Research Article

Wenjuan Yang, Zhongbin Chan*, Yi Wang, and Fuli Qi

# Application of big data technology in electromechanical operation and maintenance intelligent platform

https://doi.org/10.1515/pjbr-2022-0121
received May 8, 2023; accepted June 19, 2023

**Abstract:** Aiming at the data preprocessing requirements and label data cost issues arising from the intelligent operation and maintenance of electromechanical equipment, this article mainly studies structured data cleaning methods and fault prediction algorithms for a small number of label samples. First, this article introduces the overall architecture of the intelligent operation and maintenance system for electromechanical equipment. Second, based on the electromechanical equipment operation and maintenance data access service, data cleaning, and fault prediction, this article constructs an electromechanical equipment intelligent operation and maintenance platform based on Kafka message queue, Spark cluster, and other components, and introduces the functional composition of the system in detail. Finally, the article describes the functions of each component of data access service, data cleaning, and fault prediction in detail. To address the cost issue associated with sufficient labeled sample data for data analysis, we propose a semi-supervised learning algorithm, IF-GBDT, based on improved independent forests and Gradient Boosting Decision Tree. The independent forest algorithm supplements labels for unlabeled data based on the learning results of a small number of labeled samples. We also use the gradient lifting tree algorithm to train the model based on the new tag data set for fault prediction, thereby reducing the impact of lack of tags on the accuracy of the prediction model. Experiments show that this method improves classification accuracy and has good adaptability and concurrency performance for a small number of tags.

**Keywords:** big data technology, electromechanical operation and maintenance intelligent platform, data cleaning, fault prediction, small amount of labels

* **Corresponding author: Zhongbin Chan,** School of Information Engineering, Shanghai Zhongqiao Vocational and Technical University, Shanghai, 201514, China, e-mail: ZhongbinChan@126.com
**Wenjuan Yang:** School of Information Engineering, Shanghai Zhongqiao Vocational and Technical University, Shanghai, 201514, China, e-mail: WenjuanYang9@163.com
**Yi Wang:** School of Information Engineering, Shanghai Zhongqiao Vocational and Technical University, Shanghai, 201514, China, e-mail: YiWang877@163.com
**Fuli Qi:** School of Information Engineering, Shanghai Zhongqiao Vocational and Technical University, Shanghai, 201514, China, e-mail: FuliQi7@126.com

## 1 Introduction

The application of big data in the acquisition system includes data acquisition and number Data cleaning, data storage, data analysis, data processing, data interpretation, and Data application 6 links. This article aims at the data analysis link. The relational online analysis processing mechanism and multidimensional analysis technology are further optimized. The whole system is discussed: Big Data operations architecture collects, collects, and markets and operations data through Data Extraction Transformation and load (Extraction–Transformation Loading, ETL) classifying data adapter layer and the data storage. The layer provides data, and the data storage layer classifies and stores data backward. The data analysis layer provides the required data.

Data analysis layer on the data Optimization, data retrieval, for the application layer of abnormal work order intelligent distribution, different Intelligent processing of regular work order and multidimensional quality evaluation of acquisition and operation According to support.

The main function of the data ETL layer is to align the ETL tools. Source data is extracted, cleaned, transformed, and transferred to the data storage layer Load.

The arrival of a new era and the rapid development of information technology and Internet technology have promoted the rapid development of various fields combined with network technology in China. Based on the development and promotion of big data and cloud computing, China's radio and television operators have completed further optimization and improvement of their front-end systems in combination with new technologies, which has greatly improved their equipment and operation methods compared to previous ones [1]. From the actual development of the current radio and television field, it can be understood that the basic performance of the front-end system of radio and television operators has been greatly improved, and the practicality and reliability of system operation have been more optimized. However, with the continuous improvement of the development speed and level of front-end systems, there are also corresponding problems and shortcomings. Operation and maintenance personnel need to pay attention to more levels of services, resulting in increasing work pressure and workload [2,3]. In this reality, how to ensure that the operation and maintenance personnel can lock in the problem area the first time is an urgent issue to be solved. Based on the investigation of the current operation process of the front-end operation and maintenance system for radio and television, it can be seen that the analysis and collation of data information still continue to use traditional manual analysis and processing forms. However, with the progress and development of society, the amount of various types of information continues to increase, invisibly increasing the work pressure of staff. At the same time, the efficiency of data analysis work is gradually reduced due to insufficient energy, which is very easy to bring errors to the data analysis results due to personal operational errors of staff, which will lead to the failure to maximize the effectiveness of the data information collected through a large amount of manpower and material resources. At the same time, it should also be clear that the combination of front-end system operation and big data information technology will make business execution more complex and volatile. Simply using manual data analysis and processing is difficult to complete work tasks, and staff cannot flexibly respond to information overload issues. The specific causes of failures cannot be identified in a timely manner, which to some extent increases the investment costs of front-end system operation and maintenance. Therefore, in the current work environment, it is necessary to combine more advanced modern technology to make up for the shortcomings in work forms. Modern technology can more effectively integrate the alarm information of the front-end system to quickly determine the cause of equipment operation failures and propose targeted solutions to solve them. More importantly, the use of modern technology can help staff predict possible failures in advance, prevent them in the process of operation and maintenance, and reduce the adverse impact on system operation [4]. This article presents a large number of operation and maintenance data and system history data. Multi-dimensional analysis, screening the severity of various anomalies, output collection Operation, and maintenance of utility value model to determine the order of exception handling, gradually improve the abnormal work order processing completion rate and work effectiveness control, further improving the availability of equipment and data acquisition system Integrality.

## 2 Methods

### 2.1 Intelligent operation and maintenance system architecture of electromechanical equipment

To be specific, the big data analysis technology system mainly relies on key technical measures such as relational online analytical processing and multidimensional online analytical processing to provide various data services for the power system. In the work of the power grid system, the big data analysis technology system is mainly divided into three levels, and the first is the data storage layer. The data storage layer needs to build the database system and ensure the comprehensiveness of all types of data information. The second is the data analysis layer. After power information data is stored, it is necessary to build a data analysis model to further analyze data anomalies. The third is the data application layer. The analyzed data information will be fed back to the end of the operation. From the perspective of data flow, the electromechanical equipment operation and maintenance system can be divided into three parts: data collection, data transmission, and intelligent operation and maintenance platform. The intelligent operation and maintenance platform that integrates and accesses electromechanical equipment data and extracts key information from massive data is its core part. The architecture diagram of the electromechanical equipment operation and maintenance system is shown in Figure 1 [5].

#### 2.1.1 Data collection

Key status information of equipment through sensors, locators, and other devices of electromechanical equipment and other intelligent devices, such as voltage, current, longitude, latitude, temperature, and humidity is collected. These data are then packaged and encrypted for remote transmission using communication protocols and existing networks.

#### 2.1.2 Data transmission

Based on network devices such as intelligent gateways, wired and wireless networks are accessed, remotely

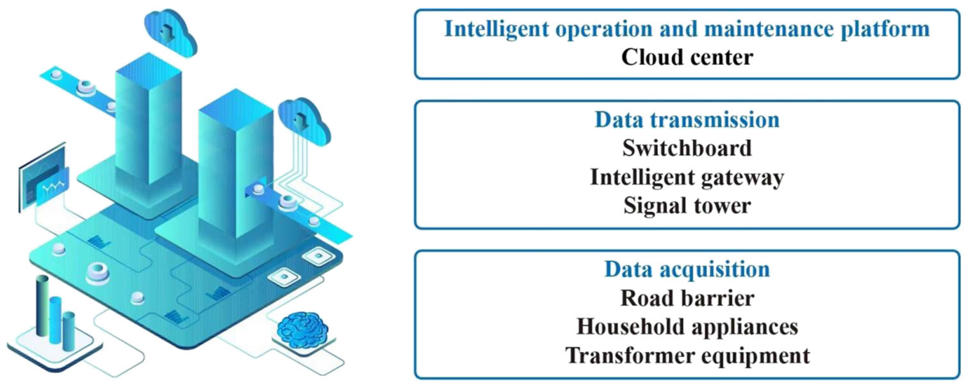


**Figure 1:** Architecture of electromechanical equipment operation and maintenance system.

transmitted to the cloud center, decrypted for preprocessing, and stored in a message queue for data analysis [6,7].

### 2.1.3 Intelligent operation and maintenance platform

The integrity, stability, and continuity of mechanical and electrical equipment status data are ensured through accessing services, preprocessing and analyzing the original mechanical and electrical equipment status data based on data cleaning algorithms. The core of mechanical and electrical equipment operation and maintenance is data analysis. The primary task of data analysis is to identify problems, namely, fault prediction. Using trained fault prediction models, a large number of mechanical and electrical equipment status data are processed efficiently and accurately, and fault prediction results are output. Finally, the results of data analysis are pushed to subsequent application services, such as monitoring interfaces/client alarm prompts, database persistent storage, and automatic recovery operations.

## 2.2 Design of intelligent operation and maintenance platform for electromechanical equipment

In the actual process of technology application, mainly from the following aspects, to realize the multi-dimensional data evaluation of power system: one is for the power grid system. Quality evaluation of all kinds of equipment products is applied, through the use of equipment. Big data analysis technology can be combined with all kinds of power grid equipment operation data information, different Often terminal data information constructs a power grid product equipment quality analysis evaluation refers to Mark. Through data operation, one can understand all kinds of equipment. At the same time, it can also evaluate the running quality of the equipment and understand the equipment property, ensuring the effective implementation of the work. The second is the operation and maintenance of the power grid system and it is overall data evaluation of the quality of the work. In an electric power system, the grid carries dimensional data that will be updated in real time, and it will affect the working quality of the entire power grid.

The intelligent operation and maintenance platform is the core carrier of intelligent operation and maintenance, realizing intelligent operation and maintenance by integrating functions such as access to operation and maintenance big data, hierarchical storage, and intelligent processing. Based on Kafka message queue, Spark big data processing framework, and MongoDB distributed database, this article constructs distributed data access service, data preprocessing, and fault prediction function modules using a cluster composed of thousands of nodes, ensuring the functionality and stability of the service. The platform architecture is divided into three layers from bottom to top: the lowest layer is data access services based on Kafka message queues and data cleaning for electromechanical equipment problems, the middle layer is fault analysis and prediction based on Spark's big data processing framework, and the top layer is analysis result output feedback. The architecture of distributed electromechanical equipment intelligent operation, and maintenance platform is shown in Figure 2 [8].

The specific design description of each layer of the platform is as follows:

Data access service and data preprocessing: A cluster composed of multiple nodes of Kafka message queues provides distributed data access services. Multiple nodes are interconnected and collaborated with each other. The accessed electromechanical device data are cached in each node, and multiple nodes can backup based on the leader to avoid data

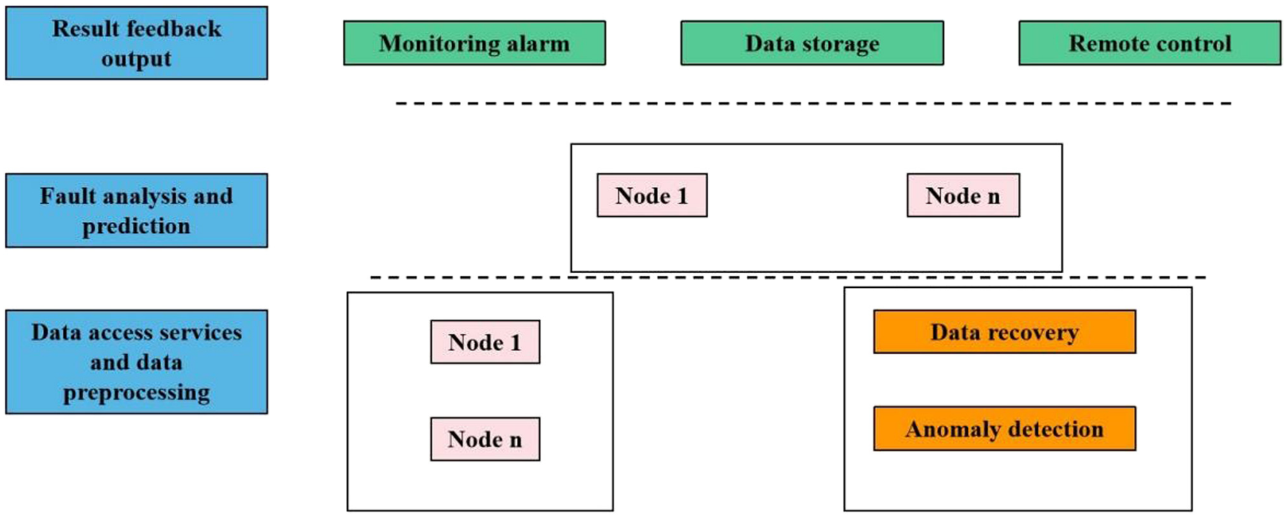


**Figure 2:** Architecture of distributed electromechanical equipment intelligent operation and maintenance platform.

loss and allow parallel writing and reading. Data cleaning algorithms are used to detect and repair cached electromechanical device data. Fault analysis and prediction is the layer composed of multiple Spark nodes. Spark clusters coordinate computing resources based on the YARN resource manager, plan task execution steps for submitted programs, and assign tasks to worker nodes to achieve parallel electromechanical device fault prediction [9,10].

Analysis result output feedback is the application layer for further processing after the electromechanical equipment intelligent operation, and maintenance platform obtains the fault analysis results. According to different needs, electromechanical equipment data are stored to a local database to complete persistence, issue monitoring alarms based on fault prediction results, trigger subsequent processing by operation and maintenance personnel, and even restore services to remote control operations for electromechanical equipment based on automation scripts.

## 2.3 Data access service and data preprocessing

### 2.3.1 Data access service cluster architecture

This article implements a distributed data access service based on Kafka to ensure the flexibility, concurrency, and stability of data access. The following is an introduction to the functions of Kafka cluster components:

Producer: is the API interface of Kafka, used to write data to the specified topic and broker.

Consumer: is the API interface of Kafka, used to read data from the specified topic and broker.

Connector: is the API interface for Kafka, used to manage data connections between Kafka and various databases.

Streams: is the API interface for Kafka, used for data stream processing, consuming data from one or more specified topics and producing data streams to one or more specified topics.

Broker: is a fragmented entity of Kafka, responsible for the implementation of functions of the Kafka cluster on the nodes, including file storage and data grouping using different topic names. Brokers between Kafka nodes can backup based on leaders, thereby forming a distributed message queue cache and ensuring the stability of data access services [11].

Topic: is a grouping identifier for Kafka file storage. It can support multiple groups of consumers and a group of multiple consumers to read data based on the reading progress Offset and the consumption group name Groupid.

### 2.3.2 Data preprocessing process design

This article designs a data cleaning algorithm for electromechanical equipment based on independent forests to achieve data preprocessing for the intelligent operation and maintenance platform of electromechanical equipment, improving data quality.

This article classifies the data problems that may occur in the generation and network transmission of electromechanical equipment status data into three categories: data loss, data redundancy, and data error. Data redundancy refers to the occurrence of data with completely consistent parameters, including equipment number and monitoring time, and the direct deletion of redundant data. Data missing refers to data whose one or more values are NULL or invalid. For missing or invalid data, the initial value is filled in and is marked as abnormal data pending data repair. Data error refers to data where one or more values of the data do not match the actual situation. These data have no obvious representation, are difficult to identify, and need to be repaired. The data preprocessing method in this article focuses on the discovery and repair of abnormal data and proposes a data cleaning algorithm for electromechanical equipment based on an independent forest algorithm. Based on the independent forest algorithm, it quickly detects data anomalies in the operation and maintenance data of electromechanical equipment and adopts the NDAM algorithm to correct the abnormal data [12].

## 2.4 Small amount of tag fault prediction

Fault prediction in intelligent operation and maintenance usually uses supervised learning algorithms to learn fault characterization knowledge from tag sample data and establish models for fault prediction. The accuracy of supervised learning algorithms requires appropriate models and sufficient label sample data. Tags for early sample data are typically generated based on the expert experience of the tester. Currently, the application field of intelligent electromechanical devices has expanded from industrial production to communication facilities, cloud services, and even household appliances. The cost of early testing corresponding to the introduction of a large number of types of electromechanical devices cannot be ignored. In order to meet the requirements of intelligent electromechanical equipment fault prediction with high accuracy, fast response, and a small number of tags, this article proposes a semi-supervised learning method independent forests and Gradient Boosting Decision Tree (IF-GBDT) based on the unsupervised learning algorithm-independent forest improved supervised learning algorithm gradient lifting tree: Using an independent forest algorithm to provide labels for unlabeled samples based on a small number of label samples, sufficient label samples are provided for model training of the gradient lifting tree algorithm, thereby reducing the label sample requirements of the gradient lifting tree algorithm. A small number of label samples are from early equipment test run results, with key status information and fault labels for the equipment. The labels contain the expert experience of testers and are used to train independent forest models [13]. A large number of unlabeled samples are collected results from the commissioning and formal operation stages of the equipment, with key status information of the equipment. Because there is no tag, it cannot be directly used for model training of supervised learning algorithms. Data to be processed contained the real-time collection results of the formal operation stage of the equipment, including the key status information of the equipment, which need to be accurately and quickly processed through sensor collection and network transmission. Independent forest model, generated by training a small number of label samples, calculates the anomaly evaluation value of the data based on the test results and obtains the anomaly evaluation threshold based on the label data. Based on the abnormality evaluation threshold, a large number of unlabeled samples are evaluated for abnormalities. Data with an abnormality evaluation value higher than or equal to the threshold are marked as faulty, and data with an abnormality evaluation value lower than the threshold are marked as faultless [14]. A small number of label samples are from early equipment test run results, with key status information and fault labels for the equipment. The labels contain the expert experience of testers and are used to train independent forest models. A large number of unlabeled samples are collected results from the commissioning and formal operation stages of the equipment, with key status information of the equipment. Because there is no tag, it cannot be directly used for model training of supervised learning algorithms. Data to be processed: The real-time collection results of the formal operation stage of the equipment, including the key status information of the equipment, which need to be accurately and quickly processed through sensor collection and network transmission. Independent forest model, generated by training a small number of label samples, calculates the anomaly evaluation value of the data based on the test results and obtains the anomaly evaluation threshold based on the label data. Based on the abnormality evaluation threshold, a large number of unlabeled samples are evaluated for abnormalities. Data with an abnormality evaluation value higher than or equal to the threshold are marked as faulty, and data with an abnormality evaluation value lower than the threshold are marked as faultless [15].

## 2.5 Fault prediction model training

The gradient lifting tree algorithm in this article is implemented based on Spark ML lib. By calling the GBT Classifier function, a large number of labeled sample data D3 is obtained, training the mechanical and electrical equipment fault prediction model and outputting the model to HDFS for invocation. Table 1 shows the parameter meanings of the GBT Classifier.

## 2.6 Deployment of online fault prediction service

To efficiently and timely process massive intelligent electromechanical equipment data, a fault prediction model is deployed to the Spark cluster. Spark clusters have three cluster resource management methods: Standalone mode, Mesos mode, and YARN mode. The Standalone mode is a Spark native cluster manager that is suitable for small-scale cluster management and easy to deploy. The Mesos mode is based on the general cluster manager Apache Mesos, with strong performance and scalability. The YARN mode is based on the Apache Hadoop resource manager, YARN, and has high cluster utilization. It is suitable for Map Reduce batch processing and streaming computing. The online fault prediction service deployment in this article is based on the YARN mode. The Spark cluster adopts the "Master Worker" working mode in the YARN mode. Users submit computing projects to the Master, who uses a directed acyclic graph to plan the execution steps of the project. Each step can be divided into parallel execution tasks for the Work to execute. Finally, the elastic data sets that have undergone multiple concurrent processing are combined and output to the specified directory, HDFS, or support database [16]. According to the collection operation and maintenance of the terminal quality of each terminal manufacturer, Line quantitative analysis, using each terminal manufacturer's operating terminal number, made the final confirmation. The number of faulty terminals, proportion of abnormal terminals, number of replaced terminals, and terminal clock are different. Constant and other data establish the acquisition terminal product quality analysis evaluation index.

# 3 Results and analysis

## 3.1 Electromechanical equipment fault prediction system

### 3.1.1 Experimental environment

In this experiment, five virtual machines were installed on three servers to configure a distributed Kafka, Spark, and MongoDB environment, and Tomcat was configured on one server to build an electromechanical device data display platform for reading MongoDB data. Based on the above experimental conditions, a functional test of the electromechanical equipment fault prediction system was conducted. The specific experimental environment parameters are shown in Table 2.

This experiment verifies the functionality of the electromechanical equipment fault prediction system by simulating the process of electromechanical equipment data from accessing Kafka to displaying data information on the electromechanical equipment display platform. The output delay and concurrency performance of the system are tested based on time intervals, and the system performance is verified. The data flow of the experiment is shown in Figure 3 [17].

### 3.1.2 System functions

In order to verify the functionality of the electromechanical equipment fault prediction system and the output delay of the test system, this experiment was conducted in environments such as configuring and deploying

**Table 1:** GBT Classifier parameters and their meaning

| Label Col | Label column name, default to label |
|---|---|
| Features Col | Feature vector column name, default to features |
| Prediction Col | The default prediction column name is prediction |
| Max Depth | The maximum depth of the tree, which can only be a natural number, defaults to 5 |
| Max Bins | The maximum number of buckets for discrete continuous features, which defaults to 32 |
| Min Instances Per Node | The minimum number of samples contained in each sub-node after splitting. If it is less than this value, it will no longer split |
| Min Info Gain | The minimum information gain during tree node splitting, which is 0 by default |

**Table 2:** Fault prediction system experimental environment parameters

| Server configuration | Virtual machine configuration | Software deployment |
|---|---|---|
| CPU Intel Xeon Silver 4114 20 Core memory 128 G | CPU8 core memory 32 G | Spark Master *1 Kafka*1 MongoDB*1 |
| | CPU4 core memory 16 G | Spark Worker *1 Kafka*1 MongoDB*1 |
| | CPU4 core memory 16 G | Spark Worker *1 Kafka*1 MongoDB*1 |
| CPU Intel Xeon Silver 4114 20 core memory 128 G | CPU4 core memory 16 G | Spark Worker *1 |
| CPU Intel Xeon Silver 4114 20 core memory 128 G | CPU4 core memory 16 G | Spark Worker *1 |
| CPU Intel i7-8700 K CPU4 core memory 32 G | No virtual machine adoption | Tomcat |

Kafka data access service, Spark online fault prediction service, MongoDB data storage, and tomcat based electromechanical equipment data display platform.

Experimental Step: The Kafka Producer API is used to input a complete electromechanical device data without a fault tag to the specified topic {"'i" : 0.19, "x" : 0, "id" : "542181435", "y" : 0, "u": 0.4, "c" : 0, "b": 3.64, "nRssi" :30, "h" : 45.5, "time" : "20 19-08-30 09:38:34", "t" : 25.6}; Refresh the electromechanical equipment data display platform until the electromechanical equipment data containing the fault tag appears and the recording time is different.

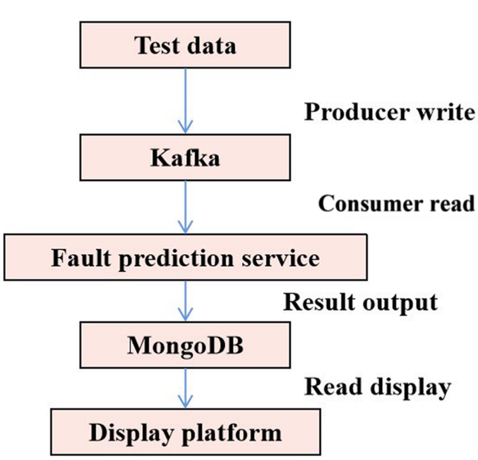


**Figure 3:** Flow chart of fault system test experiment data.

After the above experimental steps, it has been proved that the electromechanical equipment fault prediction system can complete the entire process from accessing electromechanical equipment data to storing and outputting fault results in a relatively short time, and the functionality of the electromechanical equipment fault prediction system has been verified [18].

### 3.1.3 Concurrent performance

In order to verify the functionality of the electromechanical equipment fault prediction system and the concurrent performance of the test system, this experiment was conducted on environments such as configuring and deploying Kafka data access service, Spark online fault prediction service, MongoDB data storage, and tomcat-based electromechanical equipment data display platform. In addition, in order to meet the needs of the experiment, the GJHEE data set was copied and expanded to 100,000 and 200,000 data sets, and the special id "END" was used as the tail identifier of the data set.

Experimental steps: The Kafka Producer API is used to input the expanded data set to the specified topic and the Robot 3T tool to query the specified collection in the MongoDB database until the number of documents added to the collection is the same as the number of input data, and the last piece of electromechanical equipment data with a special ID appears, resulting in a poor recording time; 100,000 and 200,000 pieces of electromechanical equipment data are entered and the processing time difference corresponding to enabling a different number of nodes as shown in Figure 4 [19].

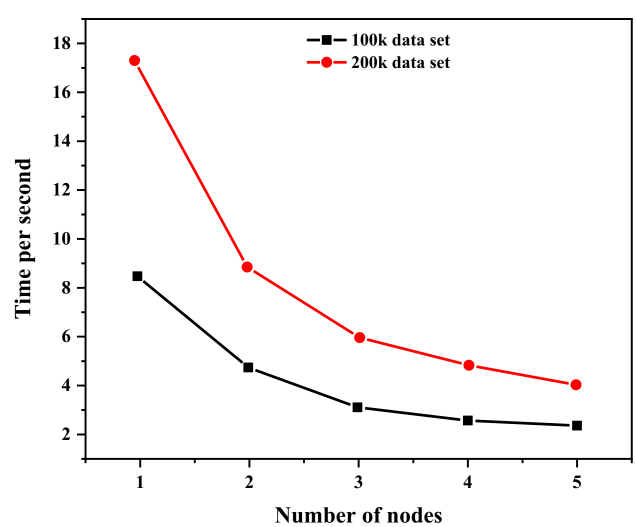


**Figure 4:** Processing time chart for 100k and 200k data sets.

As shown in Figure 4, when the number of nodes enabled by the online fault prediction service increases from 1 to 5, the processing latency of the electromechanical equipment fault prediction system is significantly shortened. Therefore, the parallelization of online fault prediction services based on Spark can greatly improve the speed of data processing. In addition, we can observe that when the number of nodes is 5, it takes 1.93 s to process 100,000 pieces of electromechanical device data and 3.67 s to process 200,000 pieces of electromechanical device data. It can be estimated that the implemented electromechanical equipment fault prediction system can provide data access and online fault prediction services for approximately 54,500 pieces of electromechanical equipment data per second [20]. Out of the use of big data multidimensional analysis technology, intelligent distribution from abnormal work order, abnormal work starting from the three aspects of single intelligent processing and multidimensional quality evaluation of collection operation and maintenance, the overall collection operation and maintenance work is optimized to improve the overall operation efficiency and quality, and the management level.

## 4 Conclusion

Through the introduction of big data technology, the collected data are deeply dug, excavating, analyzing, and combing to realize the intelligent collection, operation and maintenance of abnormal work order. Generation, analytical processing, and operation dimension multidimensional quality evaluation can enable operation and maintenance. Workers from extensive to intensive, lean transformation, and improved operation and maintenance work efficiently and give quality work. This article proposes an intelligent operation and maintenance platform architecture for electromechanical equipment based on the actual situation of electromechanical equipment operation and maintenance and the idea of intelligent operation and maintenance. It includes three layers: data access service and data cleaning, fault analysis module, and analysis result feedback. This article implements high concurrency, high flexibility, and high-reliability data access services based on Kafka message queues and implements persistence of access data caches based on MongoDB distributed databases. This article studies a data cleaning method for detecting errors in electromechanical equipment data, focusing on anomaly detection of electromechanical equipment access data, and implements an anomaly detection algorithm for data based on independent forest algorithms. We also propose a small number of tag fault prediction algorithm IF-GBDT. First, based on the improved independent forest algorithm, a small number of labeled samples are learned, and after obtaining the anomaly threshold, a large number of unlabeled samples are supplemented with labels. Then, based on the gradient lifting tree algorithm provided by Spark ML lib, a fault prediction model is trained using a large amount of labeled data. Finally, we deploy the fault prediction task to run on the Spark cluster. In this article, the accuracy and time cost of the proposed fault analysis algorithm are tested and determined on the real electromechanical equipment status collection data set and the open space shuttle operation status data set. Next, we should further study the applicability of the proposed method to the fault analysis of other electromechanical equipment.

**Funding information:** This study did not receive any funding in any form.

**Author contributions:** All the authors have made important personal contributions to this manuscript. Wenjuan Yang: writing and performing surgeries; Zhongbin Chan: data analysis and performing surgeries; Yi Wang and Fuli Qi: article review and intellectual concept of the article.

**Conflict of interest:** The authors declare that they have no competing interest.

**Informed consent:** Informed consent was obtained from all individuals included in this study.

**Ethical approval:** The conducted research is not related to either human or animals use.

**Data availability statement:** The data used to support the findings of this study are available from the corresponding author upon request.

## References

[1] M. Zhang, H. Lv, H. Chen, and Q. Hu, "Analysis of constraints on data management of intelligent platform for smart building in the era of big data," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 768, no. 1, p. 012109, 2021.
[2] L. Geng, "Application status and development suggestions of big data technology in petroleum engineering," *Pet. Drill. Tech.*, vol. 49, no. 2, pp. 72–78, 2021.
[3] L. Liu and S. B. Tsai, "Intelligent recognition and teaching of english fuzzy texts based on fuzzy computing and big data," *Wirel. Commun. Mob. Comput.*, vol. 2021, no. 1, pp. 1–10, 2021.

[4] J. Gao, L. I. Dong, S. U. Dan, W. U. Tianqi, Z. Junmei, W. Zhang, et al., “Research progress on the extraction technology of seabuckthorn fruit oil and the application of nutritional factors,” *Sci. Technol. Food Ind.*, vol. 43, no. 13, pp. 400–407, 2022.
[5] Z. Dong, Y. Sheng, Z. Huang, T. Zhang, F. Qian, Y. Zhu, et al., “Living cell nanoporation and exosomal rna analysis platform for real-time assessment of cellular therapies,” *J. Am. Chem. Soc.*, vol. 144, no. 21, pp. 9443–9450, 2022.
[6] S. D. Yu, “Research on cloud computing in the key technologies of railway intelligent operation and maintenance sharing platform,” *J. Phys: Conf. Ser.*, vol. 1800, no. 1, p. 012010, 2021.
[7] J. Chen, W. Shi, Y. Liu, Q. Chen, P. Huang, and J. Li, “Big data analysis supports the research and pilot application of distribution network planning technology,” *IOP Conf. Ser: Earth Environ. Sci.*, vol. 632, no. 4, p. 042047 (9pp), 2021.
[8] H. Yang, “Construction and application of smart education teaching platform,” *J. Phys: Conf. Ser.*, vol. 1744, no. 4, p. 042147 (6pp), 2021.
[9] S. Wang, H. Wang, M. Miao, J. Li, D. Gu, and L. Ding, “Management and control platform of auxiliary drugs based on big data,” *J. Physics: Conf. Ser.*, vol. 1744, no. 3, p. 032009 (6pp), 2021.
[10] M. Narayanan, R. G. Kumar, J. Jayasundaram, R. Kumar, and C. Isaac, “Big data analytics and an intelligent aviation information management system,” *Turkish J. Comput. Math. Educ. (TURCOMAT)*, vol. 12, no. 11, pp. 4328–4340, 2021.
[11] X. Lv and M. Li, “Application and research of the intelligent management system based on internet of things technology in the era of big data,” *Mob. Inf. Syst.*, vol. 2021, no. 16, pp. 1–6, 2021.
[12] N. Zhang, “Research on the application of data encryption technology based on network security maintenance in computer network security,” *J. Phys. Conf. Ser.*, vol. 1744, no. 2, p. 022060, 2021.
[13] S. Yan, “Application research of big data technology in engineering cost,” *J. Phys: Conf. Ser.*, vol. 1852, no. 4, p. 042003 (7pp), 2021.
[14] X. Ni and X. Hou, “Application of big data technology in water conservancy project informatization construction,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 768, no. 1, p. 012113, 2021.
[15] M. Wang and M. Qian, “Research and application of the network learning platform in based on big data analysis and mobile communication,” *J. Phys. Conf. Ser.*, vol. 1952, no. 4, p. 042054, 2021.
[16] C. Zhang, B. Du, K. Li, Q. Li, Z. Zhang, and X. Song, “Selection of the effective characteristic spectra based on the chemical structure and its application in rapid analysis of ethanol content in gasoline,” *ACS Omega*, vol. 7, no. 23, pp. 20291–20297, 2022.
[17] H. M. Hung, L. M. Duc, N. D. Dat, H. Thai, C. L. Van, D. N. Trong, et al., “Synthesis and characterization of polypyrrole film doped with both molybdate and salicylate and its application in the corrosion protection for low carbon steel,” *ACS Omega*, vol. 7, no. 23, pp. 19842–19852, 2022.
[18] C. Mase, J. F. Maillard, B. Paupy, M. Hubert-Roux, C. Afonso, and P. Giusti, “Speciation and semiquantification of nitrogen-containing species in complex mixtures: application to plastic pyrolysis oil,” *ACS Omega*, vol. 7, no. 23, pp. 19428–19436, 2022.
[19] Y. Song, Z. Zhao, Y. Zheng, and B. Zeng, “Investigation and application of high-efficiency network fracturing technology for deep shale gas in the southern sichuan basin,” *ACS Omega*, vol. 7, no. 16, pp. 14276–14282, 2022.
[20] X. Xu, X. Li, X. Gao, D. Guo, W. Yang, C. Zhang, et al. “Application of large-scale molecular prediction for creating the preferred precursor ions list to enhance the identification of ginsenosides from the flower buds of panax ginseng,” *J. Agric. Food Chem.*, vol. 70, no. 19, pp. 5932–5944, 2022.