Research on Monitoring Data Security Sharing Method for Hydropower Station Operation and Maintenance

Research on Monitoring Data Security Sharing Method for Hydropower Station Operation and Maintenance

Yuanjiang Ma, Liang Hong*, Taisong Qin, Xia Hao, Taiquan Tan

Yingxiuwan Hydropower General Plant of State Grid Sichuan Electric Power Company, Chengdu Sichuan 611800, China

*leon_hong@hotmail.com

Abstract. The computer monitoring system centrally and uniformly manages the hydropower stations in the basin, which has high operation efficiency. The current data security sharing methods can not solve the multi-party trust problems such as data ownership and data security. In this regard, a monitoring data security sharing method for hydropower station operation and maintenance is proposed. The monitoring data of hydropower station comes from multiple independent operation systems developed by different suppliers according to different application requirements. The multi-source data are integrated and processed to meet the interactive needs of users. User permission has the characteristics of dynamic change, which can identify the permission attributes. Create a trusted container based on the user ID, locate the monitoring data, and trace the source in data sharing and storage. The access policy is formulated by the data user and bound with the private key, while the data is given an attribute set during encryption. Each data user has a corresponding access policy and a private key corresponding to the access policy, so as to realize the security sharing of monitoring data. The test results show that the design method in this paper can effectively reduce the communication overhead and storage overhead, and has good computing performance.

1. Introduction

Electric power is the important energy base of economic development, and the economic development has brought the increase of electricity consumption. At present, China's electricity is mainly thermal power and hydropower. China is rich in hydropower resources, which is less polluted than thermal power. As a result, hydropower accounts for more than 50% of China's electricity supply. The hydropower enterprise's management pattern also unceasingly innovates along with the time progress. The centralized hydropower station group can design and develop the secondary capacity expansion of the system according to its own operation and maintenance characteristics and personnel needs. The operation and maintenance managers of hydropower station usually record the operation and maintenance information by traditional manual recording method. This kind of recording method needs a lot of work, but it also has some disadvantages, such as difficult to guarantee the integrity and accuracy of data, difficult to share and save information. Therefore, in many large data scenarios, sharing data is essential. The strategy and importance of shared data has also led to the emergence of operations related to the entire life cycle of shared data, from data collection to data cleansing, desensitization, data monitoring, data analysis and data storage [1]. In the centralized control mode, the hydropower station should adopt the automation technology as far as possible to operate and maintain, reduce the manual participation. Through the introduction of advanced equipment and the use of computer monitoring system, hydropower stations in the basin are managed in a centralized and

unified manner. This paper presents a method of sharing monitoring data of hydropower station operation and maintenance safely, which aims to provide encrypted protection for user information and rationally manage shared data. Through the establishment of digital identification of shared data to ensure user data sovereignty and data integrity, improve the efficiency of shared data.

2. Monitoring data security sharing method for hydropower station operation and maintenance

2.1. Hydropower station operation and maintenance monitoring data integration

In order to improve the intelligent level of hydropower station, it is urgent to effectively integrate the data from various sources and build an integrated monitoring data platform to provide data support for the whole intelligent hydropower station. Various business function modules of the information platform can be combined freely according to the actual needs to avoid the waste of resources; Different intelligent electronic devices (IED) can carry out fast and accurate information communication, use artificial intelligence technology for intelligent analysis and comprehensive processing of power plant events, and provide man-machine interface to meet the needs of user interaction [2-3]. In order to meet the diversified needs of various engineering applications of intelligent hydropower station for state data, six levels of state data are formed through the analysis and statistical calculation of different time scales, as shown in Figure 1.

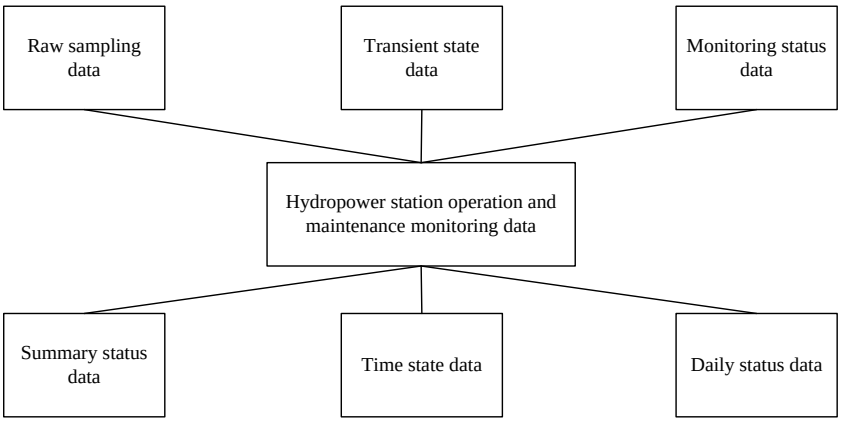

Figure 1. Composition of operation and maintenance monitoring data of hydropower station

Data preprocessing and data unification is composed of data preprocessing module and intelligent data application platform. Data preprocessing is mainly used for data extraction, classification, standardization and storage of multi-source information. Information is mainly divided into offline information and online information. The local server belongs to the field server and communicates with PLC by Modbus TCP. The lower PLC accesses through Modbus TCP protocol to obtain the sensor data of each control point. Modbus TCP protocol communication adopts C/S mode structure, the local server is the client, and the lower controller PLC is the server [4]. The monitoring data of hydropower station are typical data. In the process of data integration, the relationship between multi-source and mass data is established. Multi-level data association can be described by a six-tuple representation: {original sampling data, transient data, monitoring status data, summary status data, time status data, and daily status data}. The PLC also responds to the local server as a message, reading the content of the message and interpreting the message to get the desired data [5]. The original sampling data, transient data and performance index data of the working process are uploaded to the synthesis unit in turn.

2.2. User permission attribute identification

The user authority has the characteristics of dynamic change. When the user attribute changes or the user exits the system, it is very important for the system security to revoke the user attribute or private

key. The purpose of access control is to ensure that data information is not accessed and used by unauthorized users by limiting users' access rights. However, due to the changes in the storage mode of data in the cloud environment, users cannot actually control the data stored in the cloud server. The reliability of the cloud server and the migration technology and virtualization technology in the cloud storage will lead to changes in the security attributes of user data. The purpose of sending the user ID is that after the user creates the container, the server remotely deploying the blockchain network can still remotely access the container [6]. When identifying the user authority attribute, the proxy decryption server upgrades the key component of the attribute corresponding to the user private key associated with the revoked attribute, which is expressed as:

$$p' = \left(p, \alpha^{\beta t / \chi}, \lambda \right) \tag{1}$$

In formula (1), p and p' represent user attributes before and after upgrading respectively; α represents safety parameters; β means randomly selected elements; t represents time; χ indicates public attribute; λ represents the key component. The data owner uses the public key (identity) to encrypt the data and outsource it to the proxy server. The data owner authenticates the requester, independently generates the re encryption key and sends it to the PS. The requester with access permission can decrypt the re encrypted ciphertext. In this process, the current attribute key can be expressed as:

$$\lambda = \left(h\alpha^{\varphi} \right) \beta t / \chi \tag{2}$$

In formula (2), h represents hash function; φ indicates that a new version is randomly generated by user attributes. If the decrypted string is the same, it indicates that the user is a legal user. In this scheme, when the user needs to create a container to operate the data file, the data security sharing platform generates three files for a specific user: id_ rsa. pub (public key file), id_ rsa (private key file), name. The id_ rsa. pub is stored in KC key management center, id_ rsa (private key) and name files are transferred into the container. The name file is the identification field and plays an index role in the subsequent user identification link of the chain platform. By updating the conversion key and the attribute version in the ciphertext involving the revoked attribute, the system attribute revocation is realized to ensure the forward security of the ciphertext before revocation.

2.3. Source location of monitoring data

Operation maintenance monitoring of hydropower station has certain requirements for the storage location of files. Users need to locate and trace the source of monitoring data in the process of data sharing and storage to facilitate classification, query and statistics. The read time depends on the performance of the storage device and the number of data blocks read, while the calculation time depends on the overhead of the location scheme and CPU performance. Therefore, the cloud data location scheme design needs to consider the above factors to reduce read time and computational time. Naive Bayesian classifiers are trained for data location. In the learning phase, you need to use routing to calculate IP scores and routing scores. The IP scores are calculated as follows:

$$F_1 = \sum_{x=1}^{m} \frac{1}{e^{\mu_x - \nu_x}} \tag{3}$$

In formula (3), F_1 represents IP score; m and x represent the set and serial number of all IP respectively; e is the natural constant; μ_x indicates the number of IP addresses included in the destination route; ν_x indicates the number of hops from the source in this route. IP score is the

weighted frequency of IP in the route to the target location. In case of modification, the data to be exported is the newly generated data file, and the user ownership needs to be changed. Convert the route into numerical data, and the calculation formula of route score is as follows:

$$F_2 = \sum_{x=1}^{m} x \frac{F_1}{e^{n_1 - n_2}} \tag{4}$$

In formula (4), F_2 represents the routing score; n_1 indicates the number of IP addresses in the route; n_2 represents the weighted sum of IP scores of all IP addresses. Through the routing score, the probability density distribution is calculated by using kernel density estimation. The formula is as follows:

$$\vartheta = \frac{1}{mb} \sum_{x=1}^{m} \frac{x e^{-(F_2 - \tau)^2}}{\sqrt{2\pi}} \tag{5}$$

In formula (5), ϑ represents the probability density; b indicates bandwidth; τ indicates the delay between bits. At the same time, attribute revocation mechanism can ensure forward security and backward security. The pre decryption is outsourced to the server for calculation, and the user's attribute key needs to be sent to the server. Among the components of time delay, the reading time and calculation time are not constant, and only the network transmission time is related to the geographical location. Therefore, the larger the proportion of reading time and computing time, the smaller the correlation between time delay and geographical location.

2.4. Establish monitoring data security sharing model

In order to ensure the safe sharing of hydropower station operation and maintenance monitoring data, we should protect the whole process of data sharing from the whole data sharing process, including data provider sharing data, storage element storing data, data requester downloading data and data requester downloading data. The user access layer is a user oriented operation layer, which provides the logical realization of the main functions of the model. The meta information set is composed of data acquisition meta information, data name and other information. It can be expressed as the following formula when registered in the data management contract:

$$W = \{N, K, RP, C\} \tag{6}$$

In formula (6), W represents the meta information set; N indicates the name of meta information data; K represents data keyword; RP indicates ciphertext request access policy; C indicates encrypted information. The encryption and decryption module provides access control, and provides the acquisition of information required for encryption and decryption operation and encryption and decryption operation. The data owner only needs to care about the attribute set of the data itself, and then encrypt the data based on the attribute set and upload it to the cloud storage center. Each data user has a corresponding access policy and a private key corresponding to the access policy. The controlled requester requests the user key from the attribute authority, which is expressed as:

$$(\varpi_1, \varpi_2) = u(d, s_1, s_2, z) \tag{7}$$

In formula (7), ϖ_1 represents the user key; ϖ_2 indicates relevant information; u represents the request key of the attribute institution; d indicates dynamic access policy; s_1, s_2 are the encrypted information and public key; z indicates the user. Attribute information query and target information

encryption are the services that policy makers will provide for target information encryption in the access control mechanism. The key request and ciphertext decryption provide services for the controlled requester in the access control mechanism to obtain the target information. Firstly, the user applies for the key from all attribute institutions of the user's registered attribute through the key request, and decrypts the ciphertext with the key to obtain the target information. So far, the design of monitoring data security sharing method for hydropower station operation and maintenance has been completed.

3. Experiment

3.1. Experimental preparation

The experiment takes the monitoring data of Yingxiuwan, Yuzixi and Gengda Hydropower Stations of Yingxiuwan hydropower plant of state grid sichuan electric power company as the research object, and uses the monitoring data security sharing method for hydropower station operation and maintenance proposed in this paper to store and share the data. The server is responsible for offline data analysis and the release and push of monitoring data and analysis results at the plant and station level. In practical engineering, the functions of multiple servers can be integrated into one server to reduce the hardware cost. The server configuration is as follows: processor 2.5GHz, dual core Intel Core i7, memory capacity 8GB, operating system Centos7.4, 64 bit, development language environment Golang 1.11. The data owner encrypts the data through the terminal and uploads it to the cloud storage server, and the data user downloads and decrypts the data through the cloud. During data decryption, the hydropower station operation and maintenance platform outsources part of the calculation with large calculation cost to the proxy decryption server to reduce the calculation cost of data users. At this time, the proxy decryption key server completes part of the decryption work.

3.2. Experimental results and analysis

In order to test the performance of the monitoring data security sharing method for hydropower station operation and maintenance proposed in this paper, the communication overhead and storage overhead are calculated. The test results are compared with the monitoring data security sharing methods based on homomorphic encryption and machine learning. The experimental comparison method is Security sharing method of monitoring data based on homomorphic encryption (Method 1) and Safety sharing method of monitoring data based on machine learning (Method 2). The comparison of communication overhead of each data security sharing method is shown in Table 1.

Table 1. Comparison of communication overhead (ms)

Number of tests	The method of this paper	Method 1	Method 2
1	129.54	156.64	183.59
2	135.68	165.58	196.88
3	128.86	168.86	188.67
4	139.02	172.45	185.26
5	142.25	175.22	199.55
6	126.59	170.31	203.82
7	128.73	182.07	216.62
8	132.46	183.54	202.54
9	134.15	176.92	195.47
10	135.82	195.23	185.08

It can be seen from Table 1 that the communication overhead of the monitoring data security sharing method for hydropower station operation and maintenance designed in this paper is 133.31ms, which is 41.37ms and 62.44ms lower than the data security sharing method based on homomorphic

encryption and machine learning. Compared with the two comparison schemes, this method has obvious advantages in communication overhead. That is, the proposed scheme has smaller system parameters, key length and ciphertext size. The comparison of storage overhead of each data security sharing method is shown in Table 2.

Table 2. Comparison of storage overhead (ms)

Number of tests	The method of this paper	Method 1	Method 2
1	204.46	466.46	485.47
2	211.54	455.88	497.84
3	206.61	420.67	518.61
4	205.39	481.34	526.58
5	213.58	452.52	495.25
6	221.85	465.25	482.86
7	204.27	437.58	495.53
8	206.04	498.83	508.62
9	215.91	456.12	506.51
10	212.22	483.21	493.26

It can be seen from Table 2 that the storage cost of the monitoring data security sharing method for hydropower station operation and maintenance designed in this paper is 210.19ms, which is 251.60ms and 290.86ms lower than the data security sharing method based on homomorphic encryption and machine learning. Therefore, this design method significantly optimizes the storage overhead and achieves better computational performance than other schemes.

4. Conclusion

In order to develop and utilize hydropower, China has built more hydropower stations in different regions. Hydropower stations use the stored water energy to convert it into electric energy. The development of big data artificial intelligence technology breaks through the limitations of traditional information technology for specific applications and penetrates the data information generated in all links of power operation and maintenance. It promotes the efficient collaboration among multiple businesses of the power plant, and can realize the transformation of data resources into effective data assets. This paper presents a method of monitoring data security sharing for hydropower station operation and maintenance. This method can optimize the communication and storage overhead and achieve better computing performance. In the process of data sharing, we can also optimize the access control algorithm and encryption algorithm in time and space. This is also a direction that needs to be optimized in the future.

Acknowledgments

The study was supported by “State Grid Sichuan electric power company science and technology project support(521901170004)”.

References

[1] Cheng, L. J., Qi, Z. H., Shi, J. C. (2020) Blockchain based secure storage and sharing scheme for EHR data. Journal of Nanjing University of Posts and Telecommunications (Natural Science), 40(4): 96-102.
[2] Resende, J. S., Magalhães, L., Brandão, A. (2021) Towards a Modular On-Premise Approach for Data Sharing. Sensors, 21(17): 5805-5805.
[3] Chen, Y. W., Hu, B. W., Yu, H. J. (2021) A Threshold Proxy Re-Encryption Scheme for Secure IoT Data Sharing Based on Blockchain. Electronics, 10(19): 2359-2359.

[4] Rajkumar, V., Prakash, M., Vennila, V. (2022) Secure Data Sharing with Confidentiality, Integrity and Access Control in Cloud Environment. Computer Systems Science and Engineering, 40(2): 779-793.
[5] Chen, B., Lu, Wei. Huang, J. W. (2022) Secret Sharing Based Reversible Data Hiding in Encrypted Images With Multiple Data-Hiders. IEEE Transactions on Dependable and Secure Computing, 19(2): 978-991.
[6] Walker, D. M., Hefner, J. L., DePuccio, M. J. (2022) Approaches for overcoming barriers to cross-sector data sharing. The American journal of managed care, 28(1): 11-16.