Volume: 14 Issue: 04 Pages: 5549-5555 (2023) ISSN: 0975-0290

Private Data Protection in Social Networks Based on Blockchain

Pingshui Wang
School of Management Science and Engineering, Anhui University of Finance & Economics, China
Email: 120081049@aufe.edu.cn
Jianwen Zhu
School of Economics, Anhui University of Finance & Economics, China
Email: pshwang@163.com
Qinjuan Ma
School of Business Administration, Anhui University of Finance & Economics, China
Email: qjma321@163.com
Corresponding Author : Qinjuan Ma

Abstract: With the rapid development of big data and social networks, user data in social networks are facing huge risks of privacy leakage. It is urgent to establish a complete and effective method for protecting private data in social networks. Based on the problem of information leakage in social networks, classifies user privacy data, and constructs different privacy data protection schemes through blockchain time stamp recording data storage, hash function anonymous operation of data, asymmetric encryption and digital signature of sending information. The blockchain-based privacy data protection method in social networks can effectively solve the privacy leakage problem in social networks, and provide a reference for the research in the field of information security and social network security. This paper designs a new blockchain-based privacy data protection scheme for different privacy disclosure categories, which provides a new solution to the current privacy disclosure problem in social networks. However, the existing methods will consume a lot of computational power in the process of information interaction. The subsequent research will optimize the computational power of blockchain and try to build a better blockchain social network privacy data protection system.

Keywords: Blockchain, Social network, Privacy protection.

Date of Submission: Nov 1, 2022 Date of Acceptance: Dec 20, 2022

---

1. Introduction

With the advent of the Internet era, social network has become one of the most promising Internet application services. As a new media that allows people to write, share and comment on a given platform, social networks have gradually integrated real life and the online world. People can shorten and maintain the relationship with others through the display and exchange of information and data. As of December 31, 2021, the number of social network users in China has reached 977 million. However, in recent years, due to the development and popularization of data mining and cloud computing technology, social networks not only provide users with rich personalized services, but also have privacy leakage caused by a large amount of data storage. Users' personal information security is facing unprecedented challenges. Therefore, the design of a set of secure and efficient social network privacy data protection scheme is one of the current research focus in the field of information security.

The privacy protection of social network data is mainly aimed at the stored network data. Through some operations on these data, the data stealer cannot obtain sensitive information, cannot match the

Volume: 14 Issue: 04 Pages: 5549-5555 (2023) ISSN: 0975-0290

information with the user, and makes the data information invalid [1]. As a distributed accounting system that cannot be tampered with, blockchain can first ensure that the stored data is not tampered with, and then use the Hash function to anonymze data, and RSA algorithm to encrypt and sign data asymmetrically, which can ensure the concealment, security and integrity of data [2].Blockchain technology is regarded as the most subversive technological innovation since the invention of the Internet. It relies on the clever distributed algorithm of cryptography and mathematics. On the Internet where trust relationship cannot be established, participants can reach a consensus without the intervention of any third-party center, and the reliable transmission of trust and value can be solved at a very low cost.

Starting from the characteristics of blockchain, this paper classifies the privacy data of social network users, and builds different privacy protection methods based on each different data leakage problem using blockchain technology [3]. This paper aims to explore the privacy protection path in the background of blockchain, which is of great significance to improve the research of privacy protection scheme in social networks [4].

2. Review of relevant literature

2.1 Privacy leakage

With the popularization of computer technology and the Internet, privacy leakage has gradually become a major security vulnerability in network interaction, and personal privacy protection has attracted more and more attention and concern. Wikipedia defines privacy as the ability of an individual or group to hide themselves or their attributes in order to selectively express themselves. In the network environment, privacy is more of a symbol of information and data, which can be used to confirm the identity and characteristics of a specific individual. However, such information and data are sensitive content that individuals do not want to expose, such as patients' disease data, personal location information, financial status and so on. Literature [5] made a quantitative analysis of privacy issues in academic social networks, and concluded that personal identification information is the most easily disclosed privacy. Literature [6] divided privacy issues in social networks into three categories, namely, individual identity disclosure, connection disclosure and content disclosure.

2.2 Privacy protection technology

At present, there have been a lot of research on privacy protection. The main technical solutions in the study are as follows: By hiding or not collect the user's identity to sensitive information or data, allowing users to submit and not to reveal his identity anonymity privacy protection technology [7], the use of data and query limit the strategy of the privacy protection technology based on association rules [8] and distributed environment based on secure multi-party computation of data privacy protection service collaborative filtering method [9], The types include publishing privacy protection, storage privacy protection, mining privacy protection, access control technology and so on. In social networks, Literature [10] believes that privacy protection mainly involves some artificial operations on the original network data, such as adding, deleting or modifying a part, so that attackers cannot obtain sensitive information of users and avoid information leakage.

The above privacy protection technologies have been studied and developed for a long time, and a relatively mature technology and application framework have been formed. However, there are not many studies on the use of blockchain technology to protect privacy, and even fewer studies on the application of social networks.

2.3 Blockchain and privacy protection

Blockchain first appeared in an article titled "Bitcoin: A Peer-to-Peer Electronic Cash System" published by Satoshi Nakamoto, which explained how to establish a new set of decentralized transaction system methods that do not require any trust basis [11].Blockchain technology is a technical solution that does not rely on a third party, but stores, verifies, transmits and communicates network data through its own distributed nodes.

Volume: 14 Issue: 04 Pages: 5549-5555 (2023) ISSN: 0975-0290

Literature [12] after summarizing that blockchain technology in essence, is a decentralized database, it comes from the currency underlying derived in the core technology and infrastructure, is composed of distributed data storage mechanism, point-to-point transmission, consensus, the encryption algorithm of computer technologies such as new application mode. It can use digital summary to verify data, share and exchange data at different nodes, locations and platforms, and prevent tampering. At the same time, it can generate public and private secret keys for privacy security protection [13].

To sum up, domestic and foreign scholars' research on the privacy protection of blockchain is still very limited, mostly staying in the theoretical elaboration, or lack of specific scenario application implementation, such as social network platform. Therefore, based on blockchain technology, this paper analyzes and mines the privacy problems exposed in current social networks, and designs a privacy protection mechanism based on blockchain for the corresponding problems.

Starting from the characteristics of blockchain, this paper classifies the privacy data of social network users, and builds different privacy protection methods based on each different data leakage problem using blockchain technology. This paper aims to explore the privacy protection path in the background of blockchain, which is of great significance to improve the research of privacy protection scheme in social networks.

3. Classification of private data in social networks

Social network is a kind of Internet application service that provides social network for people. Through this network platform, anyone can communicate with friends on the network, contact with other nodes that are unfamiliar with each other, and broaden people's communication scope. Social networks allow users to express themselves by adding friends, sharing resources, making comments and other social activities according to their own wishes. In these social activities, users will provide their real personal information and leave some relevant data such as browsing history and comment history, which may lead to the disclosure of personal privacy information. Privacy in social networks mainly refers to any information or combination of information that can connect users in social networks with real individuals[14]. Based on the research of relevant literature, this paper divides the private data information involved in social networks into two categories: user basic private data and user social sharing private data.

3.1 Basic privacy data of users

At present, when users register social network accounts and use their functions, they must provide relatively complete basic information to obtain corresponding services provided by social networks. This kind of information constitutes the first kind of data information that is easy to be leaked in social networks, that is, the user's personal basic information. Most of these materials involve personal identity and characteristics closely related to the content, and in real life will not make people feel that is private information. It mainly includes the user's name, age, gender, email address, work or study unit and so on. These information is easy to be obtained by malicious users, causing unnecessary losses. The report on Internet Users' Privacy and Social Networking Site SNS Security also pointed out that after registering a personal account, users are likely to encounter risks such as mobile phone number, email number leakage and password theft.

In addition, this article will user access in the social network, such as browsing behavior of records also belong to the individual basic data for the user, through these sensitive information acquisition, to analyze the user's personal interests, habits, etc., according to the historical records to infer the current physical and financial conditions. This kind of information is defined as trajectory information, which is the reproduction of the user's life trajectory and a special kind of location information.

3.2 Users interactionshare private data

Volume: 14 Issue: 04 Pages: 5549-5555 (2023) ISSN: 0975-0290

The core of social network platform is user socialization and information sharing. In the structure of social network, each individual represents a social network node, and the two nodes are connected to each other to form an entity relationship. Each node has its own unique identifier of personal information, which is the basic personal data described in 3.1. The side information between nodes is the data shared by users in the social network.

With the continuous expansion of social networks, users weave more and more complex and extensive relationships in social network platforms. Through a node of the user can easily obtain the related node properties and relational information, such as Tencent space and Weibo, by looking at the user's friends, friends of friends can understand the message and node comments, this time can be more easily infer that user's privacy information such as social relations, friends relationship. At the same time, when users make comments and forward for their social network friends, they will leave private data such as users (starting node), users' friends (terminal node) and message information (side information), which may also be exposed to the risk of users' social relationship information.

4. Privacy protection mechanism of social networks based on blockchain

Aiming at the different types of privacy problems in the above social networks, this paper designs two different blockchain-based privacy protection mechanisms for social networks by combining the features of blockchain such as decentralization and non-tampering, Hash algorithm and RSA asymmetric encryption digital signature.

4.1 User privacy data protection based on blockchain

At present, the core of privacy protection for user basic information is to protect the relationship between private data and the user. To some extent, private data can be accessed and read by anyone, but the attacker cannot match the private data with a specific user, so it is difficult to cause user privacy leakage [15]. All social network user data will be stored in the background, each record of the centralized data is corresponding to a user entity in real life, and these data have certain privacy. Therefore, based on this, this paper uses blockchain technology to Hash function some data in the personal basic information of a single user to solve the corresponding relationship between private data and individuals.

A blockchain is like a linked list structure, with each block representing a user. The block header of a block has a hash function element information as a unique identifier, and the block chain structure is formed by successively connecting hash Pointers to the previous block. Due to the directivity of the pointer, the information stored in the blockchain is difficult to tamper with. The block body of the block stores all users' personal information data or their shared social data. All the information data is encrypted data form after Hash function transformation. The Hash function is an irreversible, unidirectional cryptosystem that transforms an input value of any length into an output value of a fixed length. It is computationally infeasible to compute the correct input value given some output values. If the input value is changed by even one letter, the subsequent converted Hash will produce a different value, thus ensuring the uniqueness and integrity of user data. In addition, the chain block to the users' personal data is divided into small blocks of data blocks of data (such as gender, age, block, etc.), each block of data is corresponding to the Hash value of the adjacent two data blocks Hash value can be merged into one large block of data (such as gender and age), the string and will form a new Hash value. The upward operation can obtain more and more data blocks and Hash values of the new level, and eventually form an upside-down tree, which is called a Merkle tree. A Merkle tree is a tree that stores hash values, increasing the security of block data.

The protection of user's personal privacy data based on blockchain is to anonymize the user's personal basic data stored in the social network through the Hash function in the blockchain and load it onto the

Volume: 14 Issue: 04 Pages: 5549-5555 (2023) ISSN: 0975-0290

blockchain.A single user contains name, gender, age and other data, which are converted into Hash values through blockchaintechnology and connected in series, such as name + gender Hash value. Each block corresponds to a user's basic data. The Hash value in the block header is the unique identifier of the user, and the Hash value in the block body is the user's basic identity information.

4.2 User sharing privacy data protection based on blockchain

Social sharing data is the data information formed by the mutual communication between users, including forwarding comments and so on. This process is similar to transactions in a blockchain, where transactions are mostly peer-to-peer, end-to-end, and user-to-user. Each user has a transaction address on behalf of himself. The input user sends transactions and information to the sender user. The process is recorded and loaded onto the block for storage. The blockchain technology is applied to the protection of sharing privacy data in social networks, and the transaction principle of blockchain is used to conceal the transaction address of each user in the form of Hash value. Every forwarding or comment between users is a transaction. At the same time, the blockchain can effectively record the whole process of forwarding from the initial user node to the end user node, and form a chained database by sorting the forwarding time according to the timestamp function. In addition, blockchain transactions use asymmetric encryption and digital signature algorithms. A user in the blockchain signs with his private key to form a new transaction, and then the user broadcasts the transaction to the blockchain system, and other nodes in the system can use the public key of the transaction user for verification. In asymmetric encryption technology, the RSA digital signature algorithm is commonly used in blockchain. After years of careful cryptanalysis, the algorithm is widely considered to be secure and reliable.

This paper tries to apply the algorithm to the protection of private data in social sharing. The application of asymmetric encryption technology and digital signature technology is the process of encrypting and decrypting data messages, which needs to use two different keys, one key A is used to encrypt data, and the other key B is used to decrypt data. Therefore, in the social network, if the user wants to share private data with A node, the user of the sender can encrypt and sign the data with his/her key A before transmitting the message, and the user of the input can decrypt the message with his/her key B after receiving the message. The RSA algorithm uses the public key to encrypt and the private key to decrypt the message. After digital signature is added, the RSA algorithm uses the private key to encrypt the message and the public key to decrypt the message, and encrypts the Hash value of the message. In this case, the input key A is called the private key, and the input key B is called the public key. The private key and public key are generated by the RSA algorithm. The message to be signed, the key, and the final signature are all represented in the form of numbers. Therefore, the message needs to be encoded into numbers before signing the message. If the information to be signed is set as M, then the signature encryption process is the result of mod N on the D power of the message. Here, D and N are the private keys of the sender. After the signature is generated, the sender can send the message and signature to the receiver. In the verification process, the input side calculates the E power of the signature and computes mod N, where E and N are the public keys of the input side. Finally, the unsigned message is obtained, and the content of the message is compared with that of the sent message. If the two are consistent, the signature verification succeeds, otherwise, the signature verification fails.

When user A shares with user B in social network, there are mainly the following steps:

(1) User A edits a piece of social network information to be sent;
(2) User A uses the Hash function algorithm to hash the social network information to be sent, and obtains the hash value H1.
(3) User A uses his own private key to sign and encrypt Hash value H1 to obtain S(H).
(4) User A sends the signature S(H) and the social

Volume: 14 Issue: 04 Pages: 5549-5555 (2023) ISSN: 0975-0290

network information to be sent to user B through the blockchain network;

(5) User B decrypts the digital signature S(H) with user A's public key to obtain H1, and hashes the received social network information again with the Hash function algorithm to obtain the Hash value H2.

(6) User B compares H1 and H2. If they are equal, the social network information is not tampered and the privacy of the information is ensured.

The sending addresses of users A and B are hidden. The traditional social network platform stores the message data on the third-party server, but after the blockchain is deployed to the social network platform, the message information sent every time is encrypted by the sender and stored on each recipient's node. Only the public key of the sender can be mastered to verify and view the data. Through the implementation of the blockchain asymmetric encryption algorithm, the information released by users in the social network will not be tampered with and the content of the message will not be learned by the third party. Only the users who participate in the delivery of the message can perform encryption and decryption verification to obtain the message.

5. Conclusion

With the rapid development and wide application of media technology, social networks have gradually become an important platform for people to exchange information and share communication. The new information transmission mechanism will continuously enlarge the information data unique to individuals. While people exchange information with others, they will gradually socialize the private domain of the network space to form the network public domain. Personal privacy in social networks is facing unprecedented challenges.

Blockchain technology is a systematic integration and innovation of multiple underlying Internet technologies, such as distributed ledger, consensus mechanism, peer-to-peer transmission, encryption algorithm and smart contract. It is characterized by decentralization, transparency, traceability, non-tamper and forgery, data security and self-establishment of credit, etc. It is a new generation of information technology after big data, cloud computing and artificial intelligence.

This paper first analyzes the different types of private data in social networks, and designs data protection schemes for different types of privacy leakage according to the characteristics of decentralization, non-tampering, Hash function, asymmetric cryptographic signature algorithm and other functions in blockchain technology, and describes the methods and processes of data protection in detail. The privacy protection scheme based on personal basic data can effectively anonymize user information and isolate the connection between user information and individual users; The privacy protection scheme based on shared data can effectively anonymize the input and output users, encrypt the signature and verify the shared information. This conclusion can provide a new solution to the current privacy leakage problem in social networks.

However, there are still many problems in the application of blockchain technology in privacy protection in social networks. First of all, data and information stored by blockchain will exist in the form of fixed hash values. How to deal with the relationship between information public query and privacy protection by blockchain in social networks is an urgent problem to be solved. Secondly, storing the data in the social network in the decentralized blockchain requires a high capacity of storage resources, which will lead to serious waste. Moreover, in the process of information interaction, the hash algorithm will also consume a lot of computational power, which needs to be continuously optimized. Finally, the problem of privacy data leakage in social networks is far more than the two mentioned in this paper. For example, the privacy leakage of location information needs to be further studied.

6. Acknowledgment

We thank the anonymous reviewers and editors for their very constructive comments. This work was supported by the National Social Science Foundation Project of China under Grant 16BTQ085.

Volume: 14 Issue: 04 Pages: 5549-5555 (2023) ISSN: 0975-0290

Reference

[1] K. Khan, W.Goodridge, 'A survey of network-based security attacks'. International Journalof Advanced Networking and Applications, 2019,10(5): 3981-3989.

[2] P. S. Wang, and Z. C. Wang, 'Research on privacy protection strategies of mobile social network users'. International Journalof Advanced Networking and Applications, 2020,12(1): 4528-4531.

[3] P. S. Wang, Z. C. Wang, and T. Chen, 'Personalized privacy protecting model in mobile social network'. Computers, Materials & Continua, 2019,59(2): 533-546.

[4] P. Zhu, J. Hu, and S.H. LV, 'Blockchain-based privacy data protection in social networks'. Information Science, 2021,39(3):94-100.

[5] C. P. Hu, R. G. Qiu, and L. L. Wang, 'Research on privacy protection of academic social network users: A case study of sciencenet blog'. Journal of Information Science,2019,38(7):667-674.

[6] R.X. Yao, and H. Li. Privacy protection in social networks. Journal of Network and Information Security,2016,2(4):33-43.

[7] G. Q. Jiang. 'Privacy-preserving algorithms and its applications in MSNS'. China: Donghua University, 2016.

[8] H. Shen, M. Zhang, and H. Wang, 'A lightweight privacy-preserving fair meeting location determination scheme'. IEEE Internet of Things Journal, 2020,7(4): 3083-3093.

[9] Z.Q. Feng, 'Research on Problems of Network privacy under the age of big data: From protection of the rights to personal choice - A Case Study of mobile social networking users'. China: Jilin University, 2016.

[10] Y. Sun, 'Survey of data privacy protection technology in social networks'. Information and Communications, 2019(1):180-181.

[11] S. Nakamoto, 'Bitcoin: A peer-to-peer electronic cash system'. https://bitcoin.org/bitcoin.pdf,2020-03-01.

[12] Y. Yuan, and F.Y. Wang, 'Development status and prospect of blockchain technology'. ActaAutomatica Sinica,2016,42(4):481-494.

[13] F.C. Kong,'Construction and management of open access resources based on blockchain'. Information Theory and Practice,2019,(5):153-158.

[14] Y. J. Ren, Y.Leng, and F. J. Zhu, 'Data storage mechanism based on blockchain with privacy protection in wireless body area network'. Sensors, 2019,19(10):2395-2406.

[15] L.S. Huang, M. M. Tian, and H. Huang, 'Preserving privacy in big data: a survey from the cryptographic perspective'. Journal of software, 2015,26 (4): 946-952.