Research on Private Data Security Model Based on Blockchain and Function Encryption

Chen Zhang

City University of Hefei, Hefei, China

Email: zc@cuhf.edu.cn

Abstract. With the advent of the digital age, privacy data security has become an important research field. The private data security model based on blockchain and functional encryption is an emerging solution that combines the decentralized nature of blockchain with the security of functional encryption to protect the security and privacy of personal private data. This article begins with a review of the risks and implications of private data breaches and introduces the basic concepts of blockchain and functional encryption. Then, aiming at the problems existing in the existing private data security model, a private data security model based on blockchain and function encryption is proposed. The blockchain-based private data security model enables decentralized data storage and verification, ensuring data integrity and traceability. At the same time, functional encryption technology provides a higher level of data protection, ensuring that only authorized users can decrypt and access data. By combining blockchain and functional encryption, the model solves the problems of data leakage and tampering existing in the traditional private data security model, providing higher data security and privacy protection. Finally, this paper evaluates and analyzes the private data security model based on blockchain and functional encryption. The results show that the model has high safety and reliability, and has wide potential in practical applications. The private data security model based on blockchain and function encryption is an innovative solution that can effectively protect the security and privacy of personal private data. However, further research and practice is still needed to refine the model and address its possible challenges and limitations.

Keywords: blockchain; Function encryption; Private data; Security protection

1 Introduction

With the rapid development of information technology and the popularization of the Internet, a large amount of personal privacy data has been widely collected and used. However, the abuse and leakage of private data is also becoming more and more serious, which brings great challenges to personal privacy and data security. In order to solve this problem, academia and industry began to study and explore various private data security models.[1] The private data security model based on blockchain

C. Chen et al. (eds.), *Proceedings of the 3rd International Conference on Digital Economy and Computer Application (DECA 2023)*, Atlantis Highlights in Computer Sciences 17,
https://doi.org/10.2991/978-94-6463-304-7_50

and functional encryption is an emerging solution that has emerged in recent years. As a decentralized distributed ledger technology, blockchain has unique advantages in protecting the integrity and traceability of data. Functional encryption, as an encryption technology, ensures that only authorized users can decrypt and access data, providing a higher level of data protection.[2] In this context, this paper aims to study the private data security model based on blockchain and functional encryption, and explore its feasibility and effectiveness in practical applications. The study will review existing privacy data breach risks and implications, and introduce the fundamentals and concepts of blockchain and functional encryption. The study will then elaborate on the blockchain-based private data security model and functional encryption technology, and analyze its advantages and disadvantages. In addition, this paper will also evaluate and analyze the private data security model based on blockchain and functional encryption, and explore its application prospects and limitations in practical scenarios. The security, scalability and performance of the model will be discussed, and some suggestions for improvement and optimization will be put forward. [3]Finally, through the research and analysis of this paper, the research hopes to provide a comprehensive understanding of the private data security model based on blockchain and function encryption methods and tools, to provide new ideas and solutions for the protection of personal privacy data. At the same time, the research also hopes to further promote the research and development in this field, and promote the further improvement and popularization of private data security.

2 Relevant theoretical basis

2.1 Design objective

This model aims to achieve the following design objectives.

Data sharing controlled by the data owner. Aiming at the problems of data easy to copy, difficult to control, difficult to confirm rights and so on, this model aims to make the data owner have complete control over the data. The model in this paper has the characteristics of non-exposure of original data, tamper-proof of cloud storage data, access control and secure sharing[4].

Privacy protection. In order to realize data sharing under the premise of protecting the original data, the model in this paper aims to combine cryptography and blockchain technology to achieve "invisible data" and verifiable results, so as to ensure the security of private data.

Reliability verification. Zero-knowledge proof technology is used to ensure the reliability of data processing results, eliminate the untrustworthy risk of data processing caused by the invisibility of original data, and protect the rights and interests of data users.

2.2 Overall architecture

This paper combines blockchain and cloud server to achieve a hybrid storage architecture of on-chain storage and off-chain storage, and uses functional encryption,

zero-knowledge proof and other technologies to achieve privacy protection and secure data sharing of verifiable results. The specific architecture is shown in Figure 1. This model includes the following key roles: trusted authority, blockchain, cloud service provider, data owner, and data consumer. The roles are described as follows.

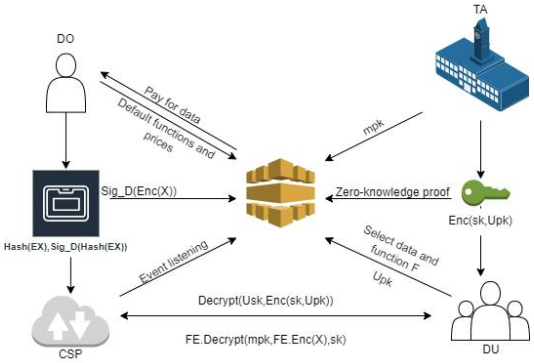

Fig. 1. Overall architecture and interaction between roles

Trusted authority: responsible for global initialization Settings. The initialization parameter Settings of function encryption and zero-knowledge proof used in this model are generated by TA, and TA is responsible for generating private keys in function encryption. [5]This article assumes that TA is completely trustworthy.

The blockchain is responsible for storing encrypted data summaries, verifying the correctness of zero-knowledge proofs, publishing pricing and functions for DO, receiving data requests from DU, and finally automated payment flow, which are implemented through smart contracts that contain related events for the delivery of messages.

Cloud service provider: is responsible for storing the ciphertext uploaded by the data owner and providing the ciphertext to the purchased data user.

Data owner: Responsible for storing data, encrypting data, specifying usage functions, and pricing.

Data users: financial institutions, technology companies, government agencies and other users with private data use needs.

The model in this paper involves three cryptographic algorithms, namely, function encryption, zero-knowledge proof and asymmetric encryption. The symbols related to function encryption all start with FE, which will be discussed in detail in Section 4.1. The generation of zero-knowledge proofs is discussed in detail in Section 4.2. Here is the asymmetric encryption section, i.e. Hash(Ex), Enc(sk,Upk), Sig_D(Hash(Ex)), Decrypt(Usk,Enc(sk,Upk)) in Figure 1. Hash(Ex) Indicates that the function encrypted ciphertext Ex is hashed. Si g _D(Hash(Ex)) indicates the digital signature of the ciphertext hash value to prevent data tampering. Enc(sk,Upk) : The user's public key Upk is used to encrypt the derived key sk. Decrypt(Usk,Enc(sk,Upk)) : The user private key Usk is used to decrypt the encrypted sk and obtain the original sk.

2.3 Operation flow

The detailed operation process of the model in this paper is as follows.

Step 1: Global initialization. First, the TA sets the global security parameters, generates the main public key and the main private key of the system, and exposes the main public key of the system in the blockchain.

Step 2: Encrypt and upload data. DO obtains the main public key from the chain and runs the function encryption algorithm to obtain the ciphertext Ct. In order to avoid overloading blockchain nodes due to excessive data, this model adopts the method of on-chain and off-chain mixed storage. DO performs hash digest and digital signature on ciphertext Ct, and then uploates the hash digest and signature to CSP together with ciphertext Ct, and sends the hash digest and signature to the blockchain for storage through transactions, thereby improving the performance and scalability of the model and reducing the on-chain space overhead. At the same time, since CSP is assumed to be honest in this model, that is, CSP will honestly execute the user's instructions, but may be curious and malicious about the user's data, this storage scheme can also avoid CSP tampering with data. The CSP listens for on-chain events to ensure that only legitimate DUs have access to encrypted data. Then, DO needs to specify the functions that can act on the encrypted data and the corresponding price. Since the model in this paper adopts inner product function encryption, in addition to the DO preset function, DU can upload its own encrypted data processing functions, including weighted sum, average value, and even simple machine learning models.

Step 3: Data access. When DU wants to get data, it does not get the raw data directly, but the result of the function processing of the encrypted data, that is, F(x). The sharing of original data in traditional data sharing is changed to the sharing of calculation results of data to realize the privacy protection and sharing of data. DU can select the default function on the chain or provide its own function F, along with a personal public key Upk, which is used to encrypt the final result and finally invoke the smart contract through the transaction.

Step 4: Key generation. After monitoring the on-chain event, TA obtains the function F purchased by DU and the personal public key Upk from the event, runs the key generation algorithm, processes F and encrypted data, and generates sk. The Proof of sk validity is generated by zero-knowledge Proof algorithm, and the proof is uploaded to the chain for verification. After verifying the validity of Proof, the smart contract sends the funds stored in DU in Step 3 to DO. If the verification is not passed or the TA does not produce Proof within the specified time, the transaction will be cancelled and the TA will be penalized. Finally, the sk encrypted using Upk is recorded in the contract, and the related event is triggered to notify DU to obtain the sk.

Step 5: Decrypt. DU first uses Usk to decrypt the UPK-encrypted sk obtained on-chain. Then the ciphertext is obtained from the CSP and compared with the hash value on the chain to verify whether the file has been tampered with, and finally run a decryption algorithm to obtain F(x).

3 Safety analysis and experiment

In this paper, a local 5-node private network was built using Ethereum. The host computers of the nodes were all VMware virtual machines (Ubuntu18.04), which were configured with Intel Core i7 CPU and 32GB memory. In order to facilitate the experiment, PoA single-node packaging block (i.e. block out) is selected, and other parameters, such as block size and gasLimit, are consistent with Rinkeby Ethereum testnet. Using Solidity language to write smart contract Market, DO chain data release, DU chain data purchase, zero knowledge proof chain verification and so on. The Rust language is used to implement the algorithmic parts that are highly computationally complex and computationally intensive and not suitable for blockchain, such as function encryption and zero-knowledge proof generation. The off-chain data is transferred to the on-chain through the form of transactions, and the on-chain transactions pass the transaction event defined in the smart contract, that is, the corresponding action will trigger a pre-designed event to notify the off-chain CSP and other relevant parties. In this experiment, local nodes were used to simulate CSP, the node configuration was the same as the host of the above nodes, and the network bandwidth was 1000MB.

Table 1. Gas and time consumed by deployment contracts and individual functions

name	Gas	Time /s
Market(Deployed)	2 031 046	0.42
SetFnAndPrice	40 091	0.01
ChooseCt	49 210	0.01
UploadFn(para_size=3.8 KB)	930 930	0.22
VerifyZk	1 014 800	0.28

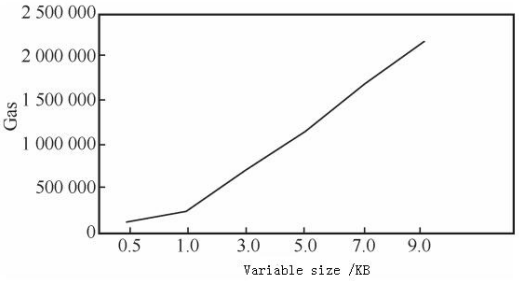

Fig. 2. The change of Gas consumed by uploading a custom calculation function with the size of the desired variable

In the initial stage of the experiment, after the network block is stable, the Market is deployed in the local private network through Remix, the ZoKrates component is adopted for zero-knowledge proof, and the Gas and time consumed by the deployment contract and each function are measured, as shown in Table 1. Comparing the ChooseCt and UploadFn functions, it can be seen that if DU uses the

function preset by DO, its Gas consumption is only 5% of that of the uploaded custom function. The Gas consumed by the uploaded custom function varies with the size of the desired variable, as shown in Figure 2. The main reason Gas is higher when using a custom function is that as the size of the desired variable increases, so does the size of the argument passed in when the UploadFn function is called. This function writes parameters to the chain and notifies the cloud service provider through transaction events, and in Ethereum, storage costs are much higher than computing costs, so DU can choose pre-defined functions whenever possible to reduce the cost of using the data.

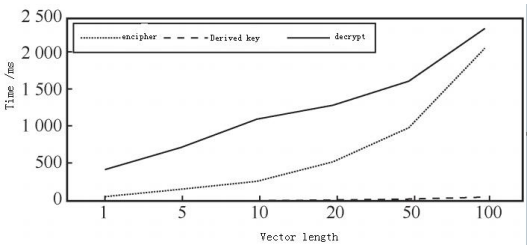

Fig. 3. The time consumed by the function encryption operation varies with the vector length

In order to determine the practicability of the research algorithm, the shared data already in the blockchain in the Byzantine fault-tolerant mechanism is used. Through comparison and analysis, the advantages and disadvantages of the proposed model and the existing algorithms are verified, and the results are shown in Table 2.

Table 2. Comparison with other models

property	POOL	POS	POW	Textual model
Verification speed	10min	<100s	<100s	<10s
Byzantine fault tolerance	50%	50%	50%	64%
Throughput (TPS)	<100	<1000	<1000	>10K

The proposed model has obvious advantages in the comparison of consensus algorithms with existing experimental data. It is obviously superior to other models in terms of verification speed, Byzantine fault tolerance and throughput, and has certain advanced nature.

4 Conclusions

The performance of the inner product function encryption algorithm will be affected by the length of the desired vector, thus affecting the performance of the entire data transaction. Therefore, this paper conducted performance tests on vectors of different lengths, and the experimental environment was the same as that of the Ethereum node host, and the experimental results obtained were shown in Figure 3. As can be seen

from FIG. 3, with the increase of vector length, the time required for encryption and decryption increases relatively, but the time required for deriving the key remains at a low level, which also proves the correctness of the proposed model for implementing encryption and decryption off-chain.

Acknowledgment

Project No.: Scientific Research Project of Anhui University (Natural Science) (Project No.: 2022AH052473)

Anhui Provincial Education Department 2022 University Outstanding Young Talents Support Project (Project No.: gxyq2022185)

Reference

1. CHEN J C, XUE Y Z. Bootstrapping a blockchain based ecosystem for big data exchange[C]//Proceedings of 2017 IEEE International Congress on Big Data (BigData Congress). Piscataway: IEEE Press, 2017: 460-463.
2. LIANG X P, ZHAO J, SHETTY S, et al. Integrating blockchain for data sharing and collaboration in mobile healthcare applications[C]//Proceedings of 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications. Piscataway: IEEE Press, 2017: 1-5.
3. THE O D O U L I A, A R A K L I O T I S S, MOSCHOU K, et al. On the design of a bl o ckch a i n-b a s e d sy st e m t o facilitate healthcare data sharing[C]//P r o c e e d i n g s o f 2 0 1 8 17 t h I E E E International Conference on Trust, Security and Privacy In Computing and Communications/12th IEEE International Conference on Big Data Science and Engineering. Piscataway: IEEE Press, 2018: 1374-1379.
4. GORDON W J, CATALINI C. Blockchain technology for healthcare: facilitating the tra nsitiontopatient-driven interoperability[J]. Computational and Structural Biotechnology Journal, 2018, 16: 224-230.
5. Menelaos K,Magda F. The cost of privacy on blockchain: A study on sealed-bid auctions[J]. Blockchain: Research and Applications,2023,4(3).