Novel Blockchain-Based Privacy Protection for Smart Home

Lifeng Pan[1], Shuguang Zhou[2]

lifengpan@ln.hk[1], 3200104014@zju.edu.cn[2]

Department of Computing and Decision Sciences, Lingnan University, 8 Castle Peak Road, Hongkong, China[1]
Department of Mathematical Sciences, Zhejiang University, No. 866, Yuhangtang Road, Hangzhou, China[2]

Abstract: Smart home data management systems are starting to be put into use in many smart home companies. To protect the private data of users in these smart home data management systems, the paper proposes Smart Home Guard, a blockchain-based system for recording smart home data operations, which creates smart contracts deployed on the Ethereum private blockchain based on Solidity. Users can customize data access policies, authorize and monitor queries on private data. By taking advantage of the immutability of the blockchain, the query records of user data are monitored and traced back to prevent user data from being stolen or maliciously tampered with. Smart Home Guard is designed to help regulators track the use of smart home data. At the same time, it allows arbitrators to verify the authenticity of evidence in the event of a privacy breach dispute.

Keywords: Blockchain, Smart Home, User Privacy, Data Protection.

1 INTRODUCTION

With the extensive use of smart home devices, their imperfect security protection capability is highly susceptible to causing large-scale leakage of user privacy data. As shown in Figure 1, the device security analysis of smart home is mainly lies in the perception layer, network layer and application layer [5]. The purpose of this paper is to design a blockchain-based smart home system to record user data operation - Smart Home Guard, by using the tamper-evident nature of blockchain to query the user data. By using the immutability blockchain, the records are monitored and traced to prevent user data from being maliciously tampered and stolen, and to achieve the protection of user data security in the information interaction stage, i.e., the network layer.

According to access and management authorities, blockchain is divided into three categories: public blockchain, consortium blockchain and private blockchain. In this paper, private blockchain is used to build Smart Home Guard for the following main reasons.

1) Smart home system does not need too many nodes, and private chain retains the advantages of node communication while avoiding network attacks brought by too many communication nodes to a great extent.

BDEDM 2023, January 06-08, Changsha, People's Republic of China

DOI 10.4108/eai.6-1-2023.2330276

2) The data on the public chain is highly transparent and not suitable for privacy protection, while the data on the private chain is highly private and users can even restrict and authorize read-and-write access by modifying the smart contract.

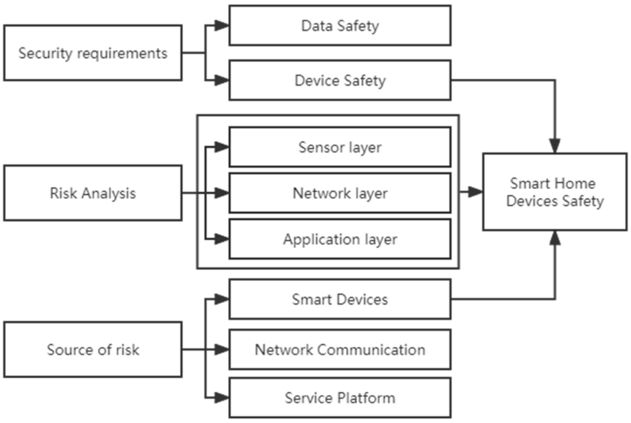

Figure 1: Smart home security analysis focus.

Compared with traditional distributed data management systems, Smart Home Guard has the following advantages.

1) Immutability: In the security solutions of traditional information system, security relies on layers of guarded access control. Smart Home Guard uses blockchain technology to record transactions in a database that can be accessed by anyone, but by clever design and supplemented by cryptography and consensus mechanisms, if an attacker modifies a data, he must change all subsequent data.

2) Heterogeneous multi-activity: Each system participant in Smart Home Guard is an off-site multi-activity node, which is a multi-activity system essentially. If a node is controlled by hackers or encounters network problems, hardware failures, software errors, it will not affect the system and other participating nodes.

Smart Home Guard, the data management system server is connected to the blockchain server, calls the smart contract interface, and adds records and queries information through the Ether client. When a user adds data, an Ethernet Virtual Machine (EVM) generates a new block. Then the Ethernet server adds it to the blockchain through a "mining" operation. Querying data does not generate new blocks. Therefore, the blockchain server in this paper stores information about the operations performed by visitors, such as additions, deletions, updates, and queries. These records are permanent and cannot be changed.

2 RELEVANT THEORETICAL FOUNDATIONS

The application of blockchain technology in IoT is one of the research hotspots in recent years. Xiubo Liang et al. [6] analyzed data storage security, privacy security, data access security and data sharing security four aspects, the blockchain technology facing data security problems and related technology solutions. Lihua Song et al. [14] proposed a model to improve the access control security of the IoT, which was based on zero-knowledge proof and smart contract technology in the blockchain. Liu Tao et al. [7] proposed a service architecture platform of blockchain technology for the terminal connection of IoT. D.M Sheeba et al. [15] believes that the use of blockchain in the IoT environment provides the flexibility to handle large amounts of data in a secure manner between IoT applications and consumers.

In terms of data storage and privacy protection, Zixiong Zhao et al. [18] proposed that the peer-to-peer nature of some nodes can be sacrificed to build a polycentric rather than peer-to-peer blockchain, where network nodes can be properly managed in a hierarchical manner. Nazar Waheed et al. [17] presented the current solutions to IoT security and privacy by utilizing machine learning algorithms, blockchain techniques, and the integration of both.

In summary, many researchers have applied blockchain technology to IoT, which proved that blockchain technology is applicable to the field of IoT. Then as an area of the Internet of Things, the smart home also applies. However, most of the research are focused on the field of data sharing, which requires significant changes to existing information systems, rather than focusing on supervision and evidence retention. Rizwan Majeed et al. [11] presented a novel idea of a smart home that uses a machine learning algorithm (Support Vector Machine) for intelligent decision making, it also uses blockchain technology to ensure identification and authentication of the IoT devices. Jung Hyun Ryu et al. [12] presented a digital forensics framework for the IoT environment based on the blockchain technology. By using blockchain technology, the integrity of the data to be analyzed has been ensured and security has been strengthened, and the preservation of integrity is more reliable by a decentralized method of integrity preservation.

3 Data collection

To maximize deployment convenience, the data operation record collection design of Smart Home Guard is decoupled from the smart home data management system as far as possible. Currently, Smart Home Guard implements two types of data: database monitoring and private API. [13] Table 1 lists the types of data currently collected by Smart Home Guard.

Table 1: Data Collected Table

Name	Description	Data Type
time	access time	timestamp
ip	access ip	string
id	name of the data visitor	string
op_type	operation type string	string
data_type	type of user data accessed	string
devices	devices in this operation	string

3.1 Data Base Monitoring

Currently, most smart home data management systems have built-in logging capabilities to record important information such as access time, visitor ID, and visitor IP. To minimize the deployment cost, Smart Home Guard has implemented a set of database middleware [2] which is used to track the log records of smart home data management systems and extract smart home user data access information from database requests based on pre-configuration. Currently, Smart Home Guard implements middleware for MySQL only, and its structure and workflow are shown in Figure 2.

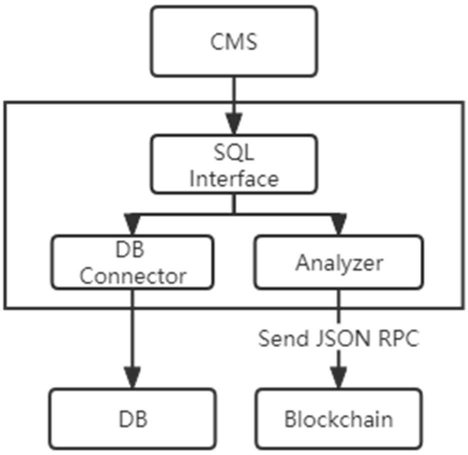

Figure 2: Database Monitoring Architecture.

The middleware implements a standard MySQL database interface, running on port 3306 by default, for receiving SQL data from the back-end program of the smart home data management system and then hand it over to the analyzer and database server. The parser is used to extract the smart home user data access information from the SQL statements, as shown in Figure 3. After finishing the information extraction, the parser processes the data to generate structured data and uploads the structured data to the blockchain through JSON RPC. [3, 16]

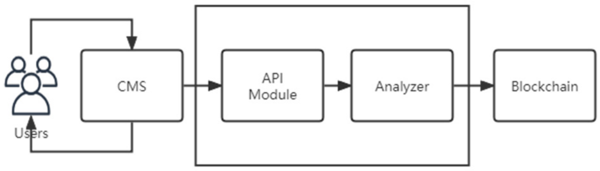

Figure 3: Analyzer Architecture.

3.2 Private API

Ideally, smart home data management systems mostly use common databases such as MySQL, Oracle, and MSSQL [8], but this is not the case, and many of them do not use these. In addition, some smart home data management systems even have proprietary designs, and information is difficult to monitor through databases. For this situation, this paper designs and implements a private API that developers can use to receive data reports from smart home data management systems. The workflow of this interface is shown in Figure 4. When the user reads and writes the smart home data management system, the system back-end program performs a series of database operations, and after all the database operations are completed, the system back-end program calls the private API of Smart Home Guard to report the data to the blockchain.

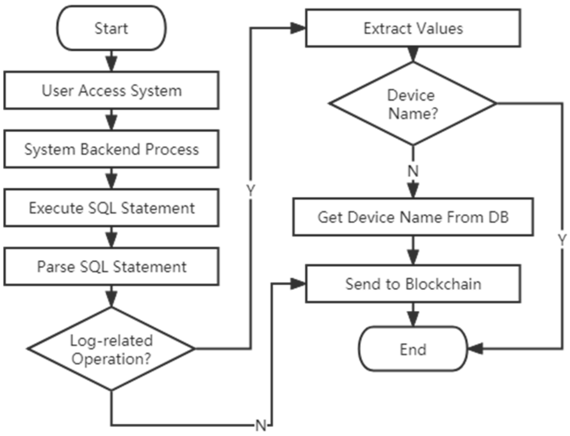

Figure 4: Private API Process.

4 SYSTEM DESIGN

In this section, we will detail the architecture design and implementation of Smart Home Guard, which uses blockchain technology to provide complete records of user access and operation data. The traceability and immutability of blockchain can be used to establish a credible third party while providing effective data support and supervision for regulators and the judiciary. In addition, with the function of monitoring data-level suspicious users, Smart Home Guard can help regulators detect internal data thieves and large amounts of unauthorized data downloads by external attackers quickly.

4.1 System Architecture

The architecture of Smart Home Guard is shown in Figure 5, which includes three components: Blockchain Infrastructure, Agent, and User Platform. The blockchain infrastructure is responsible for implementing all blockchain-related functions. As the core module, it includes data storage and permission control, and is the core module. Agent is deployed on the smart home data management system server, which used to monitor the operations of the smart home data management system and upload the relevant information to the blockchain module. For management, the user platform is responsible for monitoring suspicious users. Authorized users of the regulator can query the operation records of various smart home data management systems through this platform.

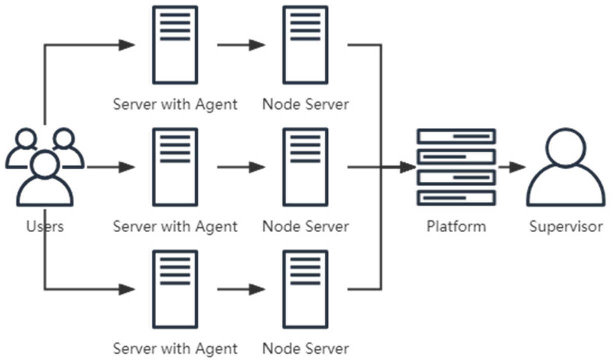

Figure 5: Architecture of Smart Home Guard.

4.2 Blockchain Infrastructure

The private blockchain of Smart Home Guard is based on Ethernet and deployed in a virtual private network, and the smart contract is deployed on the private blockchain. The blockchain infrastructure is used to complete blockchain related operations, such as logging, querying, adding, deleting, updating, checking, etc. through its own smart contract and JSON RPC interface. The blockchain infrastructure consists of two main parts, private blockchain and smart contracts. The latter is deployed on the private blockchain. It implements several interfaces which are used for reading and writing data and verifying permissions. As shown in Table 2, to facilitate the backward compatibility of smart contracts, variable names are not defined according to the actual field meanings during the contract development process. Also, several data fields are reserved for data storage.

The smart contract implements the following functions as interfaces.

1) Uint2Str, which converts unsigned integer data to strings.

2) StrConcat, stitching 4 strings into one long string.

3) TwoStrConcat, which splices 2 strings into a single string.

4) Convert, which splices 4 input parameters into a long string, and then converts it to Byte32 type.

Table 2: Field Correspondence Table.

Name	Name in Smart Contract	Used in Search Record
time	timestamp	Yes
ip	keyArgl	Yes
id	keyArg2	Yes
N/A	keyArg3	yes (reserved, null by default)
devices	infoArgl	No
op_type	infoArg2	No
data_type	infoArg3	No
N/A	infoArg4	No

The smart contract implements the interface functions for manipulating relevant information of blockchain, including adding, deleting, updating and querying. The specific code is as follows:

```
Input: timestamp, keyArg1, keyArg2, keyArg3, infoArg1, infoArg2, infoArg3, infoArg4;
Function AddRecord Input Arguments
    put timestamp, keyArg1, keyArg2 and keyArg3 into a structure keys;
    put the new keys into the global array of structre named keysArray, to store the index information;
    A = Convert(timestamp, keyArg1, keyArg2, keyArg3);
    calculate the hash value of A and assign it to key as index;
    put infoArg1, infoArg2, infoArg3, infoArg4 and the address of user into a structure info;
    put the key and info into global map named records, use key as index;
```

In addition, the contract implements an interface that allows users to retrieve, add, delete, modify and search records of operations from the blockchain, the specific code is as follows:

```
Input: timestamp, keyArg1, keyArg2, keyArg3;
Function GetRecord Input Arguments
    A = Convert(timestamp, keyArg1, keyArg2, keyArg3);
    calculate the hash value of A and assign it to key as index;
    get the values corresponding to key from global map records;
    return values;
```

4.3 Agent

Agent runs on the smart home data management system server, and each system corresponds to an Agent, which collects data from it and reports to the blockchain. The operation and maintenance engineer can modify the configuration file and then choose the operating mode of the Agent (database monitoring or private API). The workflow of the Agent is shown in Figure 6.

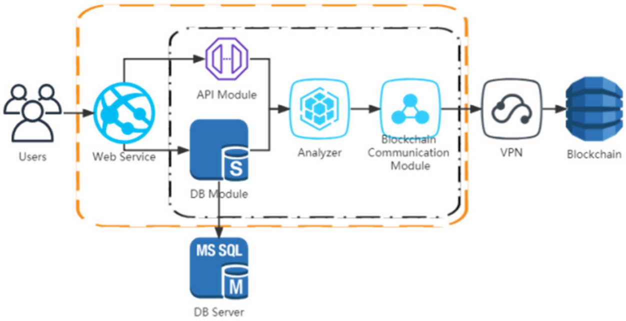

Figure 6: Workflow of Agent.

Agent is shown in Figure 6. The blockchain communication module interacts with the blockchain based on JSON RPC. After encoding, analyzing and recordsing, the module sends the obtained log data to the blockchain. The database module receives SQL statements from the smart home data management system and then forwards them to the parser and database server of the smart home data management system. The private API module implements a microservice API locally to process and clean the data manipulation interface provided by the smart home data management system. It can also provide structured data to the analyzer. The analyzer extracts the required information from the data passed by database module and private API module, and invokes the blockchain communication module to complete the relevant operations. The analyzer maintains a message queue for asynchronous operation to prevent the Agent from interfering with the operation of the smart home data management system.

4.4 User Platform

The user platform is an interaction platform provided by Smart Home Guard to the regulator. The regulator can query the data creation, modification and other operation records of all smart home data management systems connected to Smart Home Guard through this platform, while the platform sends alerts to the regulator for unauthorized bulk data access through the data recorded in the blockchain. The user platform includes 3 components: the Web front-end server, the Web back-end server, and the security audit server, as shown in Figure 7.

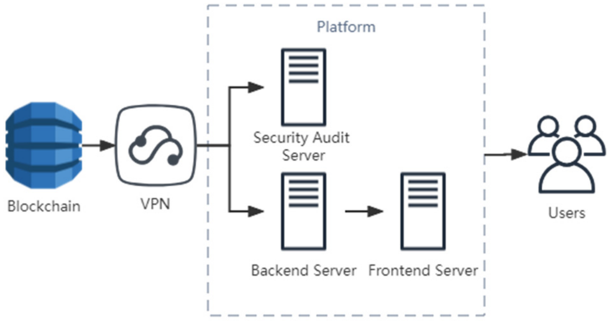

Figure 7: Workflow of User Platform.

In the user platform, the security audit server obtains data from the blockchain infrastructure and performs security analysis. The web front-end server is based on the Apache HTTP Server [4]. Smart Home Guard provides visitors with a user-friendly and beautiful React-based human-computer interface. The web back-end server is implemented on top of Sanic, a high-performance web framework built on Python. The web backend server is mainly responsible for the actual data processing of the user platform, including operations such as obtaining smart home data operation records from the blockchain infrastructure, obtaining security alerts from the security audit server, and providing information to users through the Web frontend server.

5 Security Audit

Currently, the main threat to the smart home data management system is the theft of data by attackers, including a variety of attack types. Except for external penetration, the attackers may also be internal malicious users. Most of the traditional smart home data management systems are deployed on internal networks, which are difficult to access by external users, so the risks mainly come from internal. However, with the further development of the Internet, more and more smart home companies are providing APP remote services, which allow users to check their smart devices and operation records through the Internet. This service is usually realized through secondary development of existing smart home data management systems by smart home organizations, which enables external attackers to use vulnerabilities to attack the smart home data management system, which undoubtedly increases the risk of data. Smart Home Guard mainly implements security monitoring for batch operations (data downloading,data modification and so on) identified as authorized users by the smart home data management systems.This is mainly to respond to internal attackers and vulnerabilities like horizontal privilege escalation and vertical privilege escalation, which are difficult to handle by existing Web security defense such as Web Application Firewall (WAF) and Intrusion Detection System (IDS) (BACUDIO 2011, LIAO 2013).

5.1 Design and Implementation

Smart Home Guard's security auditing is implemented based on rules for detecting smart home data access records in the blockchain, with a core approach of raising alerts when abnormal batch data operations are detected. The security audit server is a part of the Smart Home Guard user platform. It includes a blockchain communication module for fetching data from the blockchain infrastructure, an analyzer for analyzing recent smart home data operation records, a lightweight database for storing alert information, and a microservice API module for receiving user calls and monitoring alerts, as shown in Figure 8.

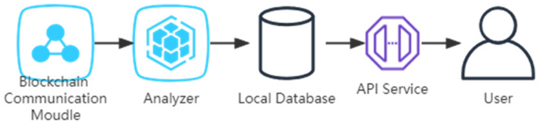

Figure 8: Security Audit Server.

5.2 Rule Definition

As for detecting rule-based malicious behavior, the most important thing is to institute reasonable and effective rules. In this paper, we extract following features and set thresholds in the data recorded by Smart Home Guard's blockchain infrastructure:

1) S_{user}: The set of data operations for a specified user.

2) $S_{decvices}$: The set of operations performed on a specified device in a day.

3) $S_{\Delta t}$: Set of a specific user's interval time between two consecutive data operations.

4) l_{time}: Specifies the threshold value of the interval between two consecutive operations by the user, which is determined by the regulator, i.e., the enterprise.

5) l_{count}: A threshold value specifying the number of device operations, determined by the regulator, i.e., the enterprise, and must be smaller than the buffer size.

6) S_{optype}: The set of all operation types of a certain operator.

This paper uses these characteristics to define the following rules. The analyzer sends an alert when user's behavior meets any of them.

1) Fast and continuous operations. Based on existing experience and historical vulnerability reports, attackers usually use fast and continuous operations when they try to steal data.

Therefore, when Δ_{ti} is the ith item in $S_{\Delta t}$, we define the following rule:

$$\forall \Delta t_i \in S_{\Delta t}, \Delta t_i < l_{time} \qquad (1)$$

2) Similar operation interval. Experienced attackers would try to hide themselves. They may perform a data operation every few seconds, which means that there tends to be a regular interval between the two operations. Therefore, we define the following rule:

When Δ_{ti} is the ith item in $S_{\Delta t}$, if i meet

$$0 < i < |S_{\Delta t}| \qquad (2)$$

We have

$$\forall \Delta t_i \in S_{\Delta t}, |\Delta t_i - \Delta t_{i+1}| < l_{time} \qquad (3)$$

3) Too many data operations in one day. An attacker who aims to obtain data in batches through illegal means has a much larger amount of data access than normal users, so we designed a threshold. The rule defined with this threshold are shown as following. Users can set this threshold value according to the scale of the smart home organization.

$$|S_{device}| < l_{count} \qquad (4)$$

4) Same type of operation. Experienced attackers typically use scripts to perform slow batch processing of data sets during unobtrusive hours such as midnight. These scripts usually do not perform operations outside of the task and work in a completely different mode from the average user. If a smart home data management system user performs the same type of operation continuously, then the operator is probably a malicious user. Therefore, we define the following rule:

If s meet

$$s \in S_{optype} \tag{5}$$

We have

$$\forall p \in S_{optype}, s = p \tag{6}$$

6 CONCLUSION AND FUTURE WORK

This paper aims to design a blockchain-based smart home user privacy protection system - Smart Home Guard, which will be applicable to the existing data systems at all ends of the smart home field. Smart Home Guard can help users and enterprises realize unmodifiable monitoring of data access, facilitate users and enterprises to view data usage records, and serve as a strong evidence when data leakage is found. Smart Home Guard has proven to be effective in meeting design requirements and providing regulators with a low-cost solution for monitoring smart home data usage. However, the current version of Smart Home Guard has some problems: firstly, it is based on the private blockchain of Ether, and its performance and security are limited by Ether. Secondly, the tracking of smart home device operation is not accurate enough. For example, in order to be more compatible with the existing smart home data management system, the existing data collection technology cannot distinguish between the same-named devices in the same system. Finally, the security audit mainly guards against bulk data operations by malicious attackers. Because of its over-reliance on rules, it is unable to find attackers with more ambiguous behavioral characteristics. The next step will be developing a new blockchain infrastructure for Smart Home Guard and using new algorithms, such as Byzantine Fault Tolerance (BFT) [10], to enhance security.

Acknowledgements: Thank Dr. Liu Xin of Lanzhou University for his help in this paper.

REFERENCES

[1] BACUDIO A G, YUAN X, CHU B T B, et al. An Overview of Penetration Testing[J]. International Journal of Network Security & Its Applications, 2011, 3(6): 19-38.

[2] Bakken D. Middleware[J]. Encyclopedia of Distributed Computing, 2001, 11.

[3] Bray T. The javascript object notation (json) data interchange format[R]. 2014.

[4] FIELDING R T, KAISER G. The Apache HTTP Server Project[J]. IEEE Internet Computing, 1997, 1(4): 88-90.

[5] HAN Zhijun, HU Huapeng, SUN Kai.A Survey for Sercurity Analysis Technologies of Smart Home Devices[J]. Information Security and Communications Privacy,2022(6):144-153.

[6] LIANG X, WU J, ZHAO Y, et al. Review of blockchain data security management and privacy protection technology research[J]. Journal of ZheJiang University (Engineering Science), 56(1): 1-15.

[7] Liu T, Yuan Y, Yu Z. The service architecture of Internet of things terminal connection based on blockchain technology[J]. The Journal of Supercomputing, 2021, 77(11): 12690-12710.

[8] LONEY K. Oracle Database 10g: the Complete Reference[M]. London: McGraw-Hill/Osborne, 2004.

[9] LIAO H J, LIN C H R, LIN Y C, et al. Intrusion Detection System: A Comprehensive Review[J]. Journal of Network and Computer Applications, 2013, 36(1): 16-24.

[10] Li Y, Qiao L, Lv Z. An optimized byzantine fault tolerance algorithm for consortium blockchain[J]. Peer-to-Peer Networking and Applications, 2021, 14(5): 2826-2839.

[11] Majeed R, Abdullah N A, Ashraf I, et al. An intelligent, secure, and smart home automation system[J]. Scientific Programming, 2020, 2020.

[12] Ryu J H, Sharma P K, Jo J H, et al. A blockchain-based decentralized efficient investigation framework for IoT digital forensics[J]. The Journal of Supercomputing, 2019, 75(8): 4372-4387.

[13] ROLSTON M E, FICKLIN J L, NAIRN M A, et al. User-application Interface: U.S. Patent 9, 176, 747[P]. 2015-11-03.

[14] Song L, Ju X, Zhu Z, et al. An access control model for the Internet of Things based on zero-knowledge token and blockchain[J]. EURASIP Journal on Wireless Communications and Networking, 2021, 2021(1): 1-20.

[15] Sheeba D M, Jayalakshmi S. Lightweight Blockchain to Improve Security and Privacy in Smarthome[J].

[16] Thurlow R. RPC: Remote procedure call protocol specification version 2[R]. 2009.

[17] Waheed N, He X, Ikram M, et al. Security and privacy in IoT using machine learning and blockchain: Threats and countermeasures[J]. ACM Computing Surveys (CSUR), 2020, 53(6): 1-37.

[18] Zhao Z, Ma J. Application of Blockchain in Trusted Digital Vaccination Certificates[J]. China CDC Weekly, 2022, 4(6): 106.