Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9850836, 1 page
https://doi.org/10.1155/2023/9850836

Retraction

Retracted: A System for Trusted Recovery of Data Based on Blockchain and Coding Techniques

Wireless Communications and Mobile Computing

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Chen, Y. Yan, S. Guo, Y. Ren, and F. Qi, "A System for Trusted Recovery of Data Based on Blockchain and Coding Techniques," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 8390241, 12 pages, 2022.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 8390241, 12 pages
https://doi.org/10.1155/2022/8390241

Research Article

A System for Trusted Recovery of Data Based on Blockchain and Coding Techniques

Jinqian Chen,[1] **Yong Yan**,[2] **Shaoyong Guo**,[1] **Yinlin Ren**,[1] **and Feng Qi**[1]

[1]*State Key Laboratory of Networking & Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China*
[2]*Electric Power Research Institute, Zhejiang Electric Power Corporation, Hangzhou 310007, China*

Correspondence should be addressed to Feng Qi; qifeng@bupt.edu.cn

Received 9 August 2021; Revised 26 October 2021; Accepted 3 December 2021; Published 17 January 2022

Academic Editor: Chi-Hua Chen

With the continuous development of information technology, the Internet of Things has also been widely used. At the same time, in the power Internet of Things environment, reliable data is essential for data use and accurate analysis. Data security has become a key factor in ensuring the stable operation of the power grid. However, the power Internet of Things devices is extremely vulnerable to network attacks, leading to data tampering and deletion. Resisting tampering, preventing data loss, and reliably restoring data have become difficult to ensure data security. In order to solve this problem, this paper proposes a trusted data recovery system based on blockchain and coding technology. Data nodes of the power Internet of Things encode key data and back them up to the blockchain network through a data processing server located on the edge. The data processing server performs real-time detection of the data integrity of the data nodes. When the data is tampered with or deleted, the data processing server promptly obtains the corresponding data encoding blocks from the blockchain network, decodes them, and sends them to the data node to complete the data recovery task. According to the test result, the data backup speed of this system is increased by 15.3%, and the data recovery speed is increased by 19.8% compared with the traditional scheme. It has good security and real-time performance. Meanwhile, it reduces the network and storage resource overhead in the data backup and recovery process.

1. Introduction

With the rapid development of network information technology, network attacks have become more and more common. Network hackers can easily use Trojan horses, worms, and local vulnerabilities to attack devices on the network. In the power Internet of Things, attackers usually maliciously delete or tamper with the data (such as collected and summarized data, firewall configuration information, virus database data, and log file) stored in the terminal to interfere with the stable operation of the power Internet of Things. For example, attackers can tamper with the firewall configuration of the device to further implement a distributed denial of service (DDOS) attack. IoT administrators can use data backup and recovery technology to reconstruct the firewall. Then, the IoT environment can be repaired. So, data backup and recovery technology is an important method to ensure the stable operation of the power Internet of Things [1]. Currently, mainstream data recovery solutions mainly include two types: (1) centralized data backup and recovery solutions. For example, all data nodes in the Internet of Things back up their own data to a central server (usually, it is a cloud service) according to the rule. When the data of a node is lost, it downloads the corresponding data from the centralized server to complete the recovery; (2) distributed data backup and recovery scheme. For example, in a distributed storage system, data is usually stored redundantly. When the node data is lost or tampered with, the IoT node can use the P2P network to download the required data from the remaining surviving nodes to complete data recovery [2, 3]. However, both of these schemes have certain shortcomings. The first scheme has the

advantages of low cost and easy management, but it also brings low security and reliability [4]. For example, a centralized storage system may be attacked by network hackers, resulting in data stored in the centralized server being tampered with or stolen. At the same time, studies have found that data recovery in a centralized manner is uneconomical [5] because it has to solve the problems of large storage data, network congestion, complex asynchronous processing, and low efficiency [6, 7]. In the second scheme, the local data node is vulnerable to hijacking or worm infection. Malicious data may be sent to the data nodes during the recovery process, such as the Stuxnet worm. Therefore, the second solution is difficult to solve the problem of untrusted data sources in the data recovery process. It is difficult to meet the security and reliability requirements of the power Internet of Things data recovery system.

The blockchain is essentially a distributed database. Blockchain has the characteristics of decentralization and can ensure that the stored data is difficult to tamper with. At the same time, the blockchain's oracle mechanism can ensure that reliable data is provided for the blockchain, and IoT data nodes can accurately and reliably obtain data on the blockchain through smart contracts. Therefore, blockchain technology provides a promising solution to the above problems. However, the storage scalability of the blockchain is poor, and each node needs a complete ledger in the storage system to maintain the decentralization and consistency of the ledger, which puts high storage requirements on each node server. This has caused the storage scalability of the blockchain to become a major bottleneck for the application of blockchain technology to data backup and recovery in the power Internet of Things. Coding technology can effectively reduce storage resource overhead, so coding technology has begun to be applied to improve the scalability of blockchain storage [8]. As a typical coding scheme, erasure coding can effectively solve the problem that the distributed storage system cannot complete data reconstruction due to partial loss of data slices during data transmission, improving data transmission efficiency and reducing network resource overhead. At the same time, the data coding block on the blockchain is publicly visible to the entire network. In order to prevent the coding matrix from being stolen by network attackers and cause IoT data leakage and to ensure the privacy of power IoT data, it is necessary to apply access control technology to protect the coding matrix.

In summary, the existing Internet of Things data backup and recovery methods have disadvantages such as high storage overhead and low data recovery efficiency. At the same time, these methods face the challenges of data leakage and untrusted data sources in data recovery. Therefore, this paper proposes a trusted recovery system for power Internet of Things data with high performance. The system can effectively ensure the privacy of power Internet of Things data while achieving decentralized and credible recovery. Our contributions are summarized below.

(i) A trusted data recovery system is implemented based on blockchain technology. Data is stored in the blockchain in the form of coded blocks. The data of the blockchain is verified based on the oracle mechanism to ensure the reliability of the data. We built the corresponding smart contract. The data processing server will automatically and accurately obtain the coding block for data recovery from the blockchain network

(ii) Based on the Jerasure coding library, an improved coding scheme is proposed. This scheme increases the data size of each matrix operation to 128 bytes and sets the word size to 64 bits. At the same time, preprocessing and concurrent computer mechanisms are added to read the data in advance. It is processed in the cache to reduce I/O operations and increase the encode and decode rate

(iii) Based on the CP-ABE algorithm, an attribute-based access control model is proposed to protect the encoding matrix with fine-grained granularity. System administrators formulate access control strategies, and the access control model will automatically distribute the coding matrix according to the access control strategy based on the subject attributes of the data nodes of the Internet of Things and the object attributes of the data

(iv) Through simulation tests, the system has better data backup and recovery efficiency than previous data recovery systems. The data backup rate has increased by 15.3%, and the data recovery rate has increased by 19.8%. The system can effectively guarantee the privacy of IoT data while reducing the storage and communication resource overhead during data backup and recovery

2. Related Work

With the continuous development of information technology, the Internet of Things has been widely used. In applying the Internet of Things and wireless sensors, the terminals in the Internet of Things will use the data stream for specific data processing and analysis. Due to unexpected power on of devices and network attacks, local data nodes may experience data tampering and loss. Reliable data is essential to data accurate analysis. The safe storage and trusted recovery of data become critical in ensuring the reliable operation of the Internet of Things.

In terms of data security storage, Tchernykh et al. [9] proposed a multicloud storage architecture called WA-MRC-RRNS in the Internet of Things environment, which combines a weighted access control scheme, threshold secret sharing, and remaining redundant data. The system also proposes a multinode fault detection/data recovery mechanism based on the redundant residual number system. Through simulation tests, this architecture can effectively ensure data security and improve the data credibility of the IoT infrastructure.

Xia et al. [10] proposed a highly robust, secure, and trust-oriented IoT data storage model (RoSES) based on

edge storage and completely partial reconstruction codes. The storage model can achieve data robustness, high security, and lightweight local computing. At the same time, the model supports a trust-oriented data access (TODA) strategy, which can realize the legal data access of uncertain access requesters under the premise of wide applicability, thereby realizing the secure sharing of IoT data.

Wang et al. [11] studied the data recovery problem based on QoS guarantee and system robustness in the information-centric Internet of Things system (IC-IoT). They proposed a data recovery algorithm based on rarity perception. Its core idea is to establish a rarity index to evaluate data copies and service demand distribution comprehensively. The algorithm will eliminate unnecessary copies and gradually restore the original data according to the rarity of the original data and the priority of recovery. Experimental and simulation results show that compared with the traditional direct data recovery method, the algorithm has better QoS performance, and the robustness has been significantly improved.

The Internet of Things (IoT) has become an emerging technology in the past decade, and the number and research prospects of smart devices and related technologies have grown rapidly. Due to IoT terminals' low processing power and storage capacity, the existing security or encryption technology is not suitable for protecting IoT data [12]. At the same time, the importance of IoT security will become more obvious and huge. However, the Internet of Things still has many security issues and is vulnerable to attacks from some potential factors [13]. Therefore, some researchers have used blockchain technology as a decentralized method to solve security and privacy issues in the Internet of Things, such as data management, access control, and data recovery.

Bae and Shin [14] proposed an automatic recovery mechanism for data systems based on the blockchain. This mechanism uploads the data copy summary to the blockchain and regularly checks and restores the data. Before the data recovery process, a comparison will be made. Blockchain copy summary value information and local redundant copy summary value information are checked to ensure that the copy is correct. This mechanism solves the problem that the local data copy cannot be restored correctly after the local data copy is tampered with. In smart cities, Mishra and others [15] have designed a data protection system for key document data of large organizations based on the blockchain, which includes inspection, location, and recovery. The malicious behaviour of data tampering is detected and recovered, but this system is only suitable for recovering document data, not real-time data. In important business areas, Zhang and Li [16] combined blockchain and smart contract technology to improve the existing backup and recovery technology and adopted role-based access control strategies to strictly audit the data backup and recovery process to prevent data from being compromised. In the field of the supply chain, Cha et al. [17] proposed a data management and recovery system that uses blockchain and key agent encryption. This system enhances data integrity, availability, and traceability while solving the failures, denial of service attacks, and undeniable problems. In order to reduce storage costs, Liu et al., N. Liang et al., and W. Liang et al. [18, 19, 20] used coding techniques to improve the storage efficiency of distributed storage systems. In the field of smart grid, blockchain technology is also beginning to be applied. Ferrag and Shu [21] proposed a novel deep learning and blockchain-based energy exchange framework for smart grids. This framework uses blockchain for facilitating the exchange of excess energy among neighboring nodes. It also uses a recurrent neural network to detect network attacks and fraudulent transaction in exchange.

Through research on data security storage and trusted recovery related literature, it is found that people usually do not need to back up their data in a central database. Data can be safely stored in distributed storage systems or different blockchain network nodes, and the blockchain can guarantee its authenticity. Moreover, it can prevent unauthorized access. However, some of the above studies are still in the discussion of theoretical concepts [15, 16], some apply blockchain to data recovery in specific fields [17], and some apply access control technology or key technology to data credible recovery to ensure data recovery security, and some just use coding technology to optimize storage efficiency [18, 19, 20]. In summary, there are few studies on improving the recovery efficiency while ensuring the credibility of the data recovery process.

At present, researchers have adopted blockchain technology to ensure the security and privacy of the data recovery process [17, 22]. However, the former blockchain-based data recovery methods use symmetric encryption to protect data privacy. Therefore, these methods have the disadvantage of low data recovery speed. These methods also ignore the reliability of data transmission. The method proposed in this paper uses code techniques to complete the storage of the backup data, which can effectively improve the reliability of data transmission. And this method uses an access control strategy to strictly protect the coding matrix. It can protect the privacy of data while ensuring the high efficiency of the system.

3. System Architecture

Figure 1 shows the architecture of the system. The system mainly includes four entities: blockchain network, local data node, data processing server, and access controller. The power Internet of Things sensors broadcast the raw data to a distributed storage system composed of multiple local nodes. According to the corresponding backup strategy, the local data sends the key data to the data processing server on the edge side. After receiving the data, the data processing server uses the corresponding encoding matrix to encode the key data to generate a data encoding block and finally upload it to the blockchain network to complete the data backup. When the data is lost, the local data node obtains the corresponding coding matrix from the access controller and then sends the data recovery request and the coding matrix to the data processing server. After receiving the data recovery request, the data processing server downloads the corresponding data encoding block from

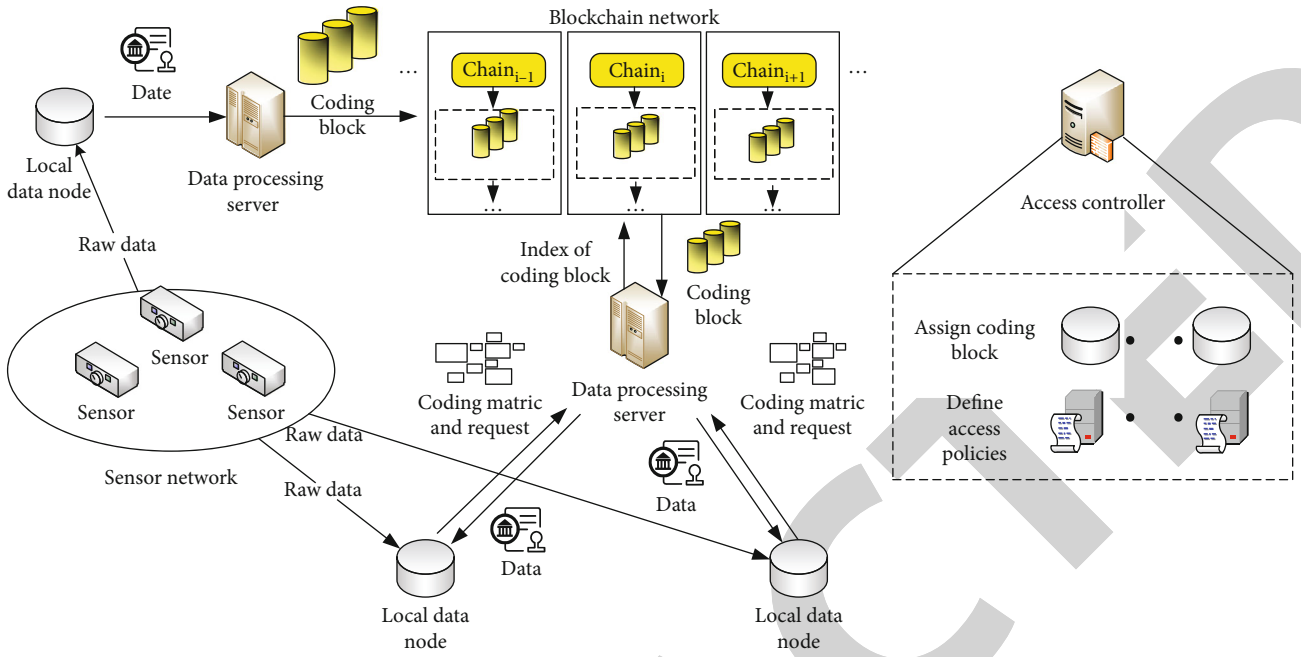

Figure 1: Architecture of system.

the blockchain network, uses the encoding matrix submitted by the local data node to complete the data decoding operation, and finally sends the data to the local data node to complete the data recovery process. The system applies coding and access control technology to data trusted recovery work to solve the problems of large data storage overhead, low transmission efficiency, and privacy leakage in the blockchain system.

3.1. Entities. As shown in Figure 1, the system includes four types of entities.

Local data nodes: the sensors in the power Internet of Things store the collected data in local data nodes, and multiple local data nodes of the same kind form a distributed storage system in full replication mode. The local data node can send data backup and recovery requests to the data processing server.

Data processing server: the data processing server is located on the edge of the local data node and is responsible for processing the data backup and recovery requests submitted by the local data node. Furthermore, it completes the data encoding and decoding operations. The data processing server is trusted for the power Internet of Things.

Access controller: the system uses an access controller to fine-grained control of the key data of the power Internet of Things. The access controller defines the data access strategy to protect the coding matrix and distributes the coding matrix to the data nodes. If the data node has sufficient permissions, it can obtain the corresponding data, and the access controller is centrally maintained and configured by the administrator in the power Internet of Things.

Blockchain network: the blockchain network is responsible for storing the coding blocks of the key data of the power Internet of Things. The local data node obtains the code block from the blockchain network through the data processing server to complete data recovery. The blockchain network is composed of multiple blockchains, and each blockchain is maintained by a single or multiple blockchain nodes.

3.2. Operations. This system provides two types of operation: data backup and data recovery.

Data backup: data backup needs to go through the following three steps:

(1) The local data node sends the data to be backed up and the coding matrix to the data processing server. The data processing server first preprocesses the data and divides it into blocks. Assuming that the data size is M, divide it into k original data blocks of fixed size, denoted as $(F_i)_{i=1,2,\cdots,n}$, and the size of each data block is M/k

(2) The data processing server uses the erasure coding scheme to encode the original data block to obtain $k+m$ coded data blocks (m is the number of check blocks), denoted as $(P_i)_{i=1,2,\cdots,k+m}$. Each coding block is a linear combination of original data blocks. After the encoding is completed, the data processing server sends the used encoding matrix E, data identification, and attributes to the access controller

(3) The data processing server sends $k+m$ code blocks to the blockchain network. The blockchain network stores the coded data blocks according to the corresponding rules. There are $k+m$ blockchains in the system, and each node stores n coded data blocks

When $k=2$, $m=2$, and $n=1$, the data backup process is shown in Figure 2.

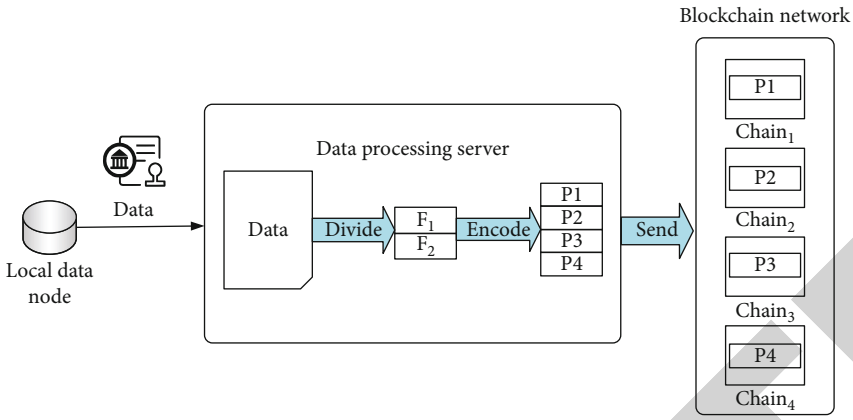

Figure 2: Data backup.

Data recovery: data recovery needs to go through the following three steps.

(1) The local data node sends its digital signature and identification of the data to be restored to the access controller. The controller assigns the coding matrix *EM* to the local data node according to the access control strategy

(2) The local data node sends the data recovery request and the coding matrix *EM* to the data processing server, and the data processing server downloads the required data coding block from the node in the blockchain network according to the data identification (usually the node with a smaller load).

(3) The data processing server decodes the downloaded data code blocks to obtain k original blocks. The server combines these blocks to obtain the original data and returns it to the local data node

When $k = 2$, $m = 2$, and $n = 1$, the data recovery process is shown in Figure 3.

3.3. Design Goals. We list the design goal of this system as follows:

(1) Integrity of data: the data recovery system should ensure that the data was submitted by an authenticated data node and has not been tampered with

(2) Data access control: the encoding matrix should be enforced with fine-grained access control policies so that only authorized data nodes can obtain raw data

(3) Data recovery efficiency: critical data should be backed up and restored efficiently through blockchain and edge computing nodes to adapt to large industrial systems

4. Proposed System

4.1. Data Node Registration Module. The system uses an attribute-based access control mechanism, and each data node has a unique ID as its identifiable identifier. The access controller formulates a registration strategy. Each data node in the system can apply to the access controller for registration and obtain subject attributes through its own MAC address and identifier *ID*. If the data node is verified, the access controller will generate the registration transaction and sign the hash value and timestamp of the registration transaction. Finally, the access controller will pack the registration transaction, signature information, and timestamp together and put it into its transaction pool. In this workflow, all parties communicate through the TCP protocol. The workflow chart is shown in Figure 4.

4.2. Data Encoding and Decoding Module. Erasure code is a forward error correction technology that originated in the field of communications. Erasure codes have low redundancy and high accuracy. We use erasure codes in a storage system, first set up an encoding matrix *E.M.* based on certain rules, and use the encoding matrix and data slices to do matrix multiplication operations to obtain a set of encoded data blocks, as shown in Figure 5 below. The coding matrix here is an 8-row 5-column matrix composed of a fifth-order unit matrix and three rows. Through calculation, the coded data obtained has a total of 8 rows; the first five rows are data slices, and the last three rows satisfy

$$\begin{aligned} C_1 &= B_{11} * D_1 + B_{12} * D_2 + B_{13} * D_3 + B_{14} * D_4 + B_{15} * D_5, \\ C_2 &= B_{21} * D_1 + B_{22} * D_2 + B_{23} * D_3 + B_{24} * D_4 + B_{25} * D_5, \\ C_3 &= B_{31} * D_1 + B_{32} * D_2 + B_{33} * D_3 + B_{34} * D_4 + n_{35} * D_5. \end{aligned} \tag{1}$$

As shown in Figure 6, if a node needs to reconstruct the original data, it needs to randomly select k different coding blocks from the data coding block set. The node obtains the inverse matrix corresponding to the coding matrix and multiplies the inverse matrix with the taken coding block, and then, the original data D is completely reconstructed.

In this system, the Vandermonde matrix is selected as the coding matrix. When calculating the coding matrix,

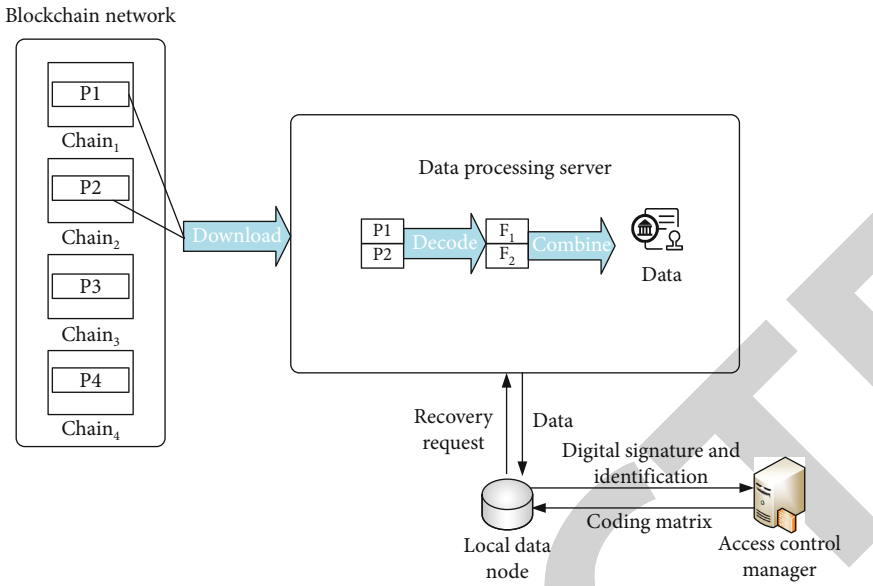

Figure 3: Data recovery.

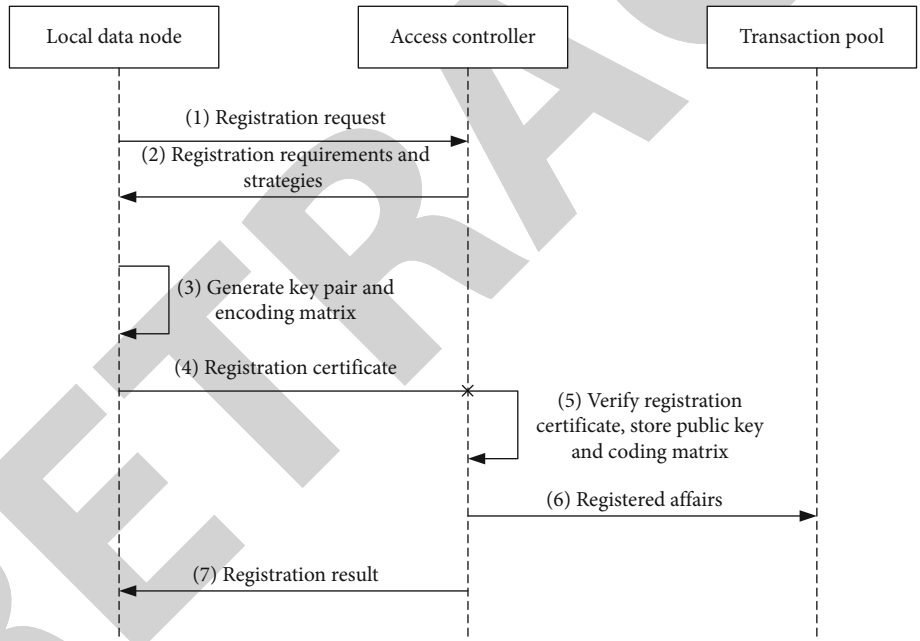

Figure 4: Data node registration flowchart.

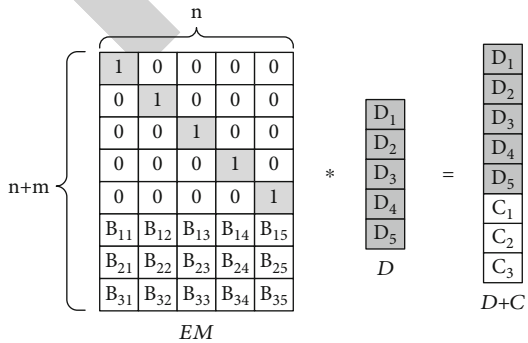

Figure 5: Process of data coding.

different matrix factors can be generated according to the different identification IDs of each data node.

4.3. Data Backup Module. Blockchain can be divided into three categories: public blockchain, private blockchain, and consortium blockchain. The private blockchain has the characteristics of fast transaction speed, and its data storage access efficiency is very close to conventional databases, which meets the real-time requirements of the system. The system designs a data blockchain network based on private blockchain. Single or multiple nodes maintain each private blockchain in the blockchain network to achieve safe and reliable data recovery. The model is shown in Figure 7.

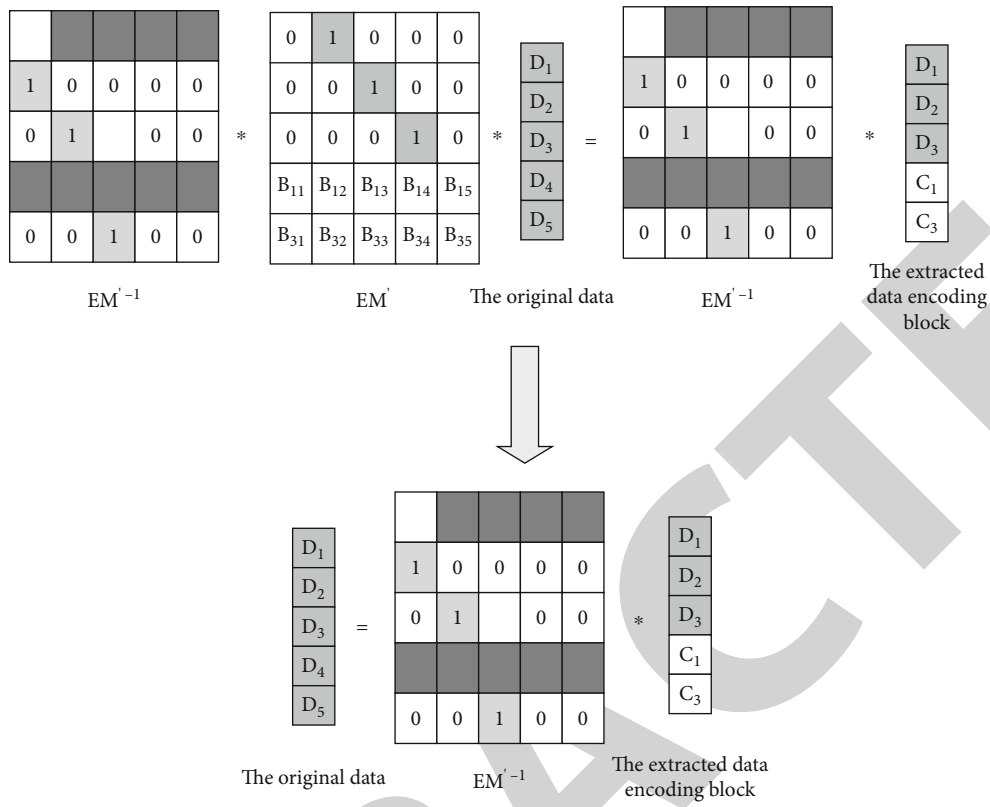

Figure 6: Process of data decoding.

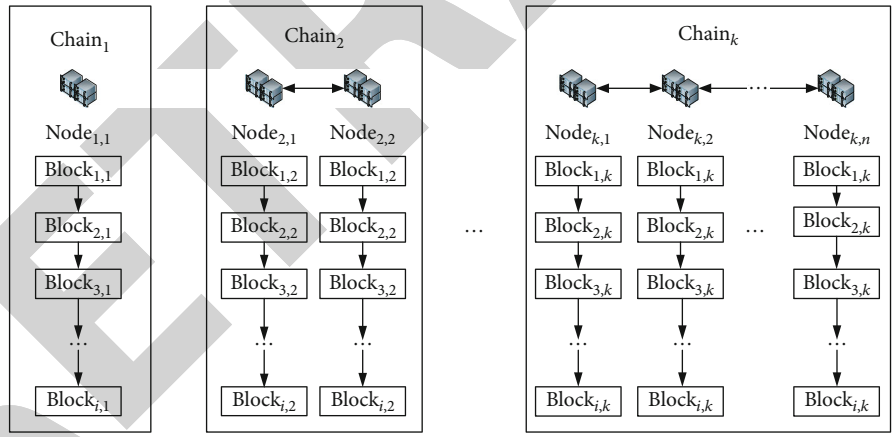

Figure 7: Data blockchain network model.

$Chain_k$ in the figure represents the k_{th} private chain in the blockchain network. $Node_{k,n}$ represents the n nodes in the k_{th} private chain in the blockchain network. $Block_{i,k}$ represents the k_{th} coding block of the i_{th} data.

The data backup process is as follows: (1) the local data node uses the private key to sign the data B and the encoding matrix, then sends them to the data processing server; (2) the data processing server checks the signature of data B. When the signature is complete, it uses the coding matrix of the data node to encode the data to obtain a set of coded blocks E; (3) the data processing server sends the set of coded data blocks E to multiple private blockchains in the blockchain network; (4) after receiving the corresponding coded blocks, the nodes in the blockchain network generate consensus results according to the hybrid consensus mechanism and feedback the consensus results to the data processing server; and (5) the data processing server adds the relevant information of the data to the data directory in the access controller after receiving the successful result of the consensus.

4.4. Data Recovery Module. Attribute-based access control (ABAC), using the attributes of the subject and object as the basis for the basic judgment of authority, can realize

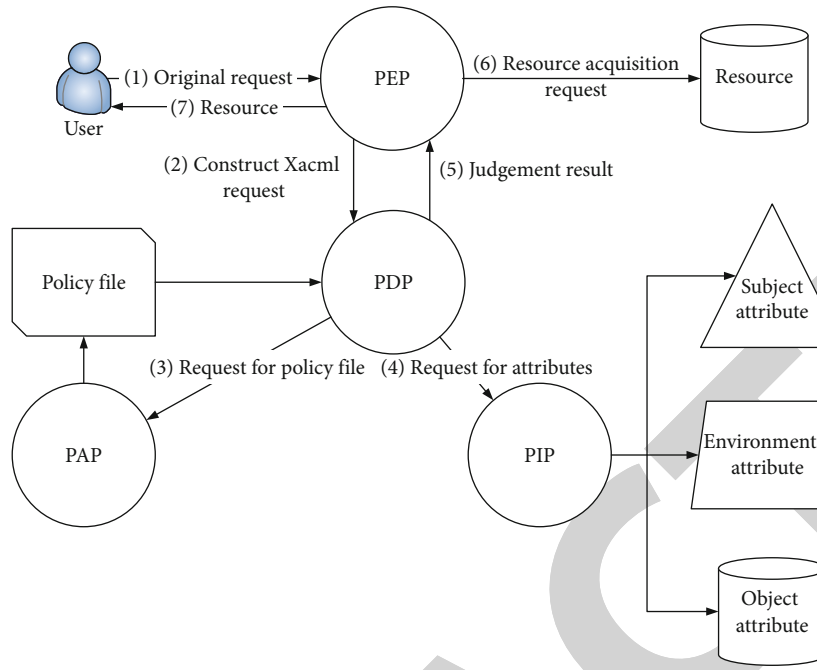

Figure 8: Authorization judgment process. (1) The user sends the original request to apply for access to the resource. (2) After receiving the original request, the policy enforcement point (PEP) constructs a request in xacml format and sends it to the policy deployment point(PDP). (3) PDP requests to obtain the policy file in the policy administration point (PAP) according to the xacml request; (4) After the PDP obtains the policy file, it sends a request to the policy information point (PIP) to obtain the required attribute values (theme attribute, environment attribute, and resource attribute) in the policy file. (5) The PDP makes the judgment result (permit, deny, uncertain, and not applicable) according to the policy file and returns it to the PEP. (6) If the result is permit, the PEP sends an access request to the resource. (7) After the PEP obtains the resource, it returns the resource to the user.

many-to-many access control. It can well separate policy management and authority judgment, has higher flexibility, and can better support fine-grained access control of large-scale information systems. The introduction of environmental attributes makes ABAC support dynamic access control. The authorization judgment process of the attribute-based access control model is shown in Figure 8.

At the same time, the data processing server will perform real-time detection of the data integrity of the local data node to prevent data loss or tampering.

Based on this control and recovery model, the process of data recovery is shown in Figure 9.

5. Simulation

5.1. Simulation Environment. We use Ethereum and the Jerasure library to build a system prototype and the Edge X Foundry framework to deploy the data server on the edge of the power Internet of Things data node. We use the Raspberry Pi 4b as a data node in the power Internet of Things, a desktop-level terminal as a data processing server. We use the Ethereum client (geth) to build a private blockchain network. The consensus mechanism of this private blockchain is PoA. Each private blockchain contains two nodes. Each node in the blockchain has a separate account. The blockchain is deployed in a server cluster with ten identical servers. The hardware configuration is shown in the Table 1. We tested the data encoding and decoding speed, data backup and recovery speed, recovery success rate, and system security under different system scales and encoding matrix parameters.

5.2. Speed of Encoding and Decoding. Assuming that the original data to be backed up by each data node is 100 MB, the original data is an arbitrary binary file. This paper proposes an improved encoding scheme based on the Jerasure encoding library. The scheme increases the data size for each encoding and decoding to 128 bytes and sets the word size to 64 bits. At the same time, it increases the preprocessing and concurrent computer mechanism to save the data in advance. The server reads data into the cache for processing to reduce I/O operations and increase the codec speed. We set the coding matrix as $(m+k)\times k$-order Vandermonde matrix, where $k = 4$, and test the encoding and decoding speed at different m (m is the number of check blocks, and its initial value is 2). As shown in Figure 10, when $m = 2$, there is the maximum encoding and decoding speed, the encoding speed is about 222 MB/s, and the decoding speed is about 196 MB/S. As m increases, the data encoding and decoding speed gradually decreases. Finally, the data encoding speed converges to 54 MB/s, and the decoding speed converges to 42 MB/s.

We tested the speed of the encoding scheme based on the Intel E.C. library in the same environment. The data size is also 100 MB. The result is shown in Figure 10. When $m = 2$, there is the maximum encoding and decoding speed. The encoding speed is about 192 MB/s, and the decoding

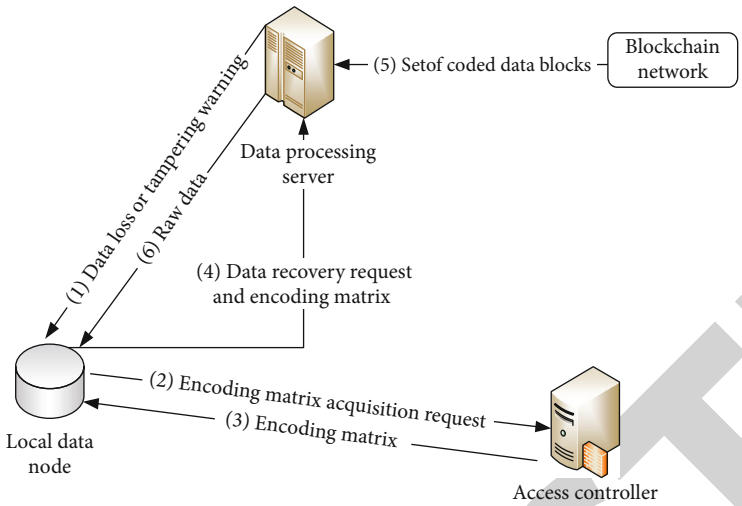

Figure 9: Data recovery. (1) After the data processing server detects that the data of the local data node is lost or tampered with, it sends an alarm message to the data node. (2) The local data node obtains the data identification from the data direction and requests the corresponding coding matrix according to the alarm information. (3) The access controller automatically distributes the coding matrix according to the access control strategy. (4) The local data node sends the data recovery request and the coding matrix to the data processing server. (5) The data processing server obtains the required set of coded data blocks from the blockchain network. (6) The data processing server decodes the coding blocks, then gets the original data and sends it to the data node.

Table 1: System hardware configuration information table.

Platform	OS	CPU	Memory
Power IoT data node	Raspbian 10	BCM2711b0	4 GB
Data processing server	Ubuntu 16.04	AMD Ryzen 7 5800H	16 GB
Server	Centos 7.4	Intel Xeon Gold 5118	64 GB

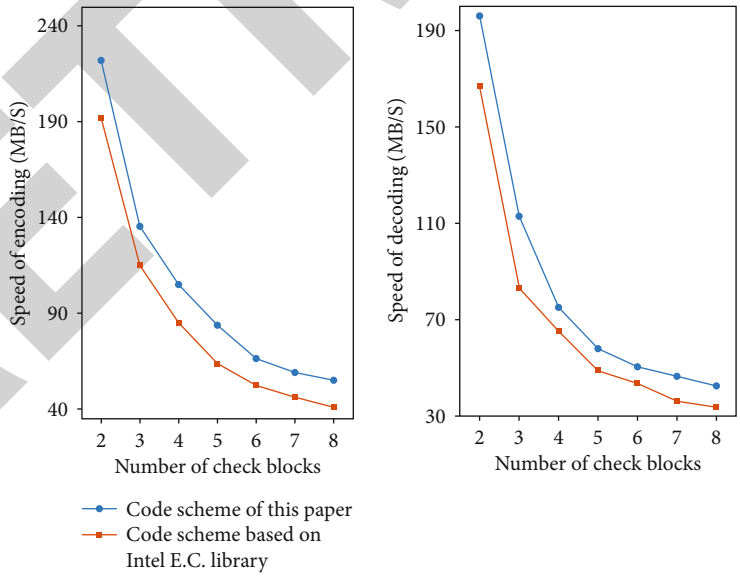

Figure 10: Data encoding and decoding speed.

speed is about 167 MB/S. As m increases, the data encoding and decoding speed gradually decreases. Finally, the data encoding speed converges to 40 MB/s, and the decoding speed converges to 33 MB/s.

We can compare the experimental results and draw the following conclusion. When the number of check blocks is less than 4, the data encoding and decoding speed of the encoding scheme proposed by this paper is greatly improved

compared with the Intel EC encoding scheme. So the encoding scheme proposed by the system is more advantageous.

5.3. Speed of Data Backup and Recovery. This article builds a blockchain network on a server cluster. The blockchain network contains five private chains, and two blockchain nodes maintain each private chain. The coding matrix is a 5×4 -order Vandermonde matrix, and each blockchain node stores one coding block of the data coding block set. Under the above conditions, the system was tested for performance comparison with the Data Protect (DP) data recovery system based on cloud computing proposed by Hewlett-Packard (HP company). DP (originally Omniback) is an automated backup and recovery software for single-server to enterprise environments, supporting disk storage or tape storage targets. It provides crossplatform, online backup of data for Microsoft Windows, Unix, and Linux operating systems. The working mode of DP is server-client. The user can install the client in the distributed database and set the relevant configuration to back up the data to the central server (usually, it is a cloud server). It is a centralized management data recovery system that can use symmetric encryption to prevent privacy leakage [23]. The test data set is sensor data randomly generated by the Raspberry Pi in a simulated environment, and the data size is 1-10 MB. The test network download speed is about 200 Mbps, and the upload speed is about 160 Mbps.

The speed of data backup is shown in Figure 11. From the results, the data backup speed of the system proposed in this paper increases as the data size increases and eventually stabilizes. When the data is small, the time required for data slicing, connection establishment, and service response is a fixed value, which will take up a large system overhead and cause the backup speed to slow down. As the total amount of data increases, edge computing technology can improve system data processing and transmission capabilities. At the same time, compared with the DP data recovery system, the system proposed in this paper has higher data backup efficiency.

Also, we use different sizes of recovery data to test the speed of data recovery. The results of the data recovery speed are shown in Figure 12 below. Compared with the DP data recovery system, the system proposed in this article also has a higher data recovery efficiency.

5.4. The Success Rate of Data Backup and Recovery. In the trusted data recovery method proposed by this system, the success rate of data recovery is related to the number of nodes in the system and the setting of the coding matrix. We set the coding matrix as $(m+k) \times k$-order Vandermonde matrix and assume that there are a total of x nodes participating in the system, and each node stores n coding blocks. In the case of $k = 4$ and $m = 1$, the failure probability of the blockchain node and the link failure rate are both 50%, and we use Matlab simulation software to test the influence of x and n on the success rate of data recovery, the success rate of data recovery = the successful number of recovery/the number of recovery requests. We test 100 times, and the test results are shown in Figure 13.

It can be seen from the figure that the success rate of data recovery increases as the number of nodes and the number of code blocks stored by nodes increase. When $n = 1$ and $c = r/n = 25\%$, the storage optimization efficiency is the highest. However, when the number of nodes in the system is small, the data recovery success rate will decrease. After the number of nodes reaches 14, the data recovery success rate will approach 100%. With the increase of n, the data

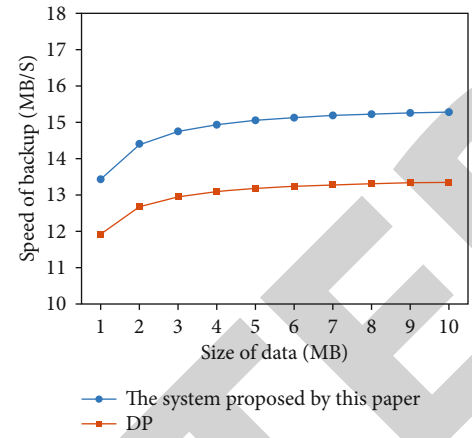

Figure 11: Data backup speed.

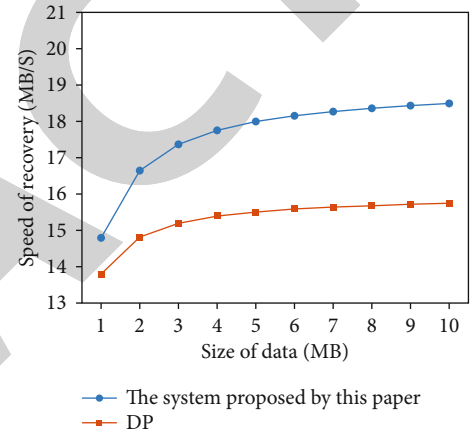

Figure 12: Data recovery speed.

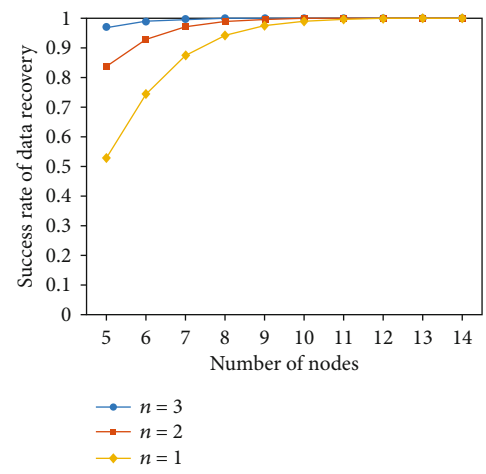

Figure 13: Data recovery success rate based on the coding scheme.

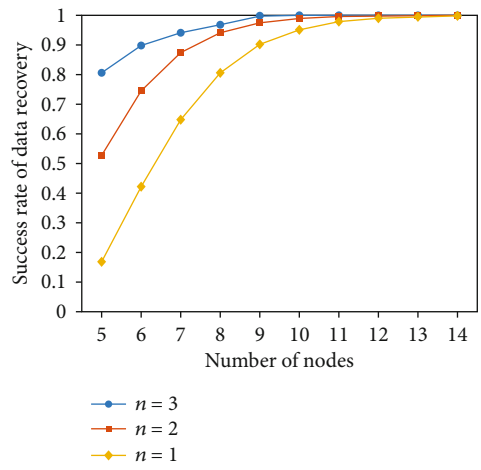

Figure 14: Data recovery success rate based on fragmentation scheme.

recovery success rate of only eight nodes can reach 100% in the end.

In the same environment, the data recovery system using the fragmentation scheme is tested consistently, and the test results are shown in Figure 14.

We compared the experimental results and drew the following conclusions. In the same environment, the data recovery system based on the coding scheme has a higher data recovery success rate than the data recovery system based on the fragmentation scheme under the same number of nodes. Therefore, the system has lower storage and network resource overhead and stronger robustness in data credible recovery.

5.5. Security. This system modifies CP-ABE and implements the access control based on attributes to protect the coding matrix. After testing, the malicious data nodes cannot obtain the encoding matrix without authorization. It is also difficult for the malicious data nodes to decode the encoding block to obtain the original data without the encoding matrix so that the system can prevent data leakage. The block data is also protected by the hash function within the block, which hashes the previous block of the verification chain and mechanically verifies the whole blockchain. The tampered block data cannot pass the verification, thus ensuring that the data in the system is trusted and complete. At the same time, in this system, the edge computing node and the central node of the blockchain will cooperate in supervising the blockchain and preventing the failure of a node. When the fault is detected, the block data will be repaired in time to ensure the security of the data and the continuity of the service.

6. Conclusion

This article proposes a data trusted recovery system in the power Internet of Things environment based on blockchain and coding technology. First, the system uses erasure coding technology to improve blockchain storage performance and data transmission efficiency, providing a basis for efficient and reliable data recovery. Then, the system uses edge computing technology to offload the coding and decoding tasks of local data nodes to the data processing server on the edge, which greatly increases the speed of data backup and recovery. At the same time, the system adopts an attribute-based access control model to protect the coding matrix in a fine-grained manner. It can ensure the privacy of power Internet of Things data and effectively prevent attackers from stealing key information collected by sensors to destroy the power Internet of Things. The system is superior to previous data recovery systems through simulation tests regarding storage performance, data recovery efficiency, and security, mainly reflected in the following aspects: the system uses the blockchain system to store the coding blocks of the key data of the power Internet of Things. Compared with ordinary distributed storage systems, the data stored on the blockchain has extremely strong integrity and is effective enough. It can prevent attackers from tampering with key data, ensure that data nodes obtain original key data after the data recovery, and ensure the stable operation of the power Internet of Things. The system supports dynamic backup of local data nodes and rapid recovery of key data and has good security and real-time performance while reducing the resource overhead in the data backup and recovery process. The system uses the data processing services to monitor the data integrity of the local data node in real time, avoiding the dependence of the data recovery system on the management personnel, so the system is universal, suitable, and easy to use.

However, this system still has some shortcomings: (1) to ensure the reliability of the data source, the system uses the blockchain to store data coding blocks, so there are certain shortcomings in the data access speed; (2) the system uses an attribute-based access strategy, which is more complicated in strategy formulation, has certain requirements for the professionalism of strategy formulation personnel; (3) in terms of coding technology, the system uses an improved RS code coding scheme. The improved coding scheme is only suitable for terminals equipped with a 64-bit operating system, so there may be insufficient compatibility; and (4) the system uses edge computing technology, but no further research has been done on task offloading optimization, which may cause the system to crash under heavy data recovery tasks.

Compared with traditional data trusted recovery system, the system proposed by this paper has many obvious advantages, especially in terms of security, data storage performance, and data recovery efficiency. The system can efficiently help local data nodes complete data backup and recovery to prevent the leakage of key data of the power Internet of Things. With the development of blockchain technology, coding technology, and the power Internet of Things, data credible recovery technology based on blockchain and coding technology will surely receive more and more attention. This advanced data credible recovery system also has high application value. In the future, we will research the optimization of the blockchain consensus mechanism to improve the speed of blockchain data access.

At the same time, we will study the optimization of task offloading in edge computing to improve the service quality of edge computing systems.

Data Availability

The result data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This work is supported by the Key Science and Technology Project of State Grid Corporation of China, No. 5700202019374A0000.

References

[1] Y. Zhao and N. Lu, "Research and implementation of data storage backup," in *2018 IEEE International Conference on Energy Internet (ICEI)*, Beijing, China, 2018.

[2] J. Lin, P. Wang, J. Zhang, Z. Zhang, and H. Sun, "Plug and play technology for power distribution terminal management based on the IoT ideas," in *2019 IEEE 8th International Conference on Advanced Power System Automation and Protection (APAP)*, pp. 196–200, Xi'an, China, 2019.

[3] B. Krishna, S. Kiran, G. Murali, and R. P. K. Reddy, "Security issues in service model of cloud computing environment," *Procedia Computer Science*, vol. 87, pp. 246–251, 2016.

[4] A. Dubey, G. Shrivastava, and S. Sahu, "Security in hybrid cloud," *Global Journal of Computer Science and Technology Cloud and Distributed*, vol. 13, no. 2, pp. 1–7, 2013.

[5] Y. Gu, D. Wang, and C. Liu, "DR-Cloud: multi-cloud based disaster recovery service," *Tsinghua Science and Technology*, vol. 19, no. 1, pp. 13–23, 2014.

[6] A. Greenberg, J. Hamilton, D. Maltz, and P. Patel, "The cost of a cloud: research problems in data center networks," *ACM SIGCOMM Computer Communication Review*, vol. 39, no. 1, pp. 68–73, 2008.

[7] S. Pujar, S. Chaudhari, and R. Aparna, "Survey on data integrity and verification for cloud storage," in *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Kharagpur, India, 2020.

[8] J. Zhang, S. Zhong, J. Wang, X. Yu, and O. Alfarraj, "A storage optimization scheme for blockchain transaction databases," *Computer Systems Science and Engineering*, vol. 36, no. 3, pp. 521–535, 2021.

[9] A. Tchernykh, M. Babenko, N. Chervyakov et al., "Scalable data storage design for nonstationary IoT environment with adaptive security and reliability," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 10171–10188, 2020.

[10] J. Xia, G. Cheng, S. Gu, and D. Guo, "Secure and trust-oriented edge storage for Internet of Things," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 4049–4060, 2020.

[11] S. Wang, J. Yuan, X. Li, Z. Qian, F. Arena, and I. Z. You, "Active data replica recovery for quality-assurance big data analysis in IC-IoT," *IEEE Access*, vol. 7, pp. 106997–107005, 2019.

[12] B. Mohanta, D. Jena, S. Ramasubbareddy, M. Daneshmand, and A. H. Gandomi, "Addressing security and privacy issues of IoT using blockchain technology," *IEEE Internet of Things Journal*, vol. 8, no. 2, pp. 881–888, 2021.

[13] J. Zhang, Y. Xin, Y. Gao, X. Lei, and Y. Yang, "Secure ABE scheme for access management in blockchain-based IoT," *IEEE Access*, vol. 9, pp. 54840–54849, 2021.

[14] S. Bae and Y. Shin, "An automated system recovery using blockchain," in *2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN)*, Prague, Czech Republic, 2018.

[15] V. Mishra, S. Yau, and C. Yenugunti, "Recovering decentralized critical archival data from tampering in smart city environment using blockchain," in *2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation*, Leicester, UK, 2019.

[16] J. Zhang and H. Li, "Research and implementation of a data backup and recovery system for important business areas," in *2017 9th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, Hangzhou, China, 2017.

[17] S. Cha, S. Baek, and S. Kim, "Blockchain based sensitive data management by using key escrow encryption system from the perspective of supply chain," *IEEE Access*, vol. 8, pp. 154269–154280, 2020.

[18] C. Liu, Q. Wang, X. Chu, Y. Leung, and H. Liu, "ESetStore: an erasure-coded Storage system with fast data recovery," *IEEE Transactions on Parallel and Distributed Systems*, vol. 31, no. 9, pp. 2001–2016, 2020.

[19] N. Liang, X. Zhang, H. Yang, X. Dong, and C. Zhang, "An optimal recovery approach for liberation codes in distributed Storage systems," *IEEE Access*, vol. 8, pp. 137631–137645, 2020.

[20] W. Liang, Y. Fan, K. Li, D. Zhang, and J. Gaudiot, "Secure data storage and recovery in industrial blockchain network environments," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 10, pp. 6543–6552, 2020.

[21] M. A. Ferrag and L. Shu, "The performance evaluation of blockchain-based security and privacy Systems for the Internet of things: a tutorial," *IEEE Internet of Things Journal*, vol. 8, no. 24, pp. 17236–17260, 2021.

[22] Y. Pu, C. Hu, S. Deng, and A. Alrawais, "R^2PEDS: a recoverable and revocable privacy-preserving edge data sharing scheme," *EEE Internet of Things Journal*, vol. 7, no. 9, pp. 8077–8089, 2020.

[23] Hewlett Packard Enterprise, *HP Data Protector Operations Guide*, 2020, https://support.hpe.com/hpesc/public/docDisplay?docId=emrnac02029306.