ITC 4/52

Information Technology and Control
Vol. 52 / No. 4 / 2023
pp. 1087-1101
DOI 10.5755/j01.itc.52.4.32642

Model Construction of Big Data Asset Management System for Digital Power Grid Regulation

Received 2022/11/02 | Accepted after revision 2023/09/31

HOW TO CITE: Wu, S., Yu, Y., Cheng, Y., Xu, M., Zhang, G. (2023). Model Construction of Big Data Asset Management System for Digital Power Grid Regulation. *Information Technology and Control,* 52(4), 1087-1101. https://doi.org/10.5755/j01.itc.52.4.32642

Model Construction of Big Data Asset Management System for Digital Power Grid Regulation

Silong Wu, Yangchen Yu
Financial Shared Service Center, Guangdong Power Grid Corporation, Guangzhou, China

Yongquan Cheng, Min Xu, Guanyu Zhang
YGSOFT INC. Co., Ltd, Zhuhai, China

Corresponding author: guanyu989@163.com

There are many and complex big data in digital power grid regulation, which leads to the difficulty of big data asset management. Therefore, a model of big data asset management system in digital power grid regulation is constructed. The model consists of three parts: data acquisition, data safe storage and data index. The data acquisition architecture is designed, and the data acquisition results are filled with missing values and corrected with grey prediction method. Using AR-Tree index organization to realize the digital power grid regulation big data index, and achieve the goal of high-quality management of digital power grid regulation big data assets. Store the filled and corrected data in the blockchain to ensure data security. The experimental results show that the average recall and precision of this method are 96.9% and 97.9%, and the data acquisition quality is high. After the application of this method, there is almost no unsafe data, and the proportion of safe data is higher, which shows that this method can ensure the security of big data storage. The response time of digital power grid regulation big data index is below 0.21s, and the index efficiency is higher.

KEYWORDS: Digital power grid; Digital big data; Asset management; System model; Grey prediction; AR-tree index organization.

1. Introduction

As the pillar energy and economic lifeline of the country, electric power plays an important role in the sustainable development of the national economy. In order to adapt to the national economic take-off and the pressure of resources and environment, power grid technology is developing in the direction of high voltage, large capacity, trans-regional and large power grid [3]. The implementation of digital power grid can improve the production efficiency of power grid enterprises and ensure the safety of power production through efficient data acquisition and advanced data processing and control methods [13]. As an important technical support for power grid operation, the power grid regulation and control system has been continuously improved with the development of the third-generation power grid. The system construction has gradually developed from introduction and digestion to comprehensive localization. After years of operation, it has accumulated rich and detailed data resources, which has been widely concerned by governments, academia and industry [12]. At present, the reason for the slow progress of research and application in power grid is that the improper management of power grid data collection, verification, storage and sharing has not been improved. When power grid enterprises or third-party research institutions need to access power grid data, they will encounter various obstacles, and need to spend a lot of time and resources to review permissions and verify data. There are many problems in the whole data management layer, such as different data formats, more false data, low security of data storage and transmission, etc. These problems have seriously delayed the construction of smart grid and the development of grid big data.

To solve this problem, a data asset management system based on data center is designed based on reference [8], which mainly discusses how to conduct data asset management from three aspects: data quality improvement and optimization, data asset management and data sharing service construction. Combined with the data management model, The system covers business areas such as enterprise operation, power grid operation and customer service as well as applications at all levels, providing strong support for the effective operation of the company's various businesses. However, the data acquisition quality of this method is poor. Authors in [14] designed a data asset management system based on power enterprise data center. Through the analysis of the current situation and demand of the existing data of the enterprise, the data asset management system conforming to the actual situation of the electric power enterprise is formed. Then the application of the new generation of information technology in the digital center is studied to break the information island of the information system between marketing and business. Design the functional architecture, data architecture and technical architecture of the data center according to the research results, so as to realize the comprehensive management of data. However, this method has low data storage security. Authors in [5] designed an enterprise data asset management system based on all-service unified data center of electric power. This paper analyzes the basic contents of enterprise data assets, expounds the principles of enterprise data assets management based on the actual situation, and proposes the construction ideas of enterprise data asset management system, such as clarifying the object of data asset management, establishing the framework of data asset management, setting up the standards of data asset management and refining the process of data quality management. However, this method has a long response to big data index of digital power grid regulation.

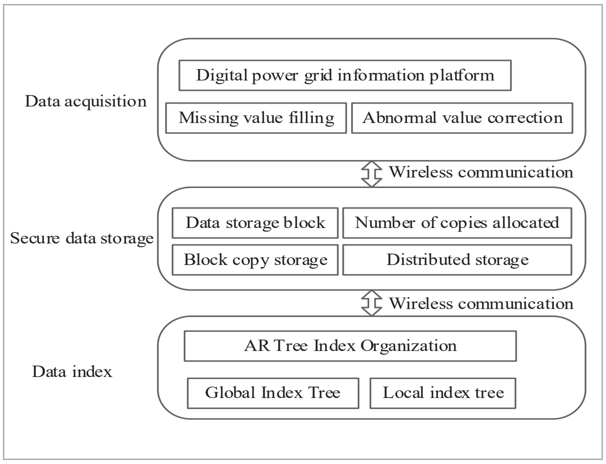
Due to poor data collection quality, low data storage security and long index response of big data for digital power grid regulation and control, the above data management methods are not suitable for managing big data asset management for digital power grid regulation and control with a large number, variety and wide distribution. Therefore, a new method to build a digital grid regulation big data asset management system model is proposed, The application effect of this method is verified by experiments. The results show that the average recall rate and precision rate of the proposed method are high, and the data acquisition quality is high. After the application of this method, there is almost no insecure data, and the proportion of safe data is higher, indicating that this method can ensure the security of big data storage. Moreover, the response time of digital power grid regulation big data index is shorter and the index efficiency is higher.

2. Model of Big Data Asset Management System for Digital Power Grid Regulation

With the development of power system, the data involved in digital power grid regulation is increasingly complex, the types of regulation big data are increasingly rich, and the amount of data is increasing. In general, the big data of digital power grid regulation can be divided into four parts: application data, collected data, attribute data and external data. There are many kinds of subsystems involved in regulating big data. In addition to the characteristics of big data such as huge data volume and numerous data types, it also has the unique characteristics of data dispersion and high security for dispatching and controlling big data. Based on the above analysis, the digital grid control big data asset management system model is designed based on the digital grid information platform, as shown in Figure 1.

Figure 1

Model of big data asset management system for digital power grid regulation

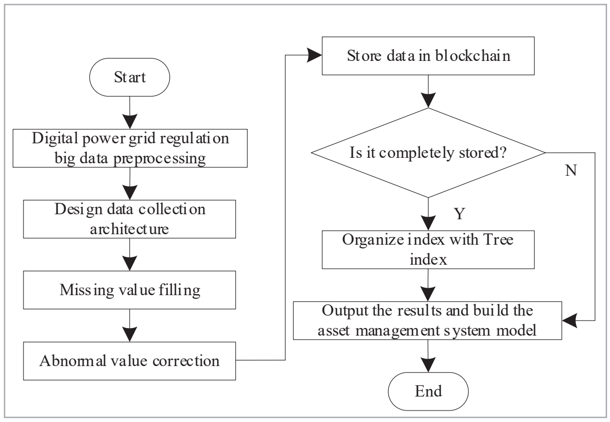


According to Figure 1, the digital grid regulation big data asset management system model consists of three parts: data collection, data security storage and data index. Based on the digital grid information platform, the data acquisition architecture is designed, and the grey prediction method is used to fill the missing values and correct the abnormal values of the data acquisition results. The filled and corrected data is stored in the blockchain through data storage blocks, block copy storage, copy quantity allocation, and distributed storage to ensure data security. The digital grid control big data global index and local index are established by using the AR-tree index organization, realize high-quality digital power grid regulation big data asset management with open, public and advanced data.

The process of building the digital grid regulation big data asset management system model is shown in Figure 2.

Figure 2

Flow Chart of Model Construction



2.1. Data Collection

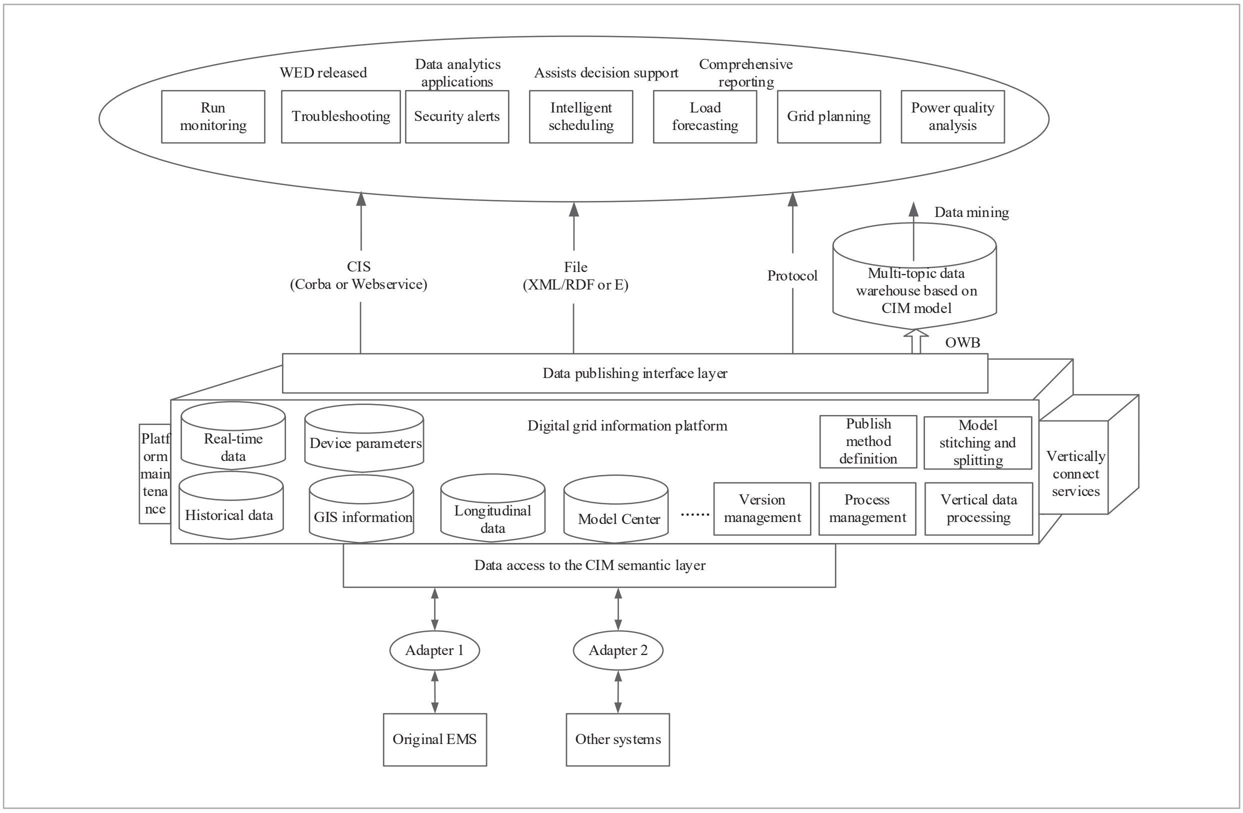
Data collection is the foundation of digital power grid regulation big data asset management system model. The specific function is to extract information related to digital power grid regulation in a period of time, including operation status, transmission quota, energy type, power price, source address, receiving time and other data information. The big data collection of digital power grid regulation not only requires real-time information, but also involves the reproduction and analysis of historical data. The architecture of digital power grid regulation big data collection can be designed by means of power grid information platform [4] and data warehouse [1], as shown in Figure 3.

For the digital power grid regulation big data acquisition architecture, the private interface is used for model transformation, and the semantic layer is formed in the information acquisition architecture for model unification. Transform the data center function of the digital grid information platform into the upper layer of data, and establish a unified model data to extract and store the digital grid regulation big data in the data warehouse, and extract effective information using the data mining technology of the data warehouse. The digital power grid regulation big data acquisition architecture provides a definable way of release, such as interface mode, file mode, protocol and mode, which facilitates data access.

Data is the foundation of digital power grid regulation and big data asset management. However, there are some missing or abnormal data in the digital power

Figure 3

Big data acquisition architecture of digital power grid regulation



grid regulation big data obtained in reality [7]. There are two main reasons for this phenomenon: one is the error of reading table of terminal equipment, deviation of data transmission and other reasons leading to abnormal data; Second, even if the data collection structure is normal, it will also cause abnormal changes in the digital power grid regulation big data due to the impact of special events, weather changes, line maintenance power outage and other factors. Therefore, it is necessary to fill in the missing values and correct the outliers in the data acquisition results.

For the processing of missing values in big data of digital power grid regulation, grey prediction method is adopted [9]. The value of the original sequence $x^{(0)}$ of the variable is the corresponding moment of the first seven days or the last seven days of the missing value. If the original number is set as $x^{(0)}(i)$, there will be $x^{(0)}(i)=\left[x^{(0)}(i-7),x^{(0)}(i-6),...,x^{(0)}(i-1)\right]$ or $x^{(0)}(i)=\left[x^{(0)}(i+1),x^{(0)}(i+2),...,x^{(0)}(i+7)\right]$. Based on the corresponding time in the previous seven days, grey model $GM(1,1)$ is established, and the grey predictive value $x^{(0)}(i)$ is finally obtained.

The processing of outliers in big data of digital power grid regulation mainly determines outliers [2]. Assume that the digital power grid regulation data series is represented by $x(i,N)$, i is the 24 periods in a day, and n is the data of N days. Then, the mean value $E(i)$ and variance $V(i)$ of the digital power grid regulation data series of N days in each period of 24 periods can be calculated:

$$E(i)=\frac{1}{N}\sum_{k=1}^{N}x(i,N) \tag{1}$$

$$V(i)=\frac{1}{N}\sum_{k=1}^{N}\left[x(i,N)-E(i)^{2}\right] \tag{2}$$

Definition $\rho(i,n)$ is the deviation rate of digital power grid regulation big data series, and the formula is:

$$\rho(i,n)=\frac{|x(i,N)-E(i)|}{\sigma_i} \tag{2}$$

In the formula, σ_i represents the actual load. Let us judge $\rho(i,n)$. When $\rho(i,n)<1.5$, the load is normal, otherwise it is abnormal. Data moving average algorithm is used to correct the data.

On the basis of formula (3), the abnormal data in the big data of digital power grid regulation are identified, and then the abnormal values are corrected according to the sliding average calculation formula (4), so that the pretreatment of the big data of digital power grid regulation can be achieved, the data can be corrected, and the accuracy of data acquisition results can be provided.

$$x^{'(0)}(t)=\frac{x^{(0)}(t-1)+x^{(0)}(t-1)+x^{(0)}(t+1)}{t} \tag{4}$$

In the formula, $x^{(0)}(t+1)$ represents the grey prediction coefficient. Formula (4) not only increases the weight of the data, but also avoids the excessive fluctuation of the value. For the calculation of two end points, the following formula can be used for calculation:

$$x^{'(0)}(1)=\frac{3x^{(0)}(1)+x^{(0)}(2)}{n} \tag{5}$$

$$x^{'(0)}(n)=\frac{x^{(0)}(n-1)+3x^{(0)}(n)}{n} \tag{6}$$

Combined with the above process, data collection, missing value filling and outlier correction are realized to ensure the quality of data collection.

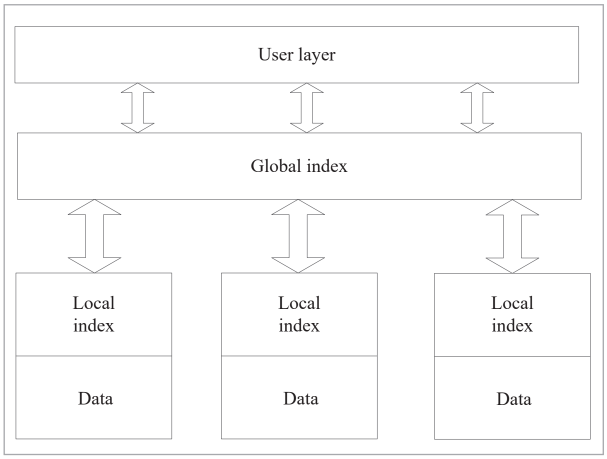
2.2. Data Index

AR-tree adopts a hierarchical design idea, which is divided into global index layer and local index layer. The local index organizes the data stored locally on the local node, and maintains the local storage range (AR for short). The global index builds an index tree again for the AR values of each local node, and together with the local index tree, forms an AR-tree index. The overall hierarchical structure of AR-tree is shown in Figure 4.

After the user gives the operation instruction of digital power grid regulation big data index, the local index pair of global index maintenance receives the operation instruction from the user and maintains the storage range AR value of the local index tree on the lower side. Local indexes are responsible for organizing data on local nodes and maintaining their own AR values. The user first indexes the first-layer global index through the access interface to locate the specific local node, and then retrieves data through the local index deployed on the local node. The choice of global index and local index is flexible and diverse. In this paper, both local index and global index use B tree to organize nodes, and the node keyword in local index is the primary key of data record, while the node keyword in global index tree is the storage range of several local indexes and the pointer to the corresponding local tree.

Figure 4

Overall layered structure of AR-tree



In the local index, the index Tree not only needs to maintain the index itself, but also needs to maintain the AR of each tree. Figure 5 shows the index organization diagram of AR-tree.

In Figure 5, group T nodes are a group of nodes on the storage cluster that store some kind of digital grid regulation big data, $T_i\,(i=1,2,\ldots,7)$. Each AR of the local index will be uploaded to the node where the global index is located, and the local node $T1$ will upload AR [100,1350] to the global index to form keywords in the global index. Therefore, the global index node key is a composite data type containing a pair of data key ranges and local index locating pointers.

C tree is a balanced retrieval tree, whose main function is to make file storage and reading more efficient.

Figure 5

Schematic diagram of AR-Tree index organization

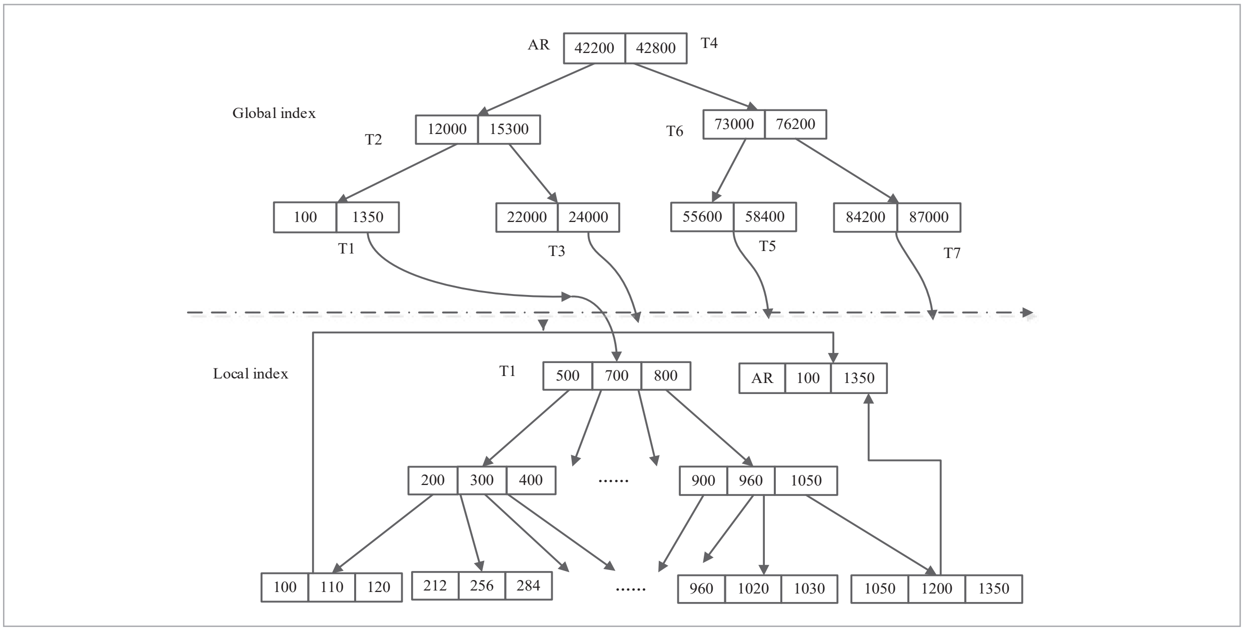
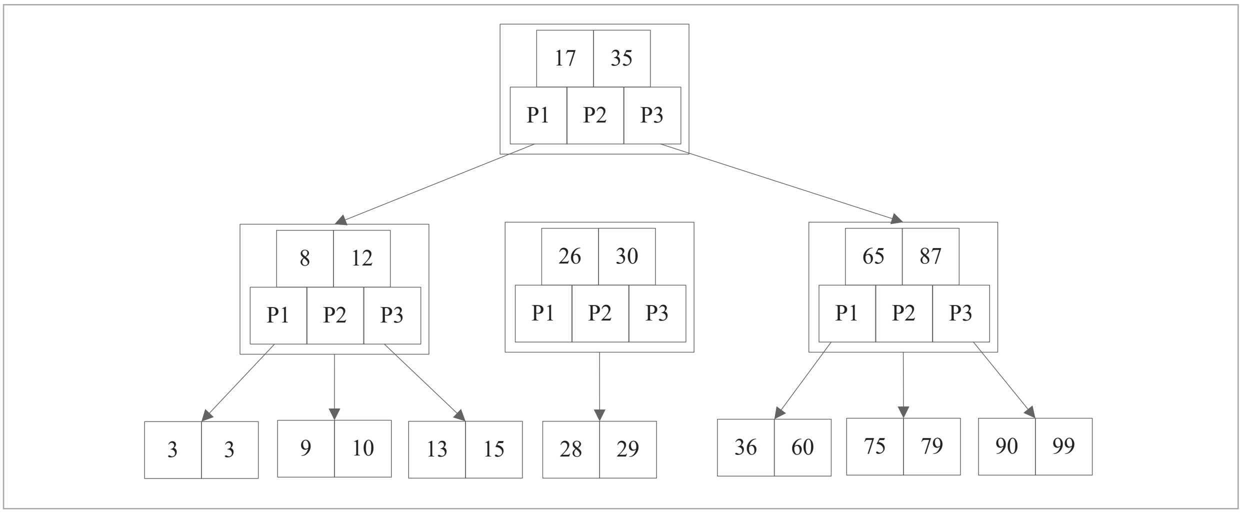


Figure 6

B tree index diagram



The B tree can reduce the number of times the disk is read during the search, thus improving the search efficiency. The B tree index is shown in Figure 6.

The top node of the B tree is the root node, that is, the root node. The node in the B tree contains the keyword data of the node. For example, node y and $y.n$ indicate the number of key words in node y. n keywords $y.key1, y.key2, \ldots, y.keyn$ are stored in non descending order, making $y.key1 \le y.key2 \le \ldots \le y.keyn$. $y.leaf$ is a Boolean value. If y is a leaf node, it is TURE; If y is an internal node, it is FLASE. Node y also stores $n+1$ pointers to its corresponding child nodes. n ordered keywords orderly separate 1 child nodes. Each leaf node has the same depth, that is, the height of the tree is h. About the height of B tree: when $n \geq 1$, B tree with n keywords stored has a minimum degree of $2n$.

For big data asset management of digital power grid regulation, index is the most basic operation and the core of AR-tree index method. Assume that key is the index key, F Tree is the first level index tree of AR-tree. This paper uses the traditional B Tree index, AR is one of the key object attributes contained in the F Tree node, and B Tree is the second level local index tree of the AR-tree index structure. The algorithm steps are:

1. First index the F Tree. If there is no AR [min, max] in the node keys in the F Tree, and the key falls in the range [min, max], the index ends;
2. If there is an AR in the node keyword of F Tree, the B Tree to which the keyword refers is indexed;
3. If there is a key in the B Tree node, the index ends;
4. The algorithm ends.

Since the data has been sorted according to the primary key when AR-tree is created, there can be at most one such AR. If there is one AR that meets the conditions. For range indexes, such as index ranges (key1, key2), the maximum and minimum value keywords key1 and key2 of the range need to be taken out, respectively, and the two keywords need to be retrieved. Finally, all keyword records between the two keywords need to be returned. The insertion operation is based on the index. If the indexed keyword already exists, it will not be inserted. On the contrary, it will be inserted according to the upward splitting principle of the B Tree. Insert algorithm description:

1. Whether the index key already exists, and if so, the insertion ends;
2. AR with key in the index. If it exists, insert the key into the response local tree, and the insertion ends;
3. If the key falls between two ARs, the insertion is completed after judgment;
4. The algorithm ends;

B tree increases upward, and before insertion, check whether the keywords in the node are full. If yes, split the node to maintain tree balance. Insert in a special case is, to insert the key word in any B Tree existing range, when a key to insert the key=10000, landed in the $T1$ of AR and $T2$ of the AR. In this case, the strategy adopted in this paper is to randomly select a local suB Tree for insertion, and once a suB Tree is inserted repeatedly, another suB Tree will be forcibly inserted next time. For AR-tree, its operation cost includes two parts: global index and local index. Suppose N pieces of data are stored in a single B Tree, the minimum degree allowed is t, t is not less than 2, and the height of the tree h meets the following inequality.

$$h_1 \leq \log_t \frac{N+1}{2} \tag{7}$$

Similarly, suppose that there are k local nodes in AR-tree, and the minimum degree in the two-layer index tree is set to t. Then the height of local index tree and global index tree h_2 and their average height h_Δ meet the following inequality.

$$h_2 \leq \log_t \frac{N+k}{2k} \tag{8}$$

$$h_\Delta \leq \log_t \frac{k+1}{2} \tag{9}$$

According to the maximum calculation value of the tree height, and from the properties of N far greater than k and log function, it is not difficult to find the following formula.

$$h_1 - \left(h_2 + h_\Delta\right) \leq \log_t \left(1 + \frac{N+1-k}{N+k+1}\right) \tag{10}$$

It can be concluded from formula (10) that the comparison times of AR-tree are less than that of B Tree in the big data indexing operation of digital power grid regulation, that is, the time cost of indexing is less. In the same way, other operation time of AR-tree is also less, so it can reduce the response time of digital power grid regulation big data index and improve the response efficiency.

A list of symbols and abbreviations is established as shown in Table 1.

Table 1

List ofsymbols and abbreviations

Symbol abbreviation	Meaning
PRBF	Practical Byzantine Fault Tolerance
ECISB	Electricity Consumption Information Storage Blockchain
POR	Packet Out of-order delivery Ratio
AR	Actual Range

2.2. Secure Data Storage

After the data obtained from the digital grid control big data acquisition architecture is aggregated and encrypted by the aggregator, the aggregation node constructs a block to load the aggregated power data into the transaction verification tree (Merkle tree) of the block body [6, 15]. The data block structure is shown in Figure 7.

It can be seen from the analysis of Figure 7 that the front block hash field records the hash value of the previous block, the block hash in the block header is the hash value allocated to it when the block is generated, the timestamp field is used to record the time when the block is generated [10], and the total power consumption field is used to record the aggregated encrypted value of multiple power consumption information collected in the current block. The Merkle root field stores the Hash value of the root node of the Merkle tree, which is a combination of the collected power consumption of the aggregation node and the Hash calculation. Since the PBFT consensus mechanism is adopted, the PBFT consensus mechanism is a state machine replica replication algorithm, that is, the service is modeled as a state machine, and the state machine replicates at different nodes of the distributed system, and the chain maintenance nodes are all selected and set as dedicated nodes, no token is required for incentive to initiate and achieve node consensus.

Since the data received by the aggregation gateway node may exceed the amount of data that can be accommodated in the block body each time the digital grid control big data is collected, it is necessary to generate multiple blocks for these data [11], apply for link up to the adjacent ECISB chain consensus maintenance node one by one, and append them to the original ECISB chain in the order of timestamp after consensus verification to achieve block copy storage. The block copy storage process is shown in Figure 8.

Figure 7
Data storage block structure

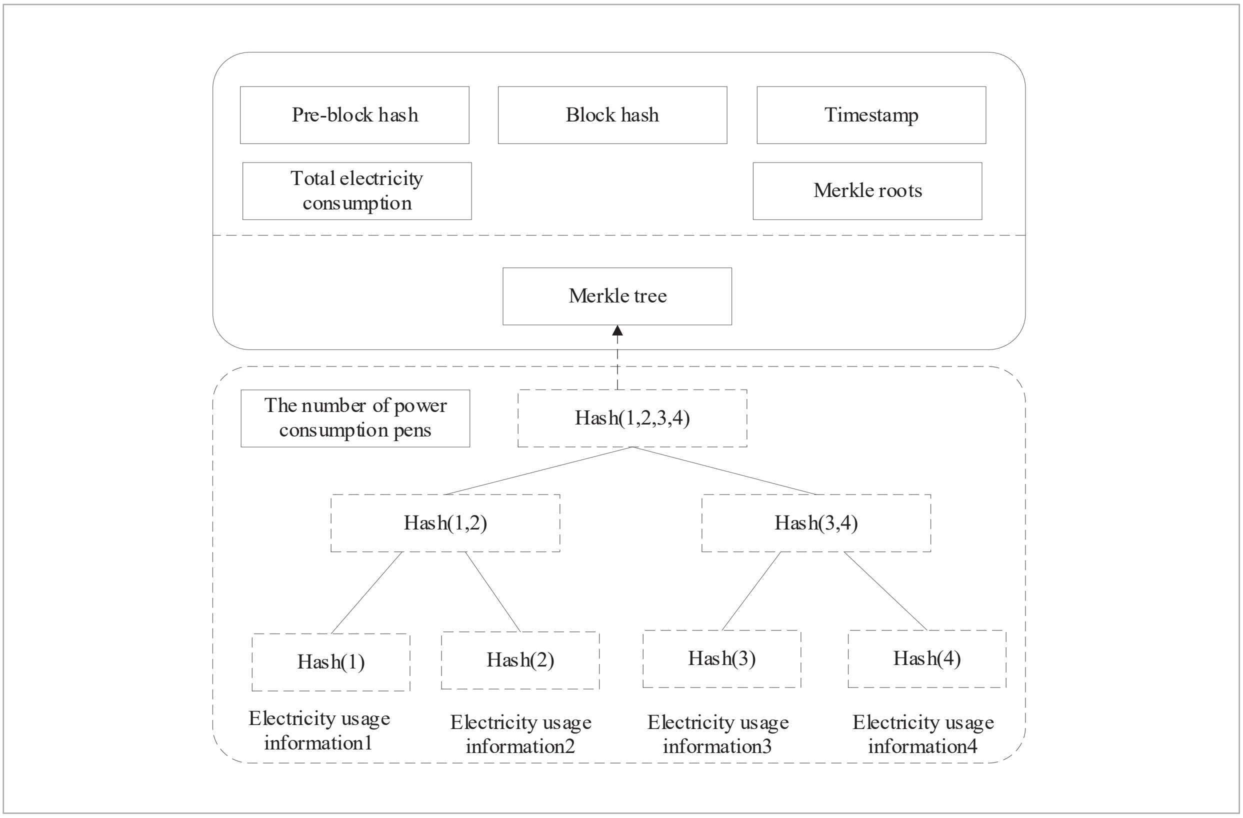
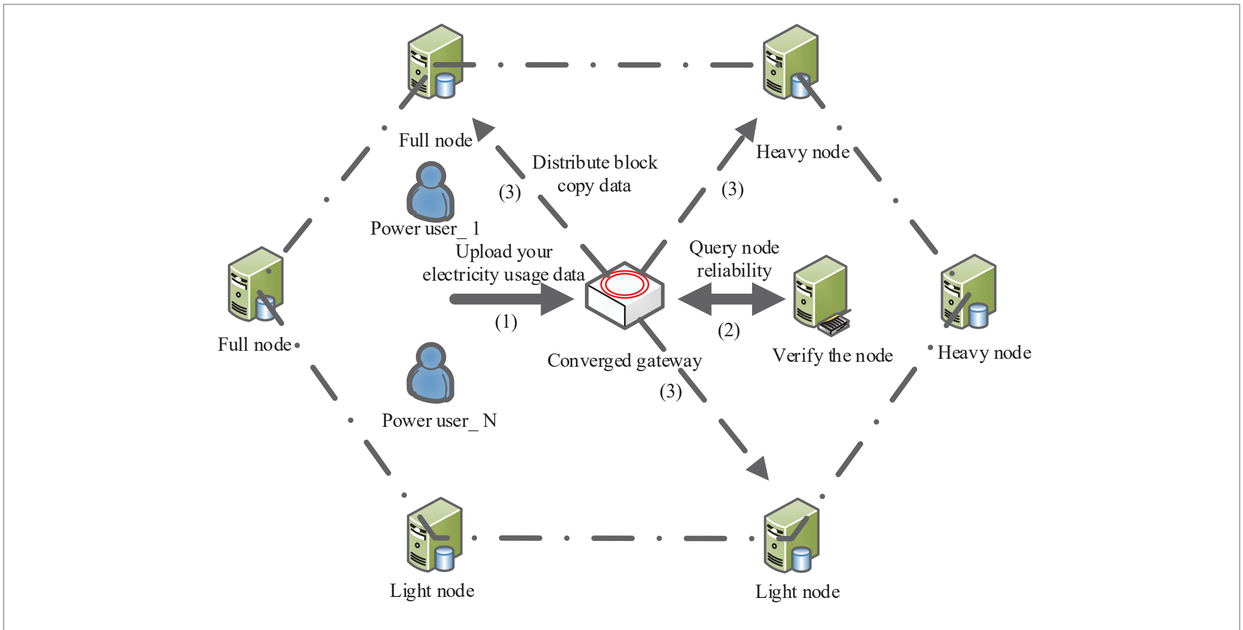


Figure 8
Block copy stored procedure



The block copy stored procedure looks like this:

Step 1: The smart meter installed at the user end and the big data acquisition architecture of digital power grid regulation will upload the data to the nearby aggregation gateway node regularly according to the data acquisition requirements, perform aggregation encryption, and generate data blocks;

Step 2: The aggregation gateway node indexes the storage nodes with high reliability in the current network to the verification node, and the verification node returns information about alternative storage nodes.

Step 3: According to the alternative nodes returned by the verification node, the aggregation gateway node copies multiple copies of block data and sends them to different storage nodes for backup storage according to the copy allocation storage policy. Meanwhile, the storage locations are recorded in the P chain.

The storage policy for copy allocation is as follows:

Each block faces different security requirements due to its location. As the length of blocks grows and the computing power of the whole network increases, it becomes more and more difficult for an attacker to tamper with block data in a way that is computationally overwritten. Therefore, the number of copies of a block needs to be assigned according to the difficulty of the block being attacked. The voltage level of nodes, especially the voltage level of junction points, is very important in power grid regulation. It plays a role in measuring the overall quality of electric energy in power grid regulation and control, and can also be reflected by the voltage level. In the network loss and voltage level, the reactive power distribution of the system and the reactive power interaction of the tie line all play a very important role. For the guarantee of voltage quality, reactive power local balance and sufficient reactive power reserve are very important, which can ensure the safe operation of digital power grid regulation big data asset management system. Suppose p is the probability of the normal "mining" node producing the next block, and q is the probability of the attack node producing the next block. Suppose that when the attack node produces the z-th block, its expected progress exists in the form of Poisson distribution:

$$\lambda = z \times \frac{q}{p} \tag{11}$$

Therefore, its total probability can be calculated by the product of Poisson probability distribution and its progress value to calculate the probability p_z of attack node's final victory:

$$p_z=\sum_{k=0}^{\infty}\frac{\lambda^k e^{-\lambda}}{k}\times\begin{cases}\dfrac{q}{p}^{(z-k)}, k\le z\\ 1, k>z\end{cases} \tag{12}$$

In the formula, λ represents the attack strength; k stands for attack frequency; To avoid the trouble of summing infinite series, the above formula can be transformed into:

$$p_z{}'=1-\sum_{k=0}^{z}\frac{\lambda^k e^{-\lambda}}{k}\times\left(1-\frac{q}{p}^{(z-k)}\right) \tag{13}$$

With the growth of the number of blocks, the possibility of attacking nodes catching up with normal "mining" nodes to produce blocks to replace the original chain is getting smaller and smaller, p_z showing a sharp decline. Therefore, it can be considered that with the increasing security of blocks on the chain, and the possibility of blocks generated earlier in the chain being tampered with is getting smaller and smaller. Therefore, in the block data copy number allocation strategy of this model, the copy backup number of the block is associated with the current location of the block, so the backup number c_k of the k block is calculated as follows:

$$c_k=\lceil P_{n-k}\times M\rceil \tag{14}$$

In formula (14), P_{n-k} is the probability when the attack node catches up with the block generation speed of the normal "mining" node when the k -th block is generated, which is also the safety factor of this block, and M is the total number of chain maintenance nodes. In the blockchain, once a block of data is recognized by more than 51% of the nodes, it will be considered as real data. Therefore, once more than 51% of the computing power in the network is controlled, the subsequent block data will be controlled. Therefore, each block, even under the premise of high security, should not only refer to formula (14), but should set a minimum value of $c_{k\min}$ for the number of copies.

According to the analysis of Borel theorem, the probability below 10-50 is regarded as an impossible event. Therefore, according to the calculation of formula (13), the number of blocks when p_z is less than 10-50 is z . At this time, it is impossible for attackers to attack block data on the chain by catching up with the original block generation speed, so z blocks can be used as a block data fragment, The minimum number of copies $c_{z\min}$ specified in the block security requirements is taken as the copy value of the partition.

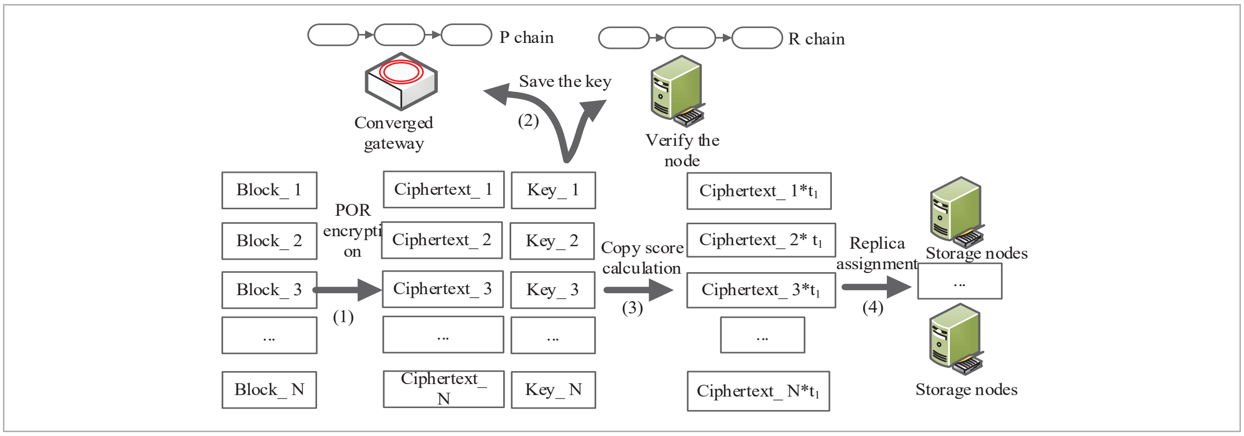
After formula (13) is substituted into formula (14) and simplified, the formula for calculating the number of storage copies of the partition can be obtained:

$$c=2^{-\left\lfloor\frac{n-t}{100}\right\rfloor}\times M \tag{15}$$

After receiving the technology of smart electricity meter and digital grid control big data acquisition architecture, the aggregation gateway node needs to package the power data into the block body and generate blocks, and allocate the appropriate number of copies to each storage node for distributed backup storage, as shown in Figure 9.

Figure 9

Distributed stored procedure



Step 1: Use the POR protocol to encrypt each block and generate the corresponding ciphertext and key.

Step 2: The aggregation gateway node and the verification node jointly save the POR encryption key, which is respectively used to recover the original data and verify the data reclaimability of the storage node when obtaining block data.

Step 3: Calculate the number of replicas to be set based on the copy allocation policy.

Step 4: Obtain the storage node with the highest reliability evaluation from the R chain maintained by the verification node, allocate the generated Nt block copy data to each storage node for backup, and record the location of the storage node in the P chain.

Table 2
Test environment

Runtime environment	Configuration	Parameter
Hardware environment	CPU	Intel(R)Core(TM) i5-9400
	Frequency	2.90GHz
	RAM	16.0GB
Software environment	Operating system	Windows 10
	Version	18362.1082 pro
	Digits	64bit
	Analog software language	APDL
	Simulation software	Matlab 7.0

3. Experimental Design

In order to verify the effectiveness of the model building method of the big data asset management system for digital power grid regulation designed in this paper, relevant experimental tests are carried out. The experimental test environment is shown in Table 2.

According to Table 2, the verification program used in the experiment is the Windows 10 emulation operating system, and its processor is Intel (R) Core (TM) i5-9400, 2.90GHz frequency, 16.0GB storage space, etc. The reference [8] method, the reference [14] method, and the reference [5] method are used as experimental comparison methods to verify the application effect of different methods by comparing the recall and precision of data acquisition of different methods, the security of big data storage of digital power grid regulation, and the index response time.

Table 3 and Table 4 show the comparison results of recall and precision of big data acquisition results of digital power grid regulation of the four methods.

By analyzing the results in Table 3, it can be seen that the maximum recall rate of the digital grid regulation big data collection results by reference [8] method is 87.7%, the average is 85.1%, and the minimum is 81.2%; The maximum recall rate of big data acquisition results of digital power grid regulation and control by reference [14] method is 79.9%, the average is 76.5%, and

Table 3
Comparison of recall rate (unit: %)

Number of experiments	Reference [8] method	Reference [14] method	Reference [5] method	Method in this paper
10	85.6	79.9	74.1	96.8
20	84.7	75.6	86.5	98.7
30	86.3	78.4	83.5	95.9
40	85.4	74.1	74.9	96.3
50	87.7	75.6	84.1	97.7
60	81.2	77.8	86.3	97.3
70	83.6	75.3	85.7	96.8
80	85.9	74.9	81.6	95.7
average value	85.1	76.5	82.1	96.9

Table 4

Comparison of precision ratio (unit:%)

Number of experiments	Reference [8] method	Reference [14] method	Reference [5] method	Method in this paper
10	84.7	79.8	71.4	98.7
20	84.9	78.7	73.6	98.4
30	86.9	74.6	74.4	97.8
40	85.5	85.6	78.9	97.6
50	84.7	71.3	85.6	98.5
60	82.6	72.6	84.7	99.3
70	81.7	74.9	81.2	95.6
80	85.4	75.6	83.1	97.1
average value	84.6	76.6	79.1	97.9

the minimum is 74.1%; The maximum recall rate of big data acquisition results of digital power grid regulation and control by reference [5] method is 86.5%, the average is 82.1%, and the minimum is 74.9%; The maximum recall rate of the digital grid regulation big data acquisition results of this method is 98.7%, the average is 96.9%, and the minimum is 95.7%. This shows that compared with the reference method, the recall rate of this method is higher, and the digital grid regulation big data acquisition results are more complete. The reason is that this method uses POR protocol to encrypt each block and generate corresponding ciphertext and key; The POR encryption key is saved by the aggregation gateway node and the verification node, which are used to recover the original data and verify the data retrievability of the storage node when obtaining block data, respectively, which is conducive to increasing the recall rate and achieving better results.

By analyzing the results in Table 4, it can be seen that the maximum precision of the digital grid regulation big data collection results using the reference [8] method is 86.9%, the average is 84.6%, and the minimum is 81.7%; The maximum precision of big data acquisition results of digital power grid regulation and control by reference [14] method is 85.6%, the average is 76.6%, and the minimum is 71.3%; The maximum precision of big data acquisition results of digital power grid regulation and control by reference [5] method is 85.6%, the average is 79.1%, and the minimum is 71.4%; The maximum precision of the digital grid regulation big data acquisition results of this method is 99.3%, the average value is 97.9%, and the minimum value is 95.6%. This shows that compared with the reference method, the precision of this method is higher, and the digital grid regulation big data acquisition accuracy is higher. The reason is that under this method, users first index the first level global index through the access interface to locate specific local nodes, and then retrieve data through the local index deployed on the local nodes, which is conducive to increasing the precision rate.

The storage security of digital power grid control big data of four methods is compared. The higher the proportion of safe data in the stored data, the higher the data security. The comparison results are shown in Figure 9.

By analyzing the comparison results of the storage security of the digital grid regulation big data in Figure 10, we can see that the security data of the reference [8] method, the reference [14] method and the reference [5] method account for about half of the total, which indicates that after the application of these three methods, the storage security of the digital grid regulation big data is low, the data is easy to be leaked, and the privacy and security of the digital grid regulation big data cannot be guaranteed. However, after the application of this method, there is almost no non security data, and the proportion of security data is higher, which shows that this method can ensure the storage security of big data of digital power grid regulation, and ensure that it will not be attacked or leaked, and the data security is good.

The response time of digital power grid regulation big data index of different methods is compared, and the comparison results are shown in Figure 11.

Figure 10

Storage security of big data for digital power grid regulation

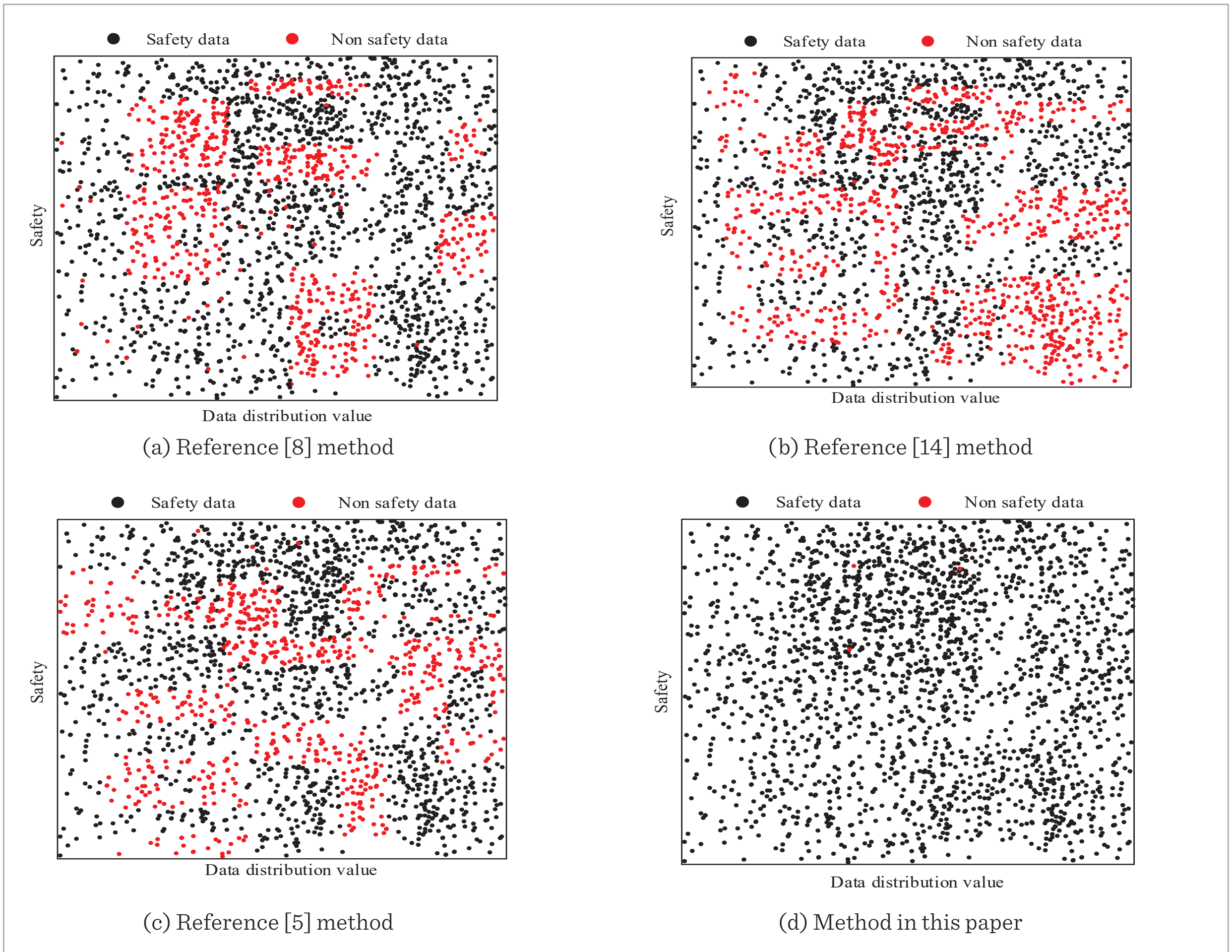
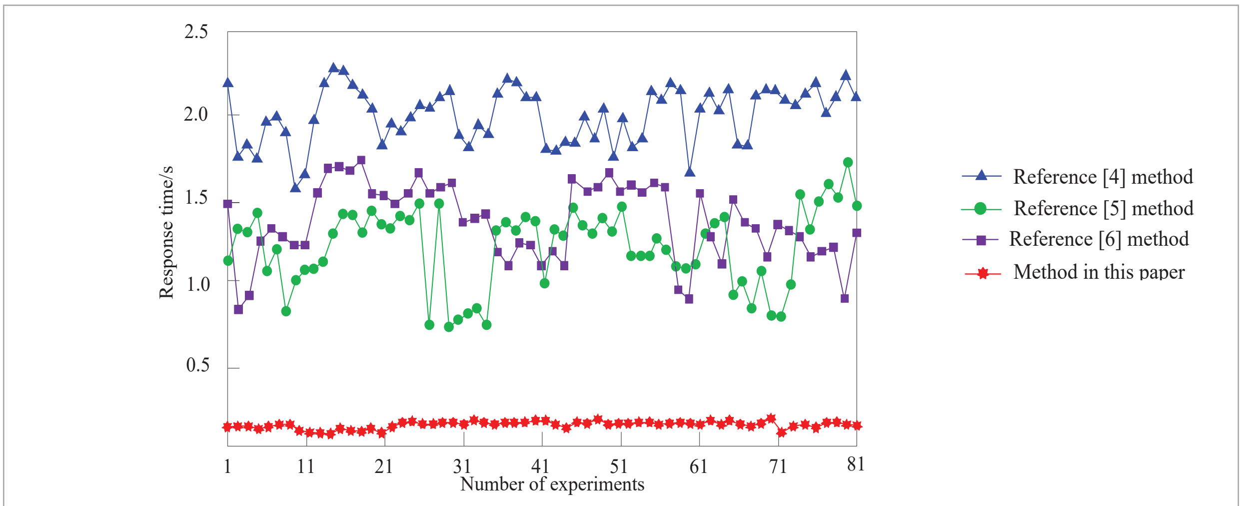


Figure 11

Response time of digital grid regulation big data index



By analyzing the results in Figure 11, it can be seen that the response time of the digital grid regulation big data index of reference [8] method varies from 1.64 seconds to 2.27 seconds, the response time of the digital grid regulation big data index of reference [14] method varies from 0.71 seconds to 1.76 seconds, and the response time of the digital grid regulation big data index of reference [5] method varies from 0.79 seconds to 1.72 seconds, The response time of the digital grid regulation big data index of this method is less than 0.21s, which shows that the response time of the digital grid regulation big data index of this method is shorter and the efficiency is higher. The reason is that the method in this paper generates multiple blocks for the big data of digital power grid regulation, applies to the adjacent ECISB chain consensus maintenance node for chain-up one by one, and appends them to the original ECISB chain according to the time stamp sequence after the consensus verification, so as to realize the storage of block copies, which is conducive to shortening the indexing time to a certain extent and making the efficiency higher.

4. Discussion

The operation process of digital power grid regulation big data involves massive and complex multi-source heterogeneous data, which constitutes regulation big data. With the help of data mining technology, the data in the asset management system model will be transformed into intelligent scheduling development trend data. Under this research background, build a digital power grid regulation big data asset management system model. It aims to solve the problem of rapid and effective analysis, processing and extraction of massive data in the big data of power grid regulation and control to explore useful rules and knowledge, bring a large number of innovative applications to the power grid and provide auxiliary decision-making for model construction. The innovation points of this method are summarized as follows:

1. The digital power grid regulation big data asset management system model consists of three parts: data collection, data security storage and data index.
2. Design the data acquisition framework, and use the grey prediction method to fill in the missing values and correct the abnormal values of the data acquisition results. Store the filled and corrected data in the blockchain to ensure data security.
3. Utilize the AR-Tree index organization to realize the digital power grid regulation big data index, and achieve the goal of high-quality management of digital power grid regulation big data assets.

5. Conclusion

With the popularization and deepening of information technology, more and more attention has been paid to the digital power grid. The future development direction is to build a real-time digital power grid with clear structure, reliable operation and comprehensive information, which can meet the various requirements of decision support for the safe operation of power grid and provide information support for the safe operation of power grid. Control system of power grid dispatching service scope expanding rapidly, various types of data, and the increase in the number of data sources data size and processing pressure increases sharply, regulating the business requirement for massive multi-source data analysis, system is facing new challenges, so this article constructs the digital grid control large data assets management system model, Through experiments, it is verified that the proposed method has certain advantages in digital power grid regulation big data collection, storage security and index response time, which can solve the problems existing in the traditional management system model and promote the further improvement of digital power grid regulation big data asset management level.

The future research work can be mainly carried out according to the following aspects:

1. Improve the intelligent level of digital power grid regulation: With the rapid development of artificial intelligence, Internet of Things and big data, we can further explore how to apply these cutting-edge technologies to improve the intelligent level of digital power grid regulation in the future.
2. Strengthen the construction of big data asset management system: Big data plays an important role in the regulation of digital power grid, so it is necessary to establish a sound big data asset management system. Future research can focus on how to optimize the way of data collection, storage and processing, and establish an effective data security and privacy protection mechanism.
3. Promote the integrated development of sustainable energy and power grid: With the rapid development of renewable energy, digital power grid regulation

and control should be combined with sustainable energy development in the future to achieve efficient use of energy and low-carbon emissions.

Data Sharing Agreement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgement

The authors received no financial support for the research, authorship, and/or publication of this article.

Conflict of Interest

The authors have no conflicts of interest to declare.

References

1. Bensalloua, C. A., Benameur, A. Towards NoSQL-based Data Warehouse Solution integrating ECDIS for Maritime Navigation Decision Support System. Slovenian Association Informatika, 2021, 45(3), 415-431. https://doi.org/10.31449/inf.v45i3.3204
2. Bin, D., Yu, T., Ling, Y., Chen, W., Lu, L., Chen, H. Power Distribution Abnormal Data Identification and Processing System Based on Correlation Classification. IOP Publishing Ltd, 2022, 28(2), 1566-1578.
3. Boopathi, D., Saravanan, S., Jagatheesan, K., Anand, B. Performance Estimation of Frequency Regulation for a Micro-Grid Power System Using PSO-PID Controller. International Journal of Applied Evolutionary Computation (IJAEC), 2021, 12(2), 36-49. https://doi.org/10.4018/IJAEC.2021040103
4. Cheng, P., Li, K., Wu, C., Ma, J., Jia, L. Flexible Power Regulation and Limitation of Voltage Source Inverters Under Unbalanced Grid Faults. China Electrotechnical Society Transactions on Electrical Machines and Systems, 2022, 6(2), 153-161. https://doi.org/10.30941/CESTEMS.2022.00021
5. Cheng, P. F., Huang, B. Analysis of Enterprise Data Asset Management System Based on Power All Business Inified Data Center. Modern Industrial Economy and Informationization, 2021, 11(2), 103-104.
6. Comuzzi, M., Cappiello, C., Daniel F., Meroni, G. Toward Quality-Aware Transaction Validation in Blockchains. IEEE Software, 2022, 39(4), 54-62. https://doi.org/10.1109/MS.2022.3152176
7. Hemanth, G. R., Raja, C. S. Proposing Suitable Data Imputation Methods by Adopting a Stage Wise Approach for Various Classes of Smart Meters Missing Data -Practical Approach. Expert Systems with Application, 2022, 187(1), 1-20. https://doi.org/10.1016/j.eswa.2021.115911
8. Li, S., Xu, Q., Xia, Y., Hua, K. Comprehensive Setting and Optimization of Dead-Band for BESS Participate in Power Grid Primary Frequency Regulation. International Journal of Electrical Power and Energy Systems, 2022, 141(10), 1-14. https://doi.org/10.1016/j.ijepes.2022.108195
9. Li, W., Li, B., Fang. R., You, P., Zou, Y., Xu, Z., Guo, S. Risk Evaluation of Electric Power Grid Enterprise Related to Electricity Transmission and Distribution Tariff Regulation Employing a Hybrid MCDM Model. Mathematics, 2021, 9(9), 989-999. https://doi.org/10.3390/math9090989
10. Malan A.J., Pol, J.-S., Strehle F., Hohmann, S. Passivity-based Power Sharing And Voltage Regulation in DC Microgrids with Unactuated Buses. arXiv e-prints, 2023, 31(1),1-10.
11. Nakamura, Y., Aoki, M., Kado, Y., Iki, H. Study on Operation in Microgrid Including Hydrogen Supplying Equipment for Contributing Power System Operation. Electrical Engineering in Japan, 2021, 214(2), 1-14. https://doi.org/10.1002/eej.23329
12. Wu, Y., Liu, J., Zhou, J. The Strategy of Considering the Participation of Doubly-Fed Pumped-Storage Units in Power Grid Frequency Regulation. Energies, 2022, 15(6), 2179-2185. https://doi.org/10.3390/en15062179
13. Zareipour, H. Digital Power Grid, Lessons from History to Imagine the Future [In My View]. IEEE Power & Energy Magazine, 2022, 20(3), 90-92. https://doi.org/10.1109/MPE.2022.3150817
14. Zhang, S.C., Yang, C.S., Zhu, H.D., Zhao, H., Zhao, J., Zhang, T.Y. Research on Data Asset Management System Based on Data Center of Electric Power Enterprise. Power Systems and Big Data, 2022, 25(1), 51-58.
15. Zhao, L., Lee, S., Jeong, S. P. Decision Tree Application to Classification Problems with Boosting Algorithm. Electronics, 2021, 10(16), 1903-1910. https://doi.org/10.3390/electronics10161903