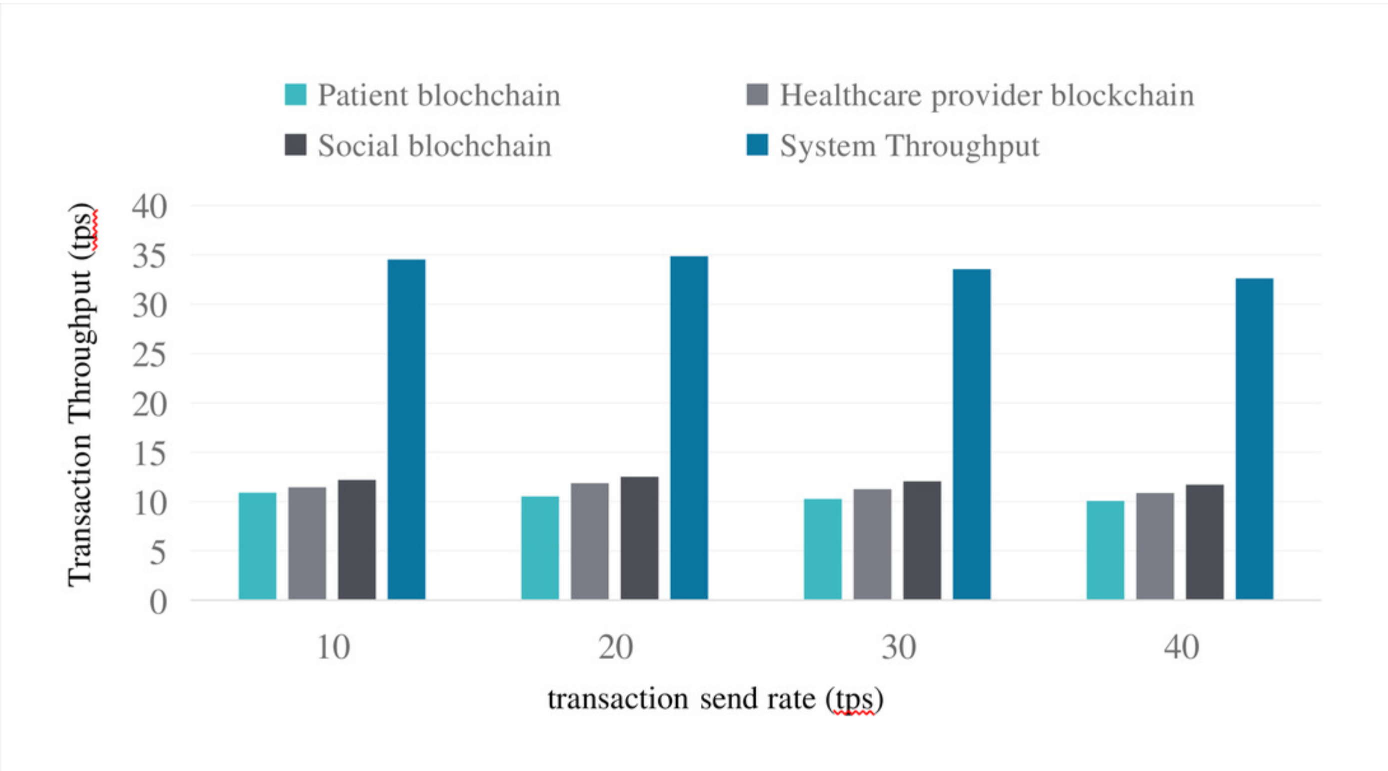

Table 1 (on next page)

Table 1. Summary of Related Work

Table 1 Summary of Related Work.

Paper	Timeline	Blockchain Type	Consensus Algorithm	Data Location	Off-chain storage	Fine-Grained Permission	Performance Eval	Cross-chain signature	Data shared externally	Data storage performance
Cerchione Rt al.[16]	2023	-	-	On-Chain	-	X	√	X	X	+
Chelladurai et al. [17]	2022	-	-	On-Chain, Off-Chain	Cloud Storage	X	X	X	X	+
Kim et al. [18]	2020	-	PBFT	On-Chain, Off-Chain	Cloud Storage	X	√	X	X	+
Fatokun et al. [19]	2021	Ethereum	-	On-Chain, Off-Chain	Cloud Storage	√	√	X	X	++
Shuaib et al. [20]	2022	Ethereum	PBFT	On-Chain, Off-Chain	Cloud Storage	√	√	X	√	++
Liu et al. [21]	2021	-	-	On-Chain, Off-Chain	Cloud Storage	X	X	X	√	+
Guo et al. [22]	2022	Hyperledger Caliper	-	On-Chain, Off-Chain	IPFS	X	√	X	X	++
Liu et al. [23]	2022	-	POS	On-Chain, Off-Chain	Cloud Storage	X	√	X	√	++
Yuan et al. [24]	2022	-	POW	On-Chain, Off-Chain	IPFS	√	√	X	X	++
Okegbile et al. [25]	2022	-	PBFT	On-Chain, Off-Chain	Cloud Storage	X	√	X	X	+
Zaabar et al. [26]	2021	Hyperledger Caliper	PBFT	On-Chain, Off-Chain	Cloud Storage	X	√	X	√	++
Hegde et al. [27]	2023	Hyperledger Caliper	PBFT	On-Chain, Off-Chain	IPFS	X	√	X	X	++
Our solution	-	Ethereum	POS	On-Chain, Off-Chain	IPFS	√	√	√	√	+++

Table 2(on next page)

Table 2. List of Symbols and Notations

Table 2. List of Symbols and Notations

Symbol	Description
$\mathbb{G}_0$	Bilinear pairing group 0
$\mathbb{G}_1$	Bilinear pairing group 1
$\mathbb{G}_T$	Target group for bilinear pairing
q	Prime order of the groups
g_0	Generator of $\mathbb{G}_0$
g_1	Generator of $\mathbb{G}_1$
H_0	Hash function treated as a random oracle
H_1	Second hash function $\mathbb{G}_1^n \rightarrow R^n$
sk	Secret key
pk	Public key
mk	Master key for CP-ABE
m	Message
e	Bilinear pairing function $e{:}\mathbb{G}_0 \times \mathbb{G}_1 \rightarrow \mathbb{G}_T$
PI_i	Personal information (name, age, gender, etc.)
C_i	Ciphertext of medical data
$H(C_i)$	Hash value of ciphertext C_i
k	Symmetric key
$\mathrm{E}_{\mathrm{P}K_u}(k)$	Symmetric key k encrypted with the patient's public key
D_i	Dataset including $D_i = C_i \| H_\mathrm{i}(C_i) \| E_{PK_u}(k)$
S_A	Aggregate signature
S_{A_s}	Aggregate signature for enhancing data privacy
S_{A_f}	Aggregate signature for achieving data fitting
T_i	Time at which a user uploads data
A	Attribute set

Table 3(on next page)

Table 3. Overview of Attack Types, Descriptions, and Defense Mechanisms in the Proposed Healthcare Blockchain Scheme

Table 3. Overview of Attack Types, Descriptions, and Defense Mechanisms in the Proposed Healthcare Blockchain Scheme.

Attack Type	Description	Defense Mechanism	Effectiveness
Data Breach	Unauthorized access to data	Attribute-based encryption, ciphertext distribution, independent signing	+++
Counterfeiting	Forging user identity or signatures	Cross-chain signature, ECC, elliptic curve cryptography	+++
Man-in-the-Middle	Message interception	Digital signatures, private key management	+++
Replay Attack	Re-sending processed transactions	Unique identifie n timestamp T	+++
Unauthorized Access	Unauthorized system access	Multi-factor authentication, access control	++
Sybil Attack	Creating multiple fake identities	Reputation system, consensus algorithms	++
DDoS Attack	Overloading system with excessive requests	Rate limiting, traffic analysis, distributed architecture	++

Table 4(on next page)

Table 4. Performance comparison

Table 4. Performance comparison

	Data generation time	Complexity of data generation	Storage space
Kim et al. [18]	$T_{ecenc} + 7T_{h}$	$O\left(\log_{n}\right)$	$M_{meta} + n \times M_{enc}$
Liu et al. [21]	$2T_{exp} + 2T_{bp} + T_{ReKeyGen}$	$O\left(\log_{n}\right)$	$M_{meta} + n \times M_{enc}$
Our solution	$n \times T_{exp} + T_{bp} + T_{s}$	$O(n)$	$M_{s} + n \times M_{enc}$

Table 5(on next page)

Table 5. Comparison of security and efficiency metrics

Table 5. Comparison of security and efficiency metrics.

Metric	Fortified Chain System (FC)	Three-Chain System (Proposed) (TC)
Minimum Secure Blocks N_{sec}	High N_{sec}^{FC}	Low N_{sec}^{TC}
Throughput (tps) T_{th}	Moderate T_{th}^{FC}	High N_{th}^{TC}
Latency (seconds) T_{lat}	High L_{lat}^{FC}	Low L_{lat}^{TC}
Redundancy R_{red}	High R_{red}^{FC}	Low R_{red}^{TC}