Blockchain-Enhanced Privacy Protection in Social Network Research and Applications

Lin Li

College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, 211106, China

ll052250102@nuaa.edu.cn

Abstract. While social networks offer users unprecedented convenience and connectivity, they inadvertently risk exposing sensitive user data. To address the juxtaposition of utility and vulnerability, this paper delves into the potential of blockchain technology as a tool for enhancing privacy protection within social networks. By first critiquing the limitations of current protective measures, it becomes clear where traditional methods fall short in counteracting common avenues of privacy breaches. This paves the way for the integration of blockchain technology into the privacy infrastructure of social networks. The paper elucidates the core cryptographic principles underpinning blockchain and, by drawing from real-world applications such as ShoCard, underscores the advantages that this nascent technology brings to the fore. Despite the evident merits, challenges remain, especially as the technology becomes a potential target for advanced cyber-attacks. The incorporation of elements like smart contracts, digital signatures, and decentralized storage manifests blockchain's promise in building a robust digital fortress, safeguarding user privacy. However, as the technology is still in its evolutionary phase, ongoing research must address vulnerabilities and ensure a holistic defense against potential threats.

Keywords: Blockchain technology, Privacy protection, social network, Distributed storage, Smart contracts.

1. Introduction

In the digital age, the meteoric rise of social networks has seamlessly integrated into daily lives, providing an expansive platform for idea exchange, social engagement, and even bolstering business efficiency and financial transactions. Global data suggests a staggering 5 billion individuals, representing 60.6% of the global populace, are active on these platforms. Yet, the symbiotic relationship between websites and gaps in current encryption paradigms renders user data vulnerable to breaches, making privacy a paramount concern. Given these challenges, there's a pressing need for a robust solution, and blockchain technology presents itself as a formidable contender for this task.

Blockchain's prowess extends beyond mere buzzwords, cementing its utility in authentication, document verification, and safeguarding assets. It's notable that over 10 million global users partake in Bitcoin transactions, with a daily transactional volume eclipsing $150 million [1]. Contemporary financial systems, too, are eyeing transformative opportunities offered by blockchain, with its decentralized ethos and resource-sharing capabilities poised to slash data upkeep expenses [2].

Drawing on blockchain's foundational strengths, such as immutable distributed storage and smart contracts, this study integrates the concepts of Merkle trees and Hash functions for bolstering personal data protection. Moreover, the digital signatures employed underscore secure data transmission. A detailed exploration of the ShoCard App serves as a practical illustration, shedding light on both the merits and potential vulnerabilities inherent in blockchain-oriented privacy measures.

2. The challenge of privacy protection on social networks

2.1. Common privacy leakage modes

2.1.1 Open API

In the quest to cater to diverse user requirements, numerous service-centric platforms, including the likes of APISpace, Facebook, and Swagger, have pivoted towards encapsulating their web functionalities as APIs. By doing so, they offer the developer community the flexibility to tailor and augment these functionalities, thereby enriching their primary platforms. However, such openness invariably escalates the risk of exposing users' sensitive data to third-party sites [3]. Consequently, there's an imperative need for API platforms to strike a balance between fostering innovation and ensuring rigorous privacy safeguards [4]. This paper delves into the potential of blockchain-centric solutions as a robust means to bolster data protection within API interactions. The adoption of such techniques posits that, even in instances where potential malefactors manage to capture these shared attributes, linking such data to individual users remains an elusive endeavor.

2.1.2 APP permission abuse

After obtaining the user's authorized permissions, third-party applications will go beyond the scope of their normal functions and conduct unauthorized or abused users' personal information. For example, after obtaining user permissions, third-party applications fail to make reasonable use in accordance with the scope and purpose agreed by the user, and collect, store, and use the user's information beyond their authority; Using users' personal information for non-consensual business activities such as advertising without authorization; Obtain sensitive user information through location information or address book.

2.2. Limitations of Traditional Privacy Protection Methods

2.2.1 K-anonymity algorithm

While many websites resort to simple anonymization of user data, this can still leave users vulnerable. There is a recognized risk where attackers might cross-reference different data sources, identifying and extracting user attributes in what's termed a 'link attack.' To counter such threats, Samarati and Sweeney proposed a robust privacy masking technique in 1998: the K-anonymity algorithm [5]. Central to this algorithm is the concept of generalization. By ensuring each published tuple's attributes match with at least k-1 other tuples, the algorithm dilutes the exactness of identification to a probability of 1/k, thereby offering a safeguard against link attacks.

However, K-anonymity is not impervious to all threats. Several notable vulnerabilities include:

Homogeneous Attack: Here, the crux is not the breach of user identity but the exposure of their private data. Even if attackers can't pinpoint a specific user's identity, if they know the target user figures within a data table, the K-anonymity model might inadvertently expose sensitive attributes, leaving users at risk.

Background Attacks: These leverage an attacker's pre-existing knowledge. Such background information could be detailed insights the attacker has on a user or even widely accepted common knowledge. As Lu Hanxiao observed, the challenge lies in the unpredictability of the background information that can compromise K-anonymous protocols. Currently, there's a scarcity of algorithms robust enough to counter such knowledge-driven attacks, making them particularly formidable.

2.2.2 Data perturbation

Random scrambling, data collation, multi-party secure computing, data masking, and data generalization are the main data disturbance technologies [6]. As a technology that balances data mining accuracy and user privacy protection, the main idea of data perturbation is to add noise perturbation to the original database or modify, generalize, map and other operations on some sensitive attributes to prevent user privacy leakage.

But it still faces the following challenges.

The contradiction between privacy protection and data accuracy.

Overprotection during noise processing can lead to a loss of accuracy and integrity of the original data. For example, in the financial industry, if the information of customers or enterprises is overprotected, there is no way to accurately quantify it. On the contrary，for medical diagnoses, patient privacy can be compromised if not properly protected.

Inevitable insider attacks.

Although data scrubbing technology largely avoids the risk of user privacy leaking to the outside, insiders may leak or abuse the data, and even use known noise algorithms to reverse exit user-sensitive attributes.

Also, for high-dimensional datasets, the raw data can be analyzed and inferred sensitive information by using correlations and statistical characteristics between data.

For example, an attacker is able to obtain sensitive user attributes by querying and comparing perturbed datasets multiple times.

3. Application of blockchain technology in social network privacy protection

3.1. Principles and mechanisms of Blockchain Technology in Social Networks

3.1.1 Distributed storage

In contrast to conventional centralized systems, blockchain networks distribute private data across multiple independent nodes in segmented blocks, eliminating dependence on centralized third-party entities. This design intrinsically curtails the potential for a singular entity to dominate or manipulate the network, while simultaneously mitigating risks associated with single point failures.

(1) The replication of data across numerous nodes ensures that even in the event of a node failure, data remains accessible and recoverable from other nodes. This exemplifies the inherent reliability of blockchain technology.

(2) Additionally, blockchain's scalability is commendable. As the network's storage capacity can be effortlessly augmented by incorporating new nodes, the system can manage ever-growing private data volumes without requiring significant structural revamps.

(3) Perhaps most crucially, within this decentralized paradigm, a public ledger facilitates data security consensus [7]. The inherent transparency of the blockchain network means that vast amounts of its information are open to scrutiny. This empowers all users to freely validate information via the consensus mechanism and to inscribe content within individual blocks. For instance, Satoshi Nakamoto delineated the application of the proof-of-work mechanism to Bitcoin in his seminal White Paper. In a similar vein, PoS, DPoS, PBFT, and POOL have been integrated into various systems as consensus mechanism algorithms.

3.1.2 Smart contracts

Blockchain smart contract is a kind of code stored in blockchain. Without the third party’s participation，this code can execute automatically when some rules are satisfied.

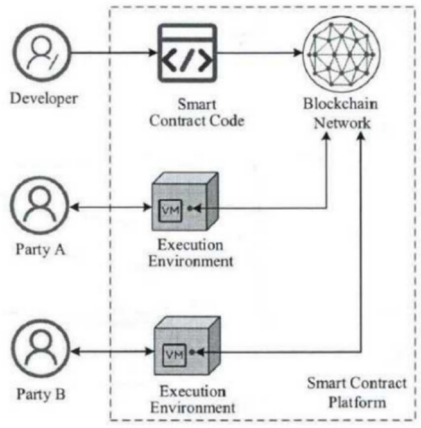


Fig. 1 Smart contract platform in blockchain [8].

When applying blockchain smart contract technology to social network，the user can control the access to their private data by making relevant rules, decide what data can be shared and how the data can be shared or disseminated. As shown in Fig 1. As is manifested in the chart above, the developer first creates smart contract programs based on user-defined rules and then stores the code to the block chain network. After that, party A and B commmence interactive execution in the smart contract platform. Thanks to the user-defined contract, some actions of A and B may be restricted, which greatly reduces the risk of privacy leakage.

3.1.3 Hash function encryption

In the Hash function, the length of the input is arbitrary while the length of the output is a string of fixed length. Once the input changes a little, the output will change a great deal, just as is shown in Fig 2. Because of unidirectionality and collision resistance of Hash function, it can be hard to infer the message from ciphertext.



Fig. 2 A hash function encryption example (Photo/Picture credit: Original).

3.2. Specific methods of blockchain protection technology

3.2.1 Personal Privacy Protection

Merkle tree combined with Hash is used for privacy protection.

As is shown in the Fig 3, users' privacy is firstly divided into several sections and store the corresponding Hash value into different blocks, in which way, leaf nodes are formed. Similarly, child nodes and root nodes can be acquired. By doing this, the user's private information is hashed, anonymized, and concatenated, resulting in a unique identifier for them.

Any falsification in the data block will result in difference in the root nodes, securing the safety of privacy a great deal.

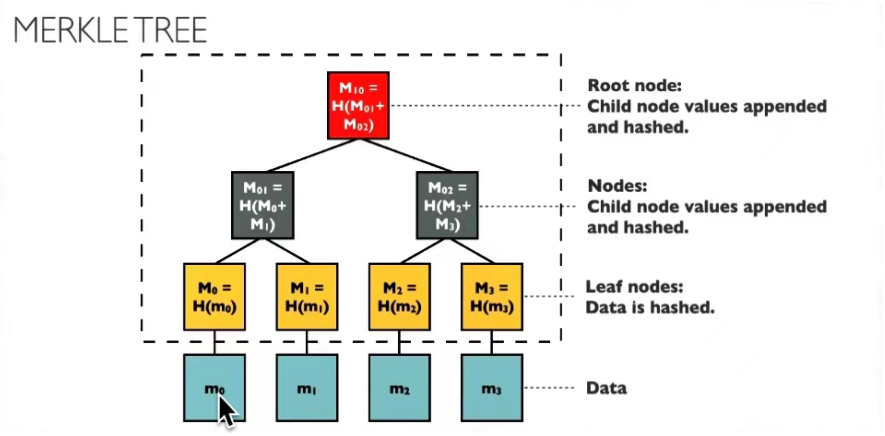


Fig. 3 Merck tree encryption principle (Photo/Picture credit: Original).

3.2.2 Secure privacy data transmission

When the user share message or privacy data through the internet, it's possible that the attackers intercept and the content. To deal with such cases, it's wise to produce digital signature used for authentication of the sender according to asymmetric cryptography like RSA.

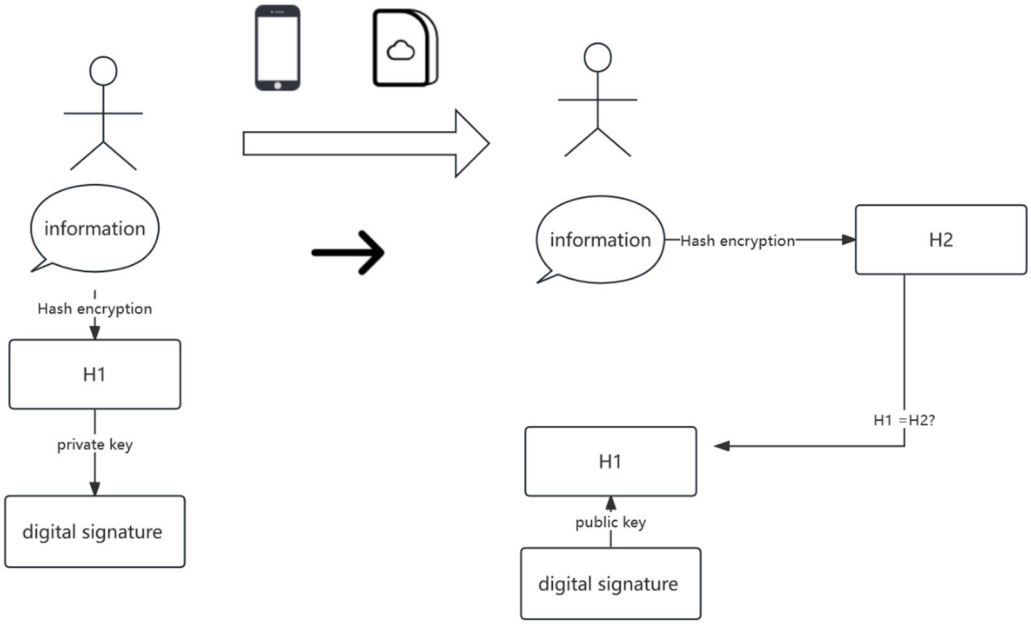


Fig. 4 Privacy protection process based on digital signature [9].

As is shown in the Fig 4, if users want to transform privacy data with each other, the sender should first generate a pair of keys, and generate the summary of message using HASH. Then, producing signature with private key. When receiving the message, the receiver uses public key to verify whether the signature was tampered.

In this way, though attacker want to tamper the content, he can't forget the signature because of the lack of corresponding key.

4. Case analysis of social network privacy protection based on blockchain

4.1. Authentication process of ShoCard app

ShoCard is provider of digital identity solutions. Its purpose is to use blockchain technology to create and manage personally identifiable information, solving problems in traditional identity

verification such as data breaches and identity theft. Its earliest application is the cooperation with SITA. The following is the certification process of ShoCard app.

Initially, when downloading this app for the first time, the system will automatically generate a ShoCardID to store user's privacy data that has been hashed. Then, digital signature will be generated by private key. Eventually, the hash value and signature will be recorded in the server for the later verification.

For the first authentication, the verifier only needs to scan user's QR code to acquire users' identity information to verify whether users' ShoCard is matched with the recording. A certification will be given to the user if the verification is successful, and it will be encrypted into the blockchain. Therefore, for the next verification, the verifier only needs to check the certification instead of privacy data.

4.2. Advantages

In addition to the high security of digital signatures, hash functions and encryption algorithms in blockchain technology, ShoCard also has the following advantages.

(1) The verifier can avoid establishing a database for password storage by not requiring users to sign up for an account, which reduces the risk of attackers' invasion.

(2) Instead of carrying proof of entity identity，users finish authentication and KYC process only by simple operating in the network, enhancing the efficiency to some degree.

(3) Based on blockchain high scalability，ShoCard can make different authentication scheme according to various business needs so that it can be applied to systems with varying scale and complexity. One example is that ShoCard supply AML and KYC function.

5. Challenges and potential attacks of blockchain technology in practical applications

5.1. Quantum computing attack

Asymmetric encryption and digital signature with a high degree of security and Hash function with unidirectionality and collision resistance are used in blockchain technology. Although these algorithms are secure in terms of current technology, their safety may be greatly affected by the development of Quantum computing [10]. For instance, the following table is a post-quantum cryptography report by NIST. It outlines the threats that large-scale quantum computers pose to the security of certain cryptographic algorithms [11, 12].

Table 1. Impact of quantum computing on common cryptogram [13].

Password type	Security implications
SHA-2, SHA-3	Attack difficulty halved
AES	Attack difficulty halved
DSA	Broken
RSA	Broken
ECDSA, ECDH	Broken

5.2. Smart contract vulnerabilities

In social network, smart contract is a program designed according to some rules Deviations in the understanding of the programming language or programming flaws may cause contract vulnerabilities. Due to the irreversibility of the blockchain, once the program is deployed into the blockchain, the contract vulnerabilities can no longer be changed, making it easy for attackers to exploit these vulnerabilities. For example, a DAO attack on Ethereum was initiated by an attacker because of the developer's misunderstanding of the calling primitive in Solidity, which resulted in losses of over 60 million US dollars [14].

6. Conclusion

This study suggests that traditional protective measures may fall short in guarding against privacy breaches stemming from open API misuse and data exploitation. In contrast, blockchain technology emerges as a more secure alternative to safeguard user privacy. The inherent encryption mechanisms in blockchain, which distribute sensitive user attributes across different blocks using the Merkle Tree format, substantially bolster storage security for personal data. Furthermore, the adoption of digital signature technology assures data integrity during transmissions, ensuring its resistance to tampering. The decentralized nature of blockchain, combined with the implementation of smart contracts, negates the dominance of a single entity over user data, further enhancing data protection.

Such findings underscore the importance of vigilant privacy preservation practices during online engagements, mitigating risks of personal data misuse. Moreover, this insight serves as a beacon for developers, presenting fresh avenues for refining the performance of social platforms. Yet, when it comes to addressing potential vulnerabilities and threats poised against blockchain, this study acknowledges the absence of comprehensive countermeasures. Future endeavors will be directed towards devising strategies to counteract these challenges, ultimately aiming to establish a more fortified privacy protection framework leveraging blockchain technology.

References

[1] ChrisB, & AdamW. (2017, June 10). Bitcoin ringing the bell for a new asset class. Retrieved.

[2] Zhu, L., Gao, F., Shen, M., Li, Y., Zheng, B., Mao, H., & Wu, Z. (2017). A Comprehensive Review on Blockchain Privacy Protection. Computer Research and Development, 10, 2170-2186.

[3] Zhu, P., Hu, J., Lü, S., & Qi, Y. (2021). Research on Privacy Data Protection Methods of Social Networks Based on Blockchain. Information Science, 03, 94-100.

[4] He, L., Hong, L., Zhou, J., Chen, P., Zhao, X., Xie, Y., & Liu, G. (2018). Comprehensive Review on Social Network Privacy Security. Information Technology, 05, 153-159.

[5] Lü, H. (2018). Privacy Protection Research of the Same Sensitive Value Data Table Based on k-Anonymity (Unpublished master's thesis). Hubei University for Nationalities.

[6] Liu, H. (2014). Privacy Protection Theory and Application Research Based on Complete Spatial Data Disturbance (Unpublished doctoral dissertation). Donghua University.

[7] Gao, Y. (2021). Research on Key Technologies of Blockchain Transaction Security and Privacy Protection (Unpublished doctoral dissertation). Beijing University of Posts and Telecommunications.

[8] Yin, H. (2023). Privacy Protection Technology Research for Blockchain Smart Contracts (Unpublished doctoral dissertation). Beijing University of Technology.

[9] Bernstein, D. J. (2009). Introduction to Post-quantum Cryptography. Berlin: Springer-Verlag, 1-14.

[10] Chen, L., Jordan, S., Liu, Y. K., Moody, D., Peralta, R. C., & Perlner, R. A. (2018, October 5). Report on post-quantum cryptography.

[11] Han, X., Yuan, Y., & Wang, F. (2019). Blockchain Security Issues: Current Research and Prospects. Acta Automatica Sinica, 01, 206-225.

[12] Atzei, N., Bartoletti, M., & Cimoli, T. (2017). A survey of attack on Ethereum smart contracts SoK. In Proceedings of the 6th International Conference on Principles of Security and Trust (pp. 164-186). Berlin, Germany: Springer-Verlag.

[13] Daian, P. (2020, December 29). Analysis of the DAO exploit.

[14] Savelyev, A. (2017). Contract law 2.0: Smart contracts as the beginning of the end of classic contract law. Information & Communications Technology Law, 26(2), 116-134.