A Study on Privacy Protection under Blockchain Data Transaction Based on Legal Perspective

Yajie Zhang

(Corresponding author: Yajie Zhang)

Law Teaching and Research Department, Party School of the CPC Shijiazhuang Municipal Committee
Shijiazhuang, Hebei 050000, China
jiepian84888@yeah.net

(Received Apr. 8, 2022; Revised and Accepted Mar. 23, 2023; First Online Apr. 27, 2023)

Abstract

The privacy protection issue in blockchain data transactions has received much attention. This paper first briefly analyzed the privacy protection issue in data transactions and then explained the current legal problems from the legal perspective. After introducing blockchain technology, the homomorphic encryption method was applied to protect privacy under blockchain data transactions. A privacy protection method based on the Paillier algorithm was designed, and the method's performance was analyzed. After verifying the correctness of the algorithm, the comparison with the zkSNARKs technique revealed that our method had better efficiency. When the key was 2,048 bits, the encryption homomorphism time was 0.17 ms, and the encryption and decryption time was 185.75 ms and 256.12 ms, i.e., the method had a slight burden on the system and could meet the needs of practical use. The results verify that the designed method was reliable. Therefore, it can be applied in actual blockchain data transactions.

Keywords: Blockchain; Data Transaction; Law; Paillier Algorithm; Privacy Protection

1 Introduction

As society continues to develop toward informationization, the importance of data is becoming more and more prominent. In people's production and life, a large amount of data is generated, and data transaction, which is gradually becoming a method of data circulation, provides services for scientific research, social management, etc. [9]. However, there is a lot of information related to personal privacy in the data, and it is easy to be illegally leaked in the process of transaction; therefore, privacy protection under data transactions has become a matter of great concern. From the legal point of view, illegal trading, illegal dumping, and illegal leakage in the process of data transaction involve the infringement of personal information and privacy, etc. Although data is not a tangible object, data transaction has much in common with the act of buying and selling and can be analyzed by referring to the legal provisions of the act of buying and selling.

The two subjects in the data transaction need to transmit the real and valid data securely and use it legally, thus ensuring a healthy data flow. However, there are still many loopholes in the legal regulation of data transactions, and the relevant legal provisions are not clear enough, which makes it very difficult to protect privacy under data transactions. In this case, in order to improve the security of data transactions, we can start from the technical aspect. Blockchain, as a decentralized, open and transparent technology [4, 21], has good security, but blockchain is not designed for data transactions [3,7]. Further research and design are still needed for privacy protection under blockchain data transactions [6, 22]. Onik et al. [17] proposed a blockchain-based personal identifiable information management system to limit the leakage problem of personal data and verified the privacy of the approach. Wang et al. [25] proposed a dual blockchain privacy protection approach to achieve secure data interaction between users, doctors, and hospitals through user chains and medical chains and found through experiments that the approach had low communication overhead.

Devidas et al. [8] designed a decentralized group signature scheme for solving the trust problem of centralized group managers and the privacy problem of users and verified the security and correctness of the scheme. Nóbrega et al. [16] proposed a privacy preserving record linking method based on blockchain technology and verified the effectiveness of the method with real-world data sources. This paper mainly studied the privacy protection problem under blockchain data transactions, designed a privacy protection method based on the Paillier algorithm, and verified the reliability of the method through experiments. This work provides a new method to further improve the security of blockchain and ensure the privacy of

data transactions.

2 Blockchain Privacy Protection Method Based on the Paillier Algorithm

2.1 Data Transactions and the Law

With the development of the network, data is also becoming more and more a new productivity. The research and analysis of data can obtain many valuable information, so it has been applied in recommendation systems, user analysis, etc. [14], but in this process, it is likely to cause infringement on privacy [12]. In the process of data transaction, it is more vulnerable to illegal attacks and tampering, leading to the leakage of private information.

According to the Personal Information Protection Law, there are personal data right and data property right, i.e., the data subject enjoys the right to decide whether his or her data can be collected and used and also has the right to request the data transaction process to ensure the privacy of personal data, inquire about the data, and delete the data, etc. However, there are still many problems in the current law: personal information data is not clearly defined, and defending the rights is difficult. Therefore, in this case, it is especially important to realize the privacy protection under blockchain data transactions through technical means.

2.2 Blockchain Technology

Blockchain technology combines distributed networks, digital signatures and other technologies [13,19,20], which is used for recording all transactions and time of all blockchain nodes. In the public ledger, every transaction is verified by consensus of most participants, and the transaction information will be permanently written in the blockchain and will not be deleted [24]. The structure of the blockchain is shown in Figure 1.

Nodes generate the Merkle root through the Hash algorithm and store it in the block head. All transactions in the blockchain are permanently stored in the blockchain through digital signatures. Data transactions under the blockchain include five steps:

1) Generate transactions: user nodes digitally sign through private keys, generate transactions, and wait for miner nodes to pack;
2) Miner nodes randomly select transactions, verify digital signatures, and pack legitimate transactions into chunks;
3) Calculate the hash value of the block, create locally packed blocks as new blocks and broadcast them to the whole network;
4) Verify new blocks and save legal new blocks into the local block chain;
5) Write the transactions into the block chain permanently to avoid malicious tampering.

The consensus mechanism is the core of the blockchain, which is used to ensure the consistency of the blockchain. Common consensus mechanisms are as follows.

1) Proof of Work (PoW) mechanism [10]: Based on the workload, the nodes are rewarded, which has high credibility.
2) Proof-of-interest (PoS) mechanism [2]: It can effectively avoid malicious attacks.
3) Delegated proof-of-stake (DPoS) mechanism [23]: It reduces the size of committers, which is efficient.
4) Practical Byzantine fault-tolerant (PBFT) consensus algorithm [15]: It is computation-based and has high consensus efficiency.

Compared with the centralized system, the decentralized feature of blockchain makes it more advantageous in terms of privacy protection, but it does not guarantee absolute privacy. Due to the transparency of transaction information on the chain, there is a threat to the user's identity privacy and transaction privacy; therefore, privacy protection under blockchain data transactions is also a very important issue.

2.3 Homomorphic Encryption and the Paillier Algorithm

Homomorphic encryption can ensure the correctness of operations while securing the original data [11], which has a wide range of applications in privacy protection. It is assumed that there is plaintext space M, ciphertext space C, and key space K, and the encryption and decryption algorithms are E and D. If there is $P(E_k(m_1), E_k(m_2), \cdots, E_k(m_n)) = E_k(L(m_1, m_2, \cdots, m_n))$, then the encryption scheme is homomorphic. For different operations L, there are several scenarios.

For any plaintext m_i, $m_j \in M$, the corresponding ciphertext $c_i = E(m_i)$, $c_j = E(m_j)$, and $c_(i), c_j \in C$, if

1) $E(m_i + m_j) = E(m_i) \oplus E(m_j)$ or $D(E(m_i) \oplus E(m_j)) = m_i + m_j$, it is an addition homomorphism;
2) $E(m_i m_j) = E(m_i) \oplus E(m_j)$ or $D(E(m_i) \oplus E(m_j)) = m_i m_j$, it is a multiplication homomorphism;
3) $E(m_i m_j) = E(m_i) \oplus m_j$ or $D(E(m_i) \oplus m_j) = m_i m_j$, it is a mixed multiplication homomorphism.

The Paillier algorithm is an algorithm that satisfies addition homomorphism [1], and its definition is as follows. Big prime numbers p and q are randomly chosen to make $\gcd(pq, (p-1)(q-1)) = 1$, and gcd is the greatest common divisor. Then, $n = pq$ and $\lambda = lcm(p-1, q-1)$ are calculated, where lcm is the least common multiple. An integer

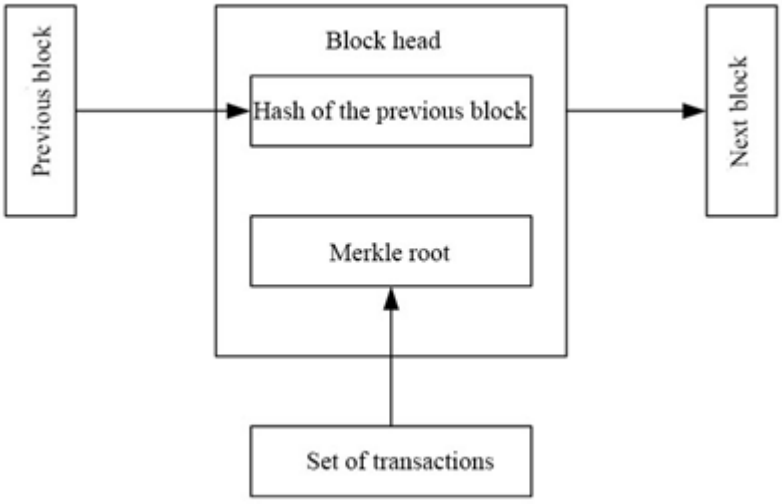

Figure 1: Blockchain structure

is randomly selected: $g \in Z_(n^2)^*$, $\mu = (L(g^\lambda \bmod n^2))^{-1}$, $L(x) = (x-1)/x$, then, public key $PK = (n, g)$ and private key $SK = (\lambda, \mu)$ are obtained. For the message to be encrypted $m(0 \le m < n)$, integer $r \in Z_(n^2)^*$ is randomly selected to ensure $\gcd(r, n) = 1$. Then, the encryption process is written as: $c \leftarrow E(m, PK)$, $E(m, PK) = g^m r^n \bmod n^2$, and the decryption process is written as: $m \leftarrow D(c, SK)$, $D(c, SK) = L(c^\lambda \bmod n^2)\mu \bmod n^2$.

According to the addition homomorphism property of the Paillier algorithm, if $c_1 \leftarrow E(m_1, PK)$, $c_2 \leftarrow E(m_2, PK)$, then:

$$\begin{aligned} c_1 \times c_2 &= g^(m_1) r_1^n \times g^(m_2) r_2^n \bmod n^2 \\ &= g^(m_1 + m_2)(r_1 r_2)^n \bmod n^2 \\ &= E(m_1 + m_2, PK). \end{aligned}$$

2.4 Privacy Protection Method Based on the Paillier Algorithm

The Paillier algorithm is applied to privacy protection under blockchain data transactions [5]. Suppose that in the blockchain, users A and B conduct transactions, the initial amounts deposited by A and B are B_0 and B_1, respectively, and the total account balance is B, the transfer amount of every transaction is $\{v_1, v_2, \cdots, v_n\}$ in the process of conducting multiple transactions $\{T_1, T_2, \cdots, T_n\}$, and the remaining amounts of A and B are $b_(0, n)$ and $b_(1, n)$ at the end of the transaction. Based on the Paillier algorithm, big prime numbers a and b are randomly selected to ensure $\gcd(ab, (a-1)(b-1)) = 1$, $n = ab$ and $\lambda = lcm(a-1, b-1)$ are calculated, and then integer $g \in Z_(n^2)^*$ is randomly chosen. $\mu = (L(g^\lambda \bmod n^2))^{-1}$ is calculated, and public key $PK = (n, g)$ and private key $SK = \lambda$ are obtained. The public-private key pair of A is (PK_0, SK_0), and the public-private key pair of B is (PK_1, SK_1).

Both A and B encrypt the balance deposited into the joint account to obtain M_1 and M_2. $E(M_1, M_2)$ is decrypted by Certificate Authority (CA), and the decryption process is $E(M_1, M_2) = (g^(B_0) r_1^n \bmod n^2)(g^(B_1) r_2^n \bmod n^2) = [g^(M_1 + M_2)(r_1 r_2)^n \bmod n^2]$.

In the first transaction, A sends the amount of the first transfer (v_1) and the account balance $(b_0, 1)$ to CA to verify T_1 after encryption. The ciphertexts after encryption are: $M_1 = g^(v_1) r_3^n \bmod n^2$, $M_2 = g^(b_0, 1) r_4^n \bmod n^2$. According to the property of addition homomorphism of the Paillier algorithm, if $m_1 \times m_2 = M_1$, then $b_0, 1 + v_1 = B_0$.

At the end of the transaction, A and B encrypt the current account balance and send it to CA for verification, and the encrypted ciphertexts are $M_3 = g^(b_(0, n)) r_5^n \bmod n^2$, $M_4 = g^(b_(1, n)) r_6^n \bmod n^2$. If $E(M_3, M_4) = B$, then the CA closes the payment channel after writing the encrypted balance to the blockchain, i.e., the transaction is completed.

3 Experiment and Analysis

From the perspective of the correctness of the algorithm, taking the first transaction as an example, CA determines whether the transaction is legitimate or not by verifying the correctness of $m_1 \times m_2 = M_1$. The proof of the process is as follows:

$$\begin{aligned} m_1 &= E(v_1, r_3) \\ &= g^(v_1) r_3^n \bmod n^2 \\ m_2 &= E(b_0, 1, r_4) \\ &= g^(b_0, 1) r_4^n \bmod n^2 \\ m_1 \times m_2 &= g^(v_1) r_3^n \bmod n^2 g^(b_0, 1) r_4^n \bmod n^2 \\ &= g^(v_1 + b_0, 1)(r_3 r_4)^n \bmod n^2 \\ &= E(v_1 + b_0, 1, r). \end{aligned}$$

Through the above calculation, b_0, $1+v_1 = B_0$ is verified; therefore, the Paillier algorithm is theoretically correct.

In order to understand the effectiveness of the privacy protection method based on the Paillier algorithm, experiments were conducted in Windows 10 environment and 8 GB memory. At the same time, the method was compared with the Zero-Knowledge Succinct Non-Interactive Argument of Knowledge (zkSNARKs) technique based on the zero-knowledge proof library libsnark [18]. The core of the zkSNARKs technique is the elliptic encryption algorithm, whose cryptosystem is different from the Paillier algorithm. Therefore, in the comparison process, the privacy protection methods under two blockchain data transactions were compared by controlling the number of keys of the Paillier algorithm and the number of constraints of libsnark.

First, the number of keys of the Paillier algorithm was set as 64 bits, 128 bits, 256 bits, 512 bits, 1,024 bits and 2,048 bits, respectively. Depending on the key sizes, the time required for the addition homomorphism of the Paillier algorithm, encryption, and decryption are shown in Table 1.

It was seen from Table 1 that the addition homomorphic time grew slowly with the increase of key size, but the change was small. The additive homomorphic time of the Paillier algorithm was 0.07 ms when the key was 64 bits and 0.17 ms when the key was 2,048 bits, indicating an increase of 0.1 ms. This suggested that the addition homomorphic time was little affected by the key, which meant that CA could complete the verification of the transaction with a low burden during the blockchain data transaction.

Then, the encryption and decryption time of the Paillier algorithm increased significantly with the increase of key size. When the key was 64 bits, the encryption time of the Paillier algorithm was 0.56 ms; when the key increased to 2,048 bits, the encryption time increased to 185.75 ms, which was about 331 times of that when the key was 64 bits. In addition, it was found that the encryption time increased mildly when the key increased from 64 bits to 512 bits, and when it increased from 1,024 bits to 2,048 bits, the encryption time showed a significant increase. Similarly, the decryption time of the algorithm increased from 0.59 ms at 64 bits to 256.12 ms at 2048 bits, which showed an increase of about 434 times. This indicated that the key change had a great impact on the encryption and decryption time, and a significant increase occurred after 1,024 bits. The decryption time was slightly longer than the encryption time. Taking 2048 bits for example, the decryption time was 256.12 ms, which was 70.37 ms longer than the encryption time. In general, the addition homomorphic time ¡ encryption time ¡ decryption time in the case of the same key.

Then, the efficiency of the zkSNARKs technique was analyzed. Different number of libsnark constraints were set, and the generation time and verification time of zero-knowledge proofs are shown in Table 2.

It was seen from Table 2 that the zkSNARKs technique took much time to generate zero-knowledge proofs, and the generation time reached 0.21 s when the number of constraints was only 1,000 and 9.87 s when the number of constraints was 50,000. Although the time required by the zkSNARKs technique for proof verification was only milliseconds, overall, its efficiency was significantly higher than that of the Paillier algorithm, which imposed a large burden on the system; therefore, the privacy protection method based on the Paillier algorithm was more applicable.

4 Discussion

With the construction of digitalization and informatization, the demand for data transactions has been growing, while at the same time, the issue of privacy protection under data transactions has gradually surfaced, reflecting the inadequacy of the current legal regulation. From the legal point of view, the provider, the buyer and the intermediary party of data transactions has a legal relationship, which is closely related to personal privacy. In June 2018, a user on the dark web sold 1 billion pieces of express data from YTO Express, which involves personal information such as the name and phone number of the user. Moreover, the behavior of collecting personal information exists in many APPs. These issues seriously threaten personal privacy; therefore, the legal regulation of data transaction is very important. For data transactions, the relevant laws are as follows.

1) According to the Network Security Law, data are generally traded with the network as the medium, then the network service provider must strictly comply with relevant laws, strengthen industry self-regulation, and keep honest and faith.

2) The Data Security Law of the People's Republic of China regulates the processing and exploitation of data, providing a new legal basis for the regulation of data transactions.

3) The Personal Information Protection Act regulates the trading of personal data to a certain extent.

According to the current legal regulation, the shortcomings are as follows.

1) The ownership of data property is still unclear. In the process of data trading, many information that has been anonymized is considered not to be private and cannot be protected according to the personal information protection law; moreover, as a kind of content that can be copied and disseminated infinitely, the subject of data rights is also difficult to be clarified.

2) There is no special legislation on data trading yet, and only some guidelines exist in other related laws, which have strong limitations.

3) There is no special supervisory department to supervise data trading.

Table 1: Efficiency analysis of the Paillier algorithm

Key size	Addition homomorphic time/ms	Encryption time/ms	Decryption time/ms
64 bits	0.07	0.56	0.59
128 bits	0.08	0.79	0.92
256 bits	0.09	2.46	2.56
512 bits	0.11	5.69	6.72
1,024 bits	0.13	27.65	35.41
2,048 bits	0.17	185.75	256.12

Table 2: Efficiency analysis of the zkSNARKs technique

Number of constraints	Zero knowledge proof generation time/s	Zero knowledge proof verification time/ms
1000	0.21	1.21
5000	1.25	5.64
10000	2.07	10.3
3 20000	3.56	11.8
9 30,000	4.69	13.5
6 40000	7.55	14.17
50,000	9.87	17.21

4) It relies too much on platform self-regulation, and the relevant regulation is not yet sound.

From the technical perspective, this paper studied the privacy protection problem under blockchain data transactions, designed a privacy protection method based on the Paillier algorithm, and verified the reliability of the method through experimental analysis, which makes some contributions to further realize the security of data transactions. In the face of the continuous development of data transactions, in addition to the technical level, the legal level should also be taken account to strengthen the supervision; therefore, this paper puts forward the following suggestion:

1) Further clarify the ownership of data property and protect the rights and interests of data owners, generators, users, practitioners, etc.;

2) Further improve the regulation of data transactions, improve the importance and recognition of data, and establish relevant regulatory sections to form a well-organized regulatory system;

3) Strengthen the legal regulation of data transactions at the level of civil, administrative, and criminal law to ensure the legitimate rights and interests of individuals while maximizing the use of data.

5 Conclusion

This paper mainly studied the privacy protection under blockchain data transactions, elaborated on the relevant legal knowledge, and then designed a method based on the Paillier algorithm to enhance the security of data transactions under blockchain. It was found through experimental analysis that the method had good correctness and also high computational efficiency, so compared with the zkSNARKs technique, it was less burdensome to the system. The proposed method can be promoted and applied in the actual blockchain to promote the security of data transactions.

References

[1] M. Alanezi, "Enhancing cloud computing security by paillier homomorphic encryption," *International Journal of Electrical and Computer Engineering*, vol. 11, no. 2, pp. 1771, 2021.

[2] K. Bala, P. D. Kaur, "A novel game theory based reliable proof-of-stake consensus mechanism for blockchain," *Transactions on Emerging Telecommunications Technologies*, vol. 33, no. 9, pp. 1-24, 2022.

[3] P. Y. Chang, M.S. Hwang, C.C. Yang, "A blockchain-based traceable certification system", in *Security with Intelligent Computing and Big-data Services*, pp. 363-369, 2018.

[4] L. Chen, Q. Fu, Y. Mu, L. Zeng, F. Rezaeibagha, M. S. Hwang, "Blockchain-based random auditor committee for integrity verification", *Future Generation Computer Systems*, vol. 131, pp. 183-193, 2022.

[5] Y. C. Chen, W. L. Wang, M. S. Hwang, "RFID authentication protocol for anti-counterfeiting and pri-

vacy protection", in *The 9th International Conference on Advanced Communication Technology*, pp. 255-259, 2007.

[6] Y. H. Chen, L. C. Huang, I. C. Lin, M. S. Hwang, "Research on the secure financial surveillance blockchain systems", *International Journal of Network Security*, vol. 22, no. 4, pp. 708-716, 2020.

[7] Y. H. Chen, L. C. Huang, I. C. Lin, M. S. Hwang, "Research on blockchain technologies in bidding systems", *International Journal of Network Security*, vol. 22, no. 6, pp. 897-904, 2020.

[8] S. Devidas, S. Rao, N. R. Rekha, "A decentralized group signature scheme for privacy protection in a blockchain," *International Journal of Applied Mathematics and Computer Science*, vol. 31, no. 2, pp. 353-364, 2021.

[9] A. Dunn, K. Hood, A. Batch, A. Driessen, "Measuring consumer spending using card transaction data: lessons from the COVID-19 pandemic," *AEA Papers and Proceedings*, vol. 111, pp. 321-325, 2021.

[10] X. Feng, J. Ma, S. Liu, Y. Miao, X. Liu, "Auto-scalable and fault-tolerant load balancing mechanism for cloud computing based on the proof-of-work election," *Science China Information Sciences*, vol. 65, no. 1, pp. 112102, 2021.

[11] C. Y. Lai, K. M. Chung, "On statistically-secure quantum homomorphic encryption," *Quantum Information & Computation*, vol. 18, no. 9 &10, pp. 785-794, 2018.

[12] F. Loukil, C. Ghedira-Guegan, K. Boukadi, A. N. Benharkat, "Towards an End-to-End IoT data privacy-preserving framework using blockchain technology," in *International Conference on Web Information Systems Engineering*, pp. 68-78, 2018.

[13] C. Y. Lin, L. C. Huang, Y. H. Chen, M. S. Hwang, "Research on security and performance of blockchain with innovation architecture technology", *International Journal of Network Security*, vol. 23, no. 1, pp. 1-8, 2021.

[14] M. A. Morris, E. L. Wilkins, M. Galazoula, S. Clark, M. Birkin, "Assessing diet in a university student population: A longitudinal food card transaction data approach," *British Journal of Nutrition*, vol. 123, no. 12, pp. 1406-1414, 2020.

[15] G. I. Navaroj, E. G. Julie, Y. H. Robinson, "Adaptive practical Byzantine fault tolerance consensus algorithm in permission blockchain network," *International Journal of Web and Grid Services*, vol. 18, no. 1, pp. 62-82, 2022.

[16] T. Nóbrega, C. Pires, D. C. Nascimento, "Blockchain-based privacy-preserving record linkage enhancing data privacy in an untrusted environment," *Information Systems*, vol. 102, no. 6, pp. 1-19, 2021.

[17] M. Onik, C. S. Kim, N. Y. Lee, J. Yang, "Privacy-aware blockchain for personal data sharing and tracking," *Open Computer Science*, vol. 9, no. 1, pp. 80-91, 2019.

[18] J. Park, H. Kim, G. Kim, J. Ryou, "Smart contract data feed framework for privacy preserving orac1e system on b1ockchain," *Computers*, vol. 10, no. 1, pp. 1-12, 2020.

[19] Z. P. Peng, M. L. Chiang, I. C. Lin, C. C. Yang, M. S. Hwang, "CBP2P: Cooperative electronic bank payment systems based on blockchain technology", *Journal of Internet Technology*, vol. 23, no. 4, pp. 683-692, 2022.

[20] S. Seybou, F. Essaf, M. Mbyamm, "Privacy protection issues in blockchain technology," *International Journal of Computer Science and Information Security*, vol. 17, no. 2, pp. 124-131, 2019.

[21] E. Sharma, "A framework of big data as service platform for access control & privacy protection using blockchain network," *Turkish Journal of Computer and Mathematics Education*, vol. 12, no. 11, pp. 476-485, 2021.

[22] M. Storm, L. Moerel, "Blockchain can both enhance and undermine compliance but is not inherently at odds with EU privacy laws," *Journal of Investment Compliance*, vol. 22, no. 2, pp. 122-132, 2021.

[23] Y. Sun, B. Yan, Y. Yao, J. Yu, "DT-DPoS: A delegated proof of stake consensus algorithm with dynamic trust," *Procedia Computer Science*, vol. 187, no. 9, pp. 371-376, 2021.

[24] S. Tsakiridi, "Blockchain and the GDPR - friends or foes?," *Privacy & Data Protection*, vol. 20, no. 3, pp. 3-6, 2020.

[25] W. Wang, L. Wang, P. Zhang, S. Xu, K. Fu, L. Song, S. Hu, "A privacy protection scheme for telemedicine diagnosis based on double blockchain," *Journal of Information Security and Applications*, vol. 61, no. 2, pp. 1-12, 2021.

Biography

Yajie Zhang is a lecturer in Law Teaching and Research Department of the Party School of the CPC Shijiazhuang Municipal Committee. She graduated from Hebei University of Economics and Business. Her research directions include civil and commercial law and administrative law.