# Research on Blockchain Privacy Preservation in Healthcare Systems

Keke Wang[1], Pengfei Lu[2,*], and Zhenghua Xin[1]
[1]School of Information Engineering, Suzhou University, Suzhou, Anhui, China
[2]School of Chemistry and Chemical Engineering, Suzhou University, Suzhou, Anhui, China
kekewang@ahszu.edu.cn, lupengfei0826@163.com, xinzhenghua@ahszu.edu.cn
*corresponding autho r

*Abstract*—Medical information is of great significance in people's daily lives, being closely related to aspects such as insurance, drug records, and driving licenses. To address the issue of medical data leakage, this paper presents a decentralized zero-knowledge proof mechanism based on blockchain technology for privacy preservation in healthcare systems. This mechanism offers a viable solution for secure data storage and application. Additionally, this paper conducts an in-depth literature review of current research, analyzes the advantages and disadvantages of existing blockchain-based zero-knowledge proof mechanisms, and elaborates and summarizes previous solutions by comparing the proposed mechanism with traditional ones. Experimental results demonstrate that the proposed mechanism has lower time and space complexity compared to traditional methods and functions effectively in a decentralized network.

*Keywords: blockchain; zero-knowledge proof; lattice; privacy protection*

## 1. INTRODUCTION

In the digital age, the privacy protection of medical data has become a crucial issue that urgently needs to be addressed globally. In recent years, frequent global medical data breaches not only pose a threat to the personal privacy and security of patients but also present a severe challenge to the trust foundation of the medical industry. For example, since 2021, some insurance agencies, in order to precisely sell insurance types such as "surgical accident insurance" [1], illegally obtained a large amount of patients' medical and health information through cooperative hospitals and conducted insurance promotions to the relevant patients. In 2024, the US government received reports of more than 700 medical data breaches, and over 180 million patient medical records were leaked [2]. The illegal acquisition and abuse of patients' medical and health information led to patients receiving scam calls from lawbreakers, triggering public concerns about medical information security. These incidents not only expose the urgency of medical data protection but also prompt us to seek more advanced and reliable technical means to strengthen data privacy protection.

Blockchain technology, as a distributed ledger technology, with its characteristics such as decentralization, transparency, and immutability, provides new ideas for solving the problem of medical data privacy protection. Especially when blockchain is combined with advanced encryption technologies such as zero-knowledge proof, it can verify the authenticity of data without exposing specific information, thus greatly enhancing the security and privacy protection level of data. In the medical field, the combined application of this technology has huge potential value.

For precision medicine, it can enable secure data sharing while protecting patient privacy. This allows researchers to access high-quality data for in-depth clinical research, which helps to more accurately identify disease-specific biomarkers and develop personalized treatment plans. In terms of medical resource allocation, blockchain-based systems can track the flow of resources in real-time to ensure transparency and fairness. Zero-knowledge proof can verify the eligibility of resource use without disclosing sensitive patient information, reducing waste, fraud, and abuse of resources, thereby making the medical system more efficient.

In summary, the urgency of medical data privacy protection cannot be ignored, and the application of the combination of blockchain and zero-knowledge proof in the medical field provides us with new solutions. This research aims to deeply explore the potential value of this technology, with a view to providing strong support for promoting the safe, efficient, and intelligent development of the medical industry.

## 2. DESIGN OF BLOCKCHAINS IN HEALTHCARE SYSTEMS

### 2.1. Architecture Design

The blockchain model can be divided into a five-layer structure, namely the application layer, access control layer, consensus layer, network layer, and data layer. As shown in Figure 1, the data layer receives transaction data from the network layer and packages it into blocks. Blocks are linked to the previous block through hash values to form a chain-like structure (as shown in Figure 2). Through hash functions and encryption technologies, the data layer encrypts and stores sensitive information (such as patient medical records) to ensure the storable, traceable, and nontamperable nature of medical data [3].

In the network layer, each node adopts the peer-to-peer (P2P) protocol to achieve decentralization [4]. The network layer receives transaction requests from the application layer and broadcasts them to the entire network. Meanwhile, it passes the verified blocks to the data layer for storage. Through anonymous communication and encrypted transmission, the network layer protects the privacy of data during transmission.

The consensus layer receives transaction data from the network layer. The selective disclosure mechanism ensures that only authorized nodes can participate in the consensus process, protecting the privacy of node identities and data.

2835-2823/25/$31.00 ©2025 IEEE
DOI 10.1109/ISSSR65654.2025.00048

The access control layer receives user requests from the application layer. After verifying the permissions, it passes the data requests to the data layer. At the same time, it decrypts the encrypted data in the data layer and returns it to the application layer. Through attribute-based encryption (ABE) and zero-knowledge proof technologies, the access control layer ensures that users can only access the data within their authorized scope [5].

The application layer is the interface for users to interact with the blockchain, providing data visualization and operation interfaces. It supports various application scenarios, such as medical data query, telemedicine, etc. The application layer receives user requests and passes them to the access control layer for permission verification. After verification, the application layer retrieves data from the data layer and presents it to the user. The application layer further protects user privacy through user identity anonymization and data desensitization technologies [6].

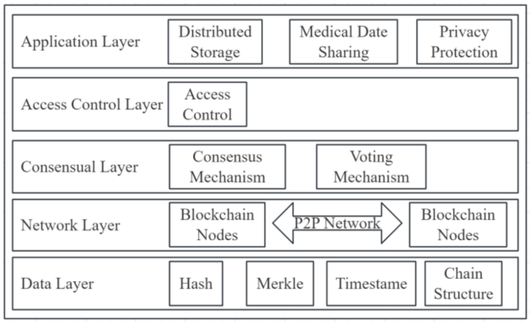

Figure 1. Blockchain architecture model

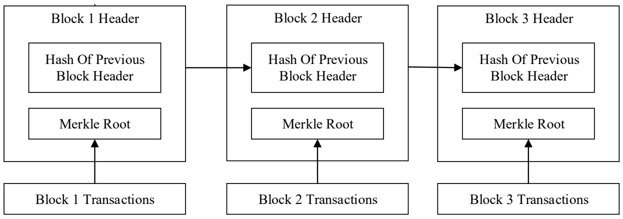

Figure 2. Blockchain chain structure

## 2.2. Functional Applications

In the medical system, the three most common entities are patients, hospitals, and research institutions, and there is a many-to-many relationship among them. To facilitate critically ill patients to receive telemedicine and multi-center joint medical treatment, and to promote the rapid development of medicine, the real-time sharing of patient medical data among different medical institutions and doctors is of utmost importance. However, the data formats and standards of various medical institutions are inconsistent, and due to data security and privacy concerns, data sharing faces numerous difficulties. For example, in the research on rare diseases, since the cases are scattered in hospitals in different regions, it takes a long time and is difficult to collect enough sample data.

The application of blockchain technology can break these barriers. By leveraging the distributed ledger feature of blockchain, various medical institutions can encrypt and store the desensitized medical data of patients on the chain, and the data ownership still belongs to each institution. Based on the smart contract of blockchain, the data access rights can be precisely set. After obtaining authorization, researchers can analyze and calculate the data through encryption algorithms without exposing the original data. For instance, in a multi-center study on tumor immunotherapy, hospitals in different regions uploaded the treatment plans and efficacy data of patients to the blockchain. Researchers used the homomorphic encryption technology to conduct statistical analysis on the data on the chain, obtained the correlation between treatment effects and patient characteristics, provided a basis for the optimization of tumor immunotherapy, greatly improved the research efficiency, and accelerated the transformation of scientific research achievements.

Blockchain technology ensures the secure transmission and storage of telemedicine data through encryption mechanisms and distributed storage. When a patient undergoes examinations in a primary medical institution, data such as images and test reports will be encrypted and uploaded to the blockchain. When an expert from a superior hospital receives a consultation request, they can obtain the access rights to the patient's data through the identity authentication and permission management system of the blockchain. During the data transmission process, the hash algorithm of the blockchain ensures the integrity of the data. Once the data is tampered with, the hash value will change and can be detected in a timely manner. Take a patient in a remote area receiving a remote consultation as an example. The local medical institution uploads the patient's vital sign data such as electrocardiogram and blood pressure to the blockchain. Experts from the superior hospital can obtain these data in real-time and securely, provide accurate diagnosis and treatment guidance for the patient, effectively improve the accessibility of medical services, and enable patients in remote areas to enjoy high-quality medical resources.

# 3. LATTICE-BASED ENCRYPTION ALGORITHM DESIGN

With the continuous development of the Internet, computing capability has achieved impressive progress in the past years. Traditional encryption and decryption algorithms are slightly inferior in the face of quantum attacks. The encryption and decryption algorithms constructed based on some average case hard problems of the lattice can effectively resist quantum attacks. For example, the LWE(Learning With Errors) hard problem [7].

A random matrix and items (A,v) are generated by randomly selecting positive integers n,m and q as parameters, where $A\in Z_q^{n\times m}$, $v\in Z_q^m$, small vectors are randomly selected by Gaussian sampling as error vectors $e\in Z_q^m$. It is difficult to find a vector $s\in Z_p^n$ that makes the equation v = As + e work and that's how the LWE problem works.Although LWE encryption has an important place in modern cryptography and is one of the core technologies of post-quantum cryptography, its security is resistant to quantum attacks, but it also has some shortcomings and challenges. For example, the public key, private key, and ciphertext sizes of LWE encryption are usually large, increased storage and transmission overhead limits its application in resource-constrained environments. And the complexity of noise management, the security of LWE encryption depends on noise e, but the management and control of noise is more complex in implementation. The noise must be small enough to ensure correct decryption. The distribution and size of noise directly affect the safety and efficiency of the scheme.

Based on the above theory, we carry out the design of a lattice-based encryption algorithm. A series of hypotheses are as follows: n is the security factor of the system itself, q is a prime number and q = poly(n)(Defining the LWE problem on a polynomial ring can significantly reduce the size of the key and ciphertext),there exists an integer m and m = O(n log q), there exists a Gaussian noise distribution as $\varphi_\beta^m(q)$ with parameter β and β = poly(n), and then in combination with the trapdoor generation algorithm GenTrap, a random matrix $A\in Z_q^{n\times m}$ and a trapdoor matrix $R\in Z_q^{n\times m}$ can be obtained based on the encryption of the data information that can be treated as public and private key matrix, respectively, which satisfies the condition AR = 0 (mod)q. Next, we denote the plain text data encoding vector by b with b = (b₁, ..., bₘ) $\in Z_q^m$, meanwhile, a vector s is randomly selected from $Z_q^m$ and then from $\varphi_\beta^m(q)$ a randomly selected error vector e, with $e\in Z_q^m$. The ciphertext C can be obtained according to equation (1), and finally, complete the encryption process.

$$c^T = s^T b^T A + qe^T + b^T \tag{1}$$

For decryption, the ciphertext C, the trapdoor function R, and the value of the prime q are known, and AR = 0 (modq)

$$(c^T \times R)\mathrm{mod}q = ((s^T b^T A + qe^T + b^T) \times R)\ \mathrm{mod}q = （s^T + b^T）AR\mathrm{mod}q + qe^T R\mathrm{mod}q + b^T R\mathrm{mod}q = b^T R\mathrm{mod}q \tag{2}$$

When q is large enough, the value of the original b vector can be found.

It is clear from the LWE problem that it is extremely difficult to find the random vectors s and e, and hence the original text b, without knowing the trapdoor function R. The influence of noise e on the decryption process is reduced. The scheme herein can effectively resist high arithmetic attacks, such as quantum attacks.

## 4. ZERO-KNOWLEDGE PROOFS FOR SMART CONTRACTS

Zero-knowledge proofs (ZKPs) enable a prover to convince a verifier of a statement's validity without leaking any extraneous information. As a versatile cryptographic tool, ZKPs have broad applications in privacy-preserving protocols and are currently a key research focus in lattice-based cryptography, particularly for post-quantum secure constructions.

Zero-knowledge proof technology belongs to the verifiable off-chain computation method of the off-chain expansion means of the blockchain. The data is firstly computed off-chain securely through the zero-knowledge proof to generate evidence, and then the computation results are submitted to the on-chain smart contract program for verification, forming a privacy protection architecture of off-chain computation and on-chain verification [8]. This forms an off-chain calculation and on-chain verification privacy protection architecture. This enables the patient's identity to be verified while protecting the patient's information. The patient, as the proving party, needs to prove his identity to the research institution, but at the same time does not want to reveal his personal information in advance. In this paper, a zero-knowledge proof scheme is designed to help patients prove their identity while protecting their privacy.

Transactions between patients and research institutions are done through smart contracts. The patient's medical data is represented by $w^T$. The patient is required to prove that his medical data is true, but not disclose any of it. The smart contract randomly generates integers n, m and q as parameters, with a small vector randomly selected by Gaussian sampling as the error vector $e\in Z_q^m$.Suppose there are n verification nodes, and each verification node generates a value to be summed Vᵢ, i=1,2,... n, each node knows only its own value Vᵢ and does not know the value of the other nodes. The node sends the generated value to be summed to the smart contract, the function of the smart contract is to sum.

$$V_m = V_1 + V_2 + V_3 + \ldots + V_n \tag{3}$$

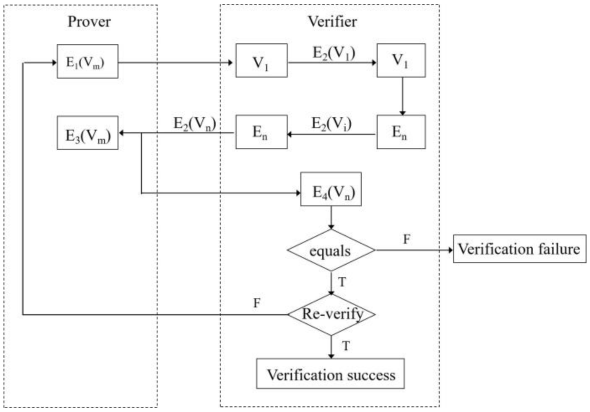

Figure 3. Zero-knowledge proof process

Step 1: The prover P (patient) calculation send the calculation result E₁(Vₘ) to the validation node by the following equation:

$$E_1(V_m) = w^T V_m + qe^T \tag{4}$$

Step 2: Verify that node V generates a secret random number $R^T$, which is first computed by node 1.

$$E_2(V_1) = R^T V_1 + qe^T \tag{5}$$

Node 1 sends the result of the calculation to node 2, and when the number of nodes i>1, iterates through the calculation in the following order.

$$E_2(V_i) = E_2(V_{i-1}) + R^T V_i + qe^T \tag{6}$$

When the calculation reaches the nth node, the result is obtained.

$$E_2(V_n) = R^T(V_1 + V_2 + V_3 + \ldots + V_n) + nqe$$
$$= R^T V_m + nqe \quad (7)$$

Step 3: The nth verification node sends the computed result $E_2(V_n)$ to the prover P.

Step 4: P-calculation send the value of $E_3(V_m)$ to the validation node.

$$E_3(V_m) = w^T E_2(V_n) \mathrm{mod} q = w^T R^T V_m \mathrm{mod} q \quad (8)$$

Step 5: Validate the node calculation.

$$E_4(V_n) = R^T E_1(V_m) \mathrm{mod} q = R^T w^T V_m \mathrm{mod} q \quad (9)$$

Step 6: The verification node compares the values of $E_3(V_m)$ and $E_4(V_n)$ and if they are equal, proves that P has the correct identity information.

Step 7: Return to the first step and regenerate the random number for verification T times. If the values of $E_3(V_m)$ and $E_4(V_n)$ are equal each time, the patient's identity is proved to be correct.

## 5. Performance analysis

The success of a zero-knowledge proof mainly depends on whether it has the following three characteristics [9]:

(1) Completeness: If the proving party P does know the identity information $w^T$ and the verification steps are all performed correctly and both parties are honest, then the verification must succeed. Through the above calculations, according to the exchange method, it is clear that the verification will pass when the information is correct. The algorithm is complete.

(2) Robustness: If the proving party P does not know the identity information $w^T$, the authentication must fail. No attacker can impersonate P to achieve successful authentication without knowing the identity information. If P does not know the $w^T$ value, the correct result will not be known in the fourth computation step, and there is only a 1/q chance of masking the correct result. However, after T rounds of verification, the probability of being correct is only$(1/q)^T$. Thus, the larger the value of the random prime q is chosen in a finite field, the greater the number of repetitions T, and the more robust the checksum is. The use of multiple nodes for verification also prevents MITM attacks. With an unknown number of verification nodes, it is almost impossible to eavesdrop on the data $V_i$ sent by all verification nodes, especially in the case of T-round verification, where continuous listening is extremely costly and the attacker is not sure if there are any missed listens and thus cannot determine the correctness of the listening results. Current research indicates that for lattice-based hard problems such as LWE, no quantum algorithm with polynomial speedup has been constructed to date, which provides a theoretical foundation for the post-quantum security of lattice-based cryptography.

(3) Zero-knowledge: After the authentication is completed, the identity information $w^T$ remains known only to P. Neither the verifier V nor the attacker can obtain any information about the identity information $w^T$ itself. This scheme does not send the $w^T$ to anyone and, as the robustness suggests, the probability of being right is small. Therefore, the scheme has zero-knowledge[10].

In all, the verification process is simple to compute, fast to verify and requires little storage space, which can save a lot of arithmetic and storage resources, so this verification method is more practical.

## 6. Conclusion and outlook

In conclusion, the integration of blockchain and zero-knowledge proof in the medical system, as proposed in this study, has demonstrated remarkable practical application effects. In real-world medical scenarios, this mechanism has successfully enabled secure and efficient sharing of medical data. For example, in multi-center clinical trials, researchers from different institutions can access and analyze the encrypted medical data of patients after obtaining proper authorization. Through the zero-knowledge proof technology, the privacy of patient data is strictly protected while the research requirements are met. The distributed storage characteristic of blockchain ensures the integrity and nontampering of medical data, reducing the risk of data loss and malicious modification. Moreover, in the process of telemedicine, doctors can securely access the necessary medical records of patients in real-time, improving the accuracy and timeliness of diagnosis and treatment.

Looking ahead, we have a clear-cut research plan. In the field of noninteractive zero-knowledge proof research, our first step is to conduct in-depth theoretical research on advanced cryptographic algorithms and protocols. We aim to optimize the existing noninteractive zero-knowledge proof models to make them more suitable for the complex data structures and high-security requirements of the medical system. Next, we will carry out simulations and experiments in a simulated medical environment to test the performance and security of the improved noninteractive zero-knowledge proof algorithms. We expect to achieve a significant reduction in the computational complexity and proof-generation time while maintaining high-level privacy protection.

To promote the application of research results to a wider range of medical scenarios, we plan to cooperate with multiple medical institutions. We will first pilot the proposed mechanism in several representative hospitals and research institutions, collect feedback and data, and continuously optimize the system according to the actual situation. Then, we will organize training and promotion activities to help medical staff and relevant personnel better understand and use this technology. Through these efforts, we hope to popularize the integration of blockchain and zero-knowledge proof in the medical field, ultimately improving the overall level of medical data security and the efficiency of medical services.

## Acknowledgment

This research was supported by School-level Key Discipline at Suzhou University in China: Computer science and technology (2019xjzdxk1), School-level Key project of natural science research at Suzhou University in China (2022ykf04, 2024yzd12), and Key Research and Development Project of Anhui Province in China (AE202102).

## References

[1] A. Atadoga, O. A. Elufioye, T. T. Omaghomi, O. Akomolafe, I. P. Odilibe, and O. R. Owolabi, “Blockchain in healthcare: A comprehensive review of applications and security concerns.” International Journal of Science and Research Archive, vol. 11, pp. 1605–1613, February 2024.

[2] I. Yaqoob, K. Salah, R. Jayaraman, and Y. A-Hammadi, “Blockchain for healthcare data management: opportunities, challenges, and future recommendations.” Neural Computing and Applications, vol. 34, pp. 11475–11490, January 2021.

[3] H. Sahu, S. Choudhari, and S. Chakole, “The use of blockchain technology in public health: lessons learned.” Cureus, vol. 16, pp. e63198, June 2024.

[4] M. Sethi, J. Arora, V. Baggan, J. Verma, and M. Sethi, Next-generation Cybersecurity, Singapore: Springer Nature Singapore, 2024, pp. 135–158.

[5] X. R. Zheng, and Y. Lu, “Blockchain technology–recent research and future trend.” Enterprise Information Systems, pp. 1–23, June 2021.

[6] A. P. Balcerzak, E. Nica, E. Rogalska, M. Poliak, T. Klie š tik, and O. M. Sabie, “Blockchain technology and smart contracts in decentralized governance systems.” Administrative Sciences, vol. 12, pp. 96, August 2022.

[7] S. Dhingra, R. Raut, A. Gunasekaran, B. K. R. Naik, and V. Masuna, “Analysis of the challenges for blockchain technology adoption in the Indian health-care sector.” Journal of Modelling in Management, vol. 19, pp. 375–406, February 2024.

[8] P. K. Ghosh, A. Chakraborty, M. Hasan, K. Rashid, and A. H. Siddique, “Blockchain application in healthcare systems: a review.” Systems, vol. 11, pp. 38, January 2023.

[9] O. Kuznetsov, A. Rusnak, A. Yezhov, D. Kanonik, K. Kuznetsova, and S. Karashchuk, “Enhanced security and efficiency in blockchain with aggregated zero-knowledge proof mechanisms.” IEEE Access, vol. 12, pp. 49228–49248, April 2024.

[10] L. Zhou, A. Diro, A. Saini, S. Kaisar, and P. C. Hiep, “Leveraging zero knowledge proofs for blockchain-based identity sharing: A survey of advancements, challenges and opportunities.” Journal of Information Security and Applications, vol. 80, pp. 103678, February 2024.