Research on Power Cable Operation Fault Alarm System Based on Big Data

Yun Guo*, Minhu Xu, Jian Zhang, Long Tan, Kexin Zhang and Rong Cao
State Grid Heilongjiang Electric Power Co., Ltd. Electric Power Research Institute, Harbin 150040, China

*Corresponding author email: yunguo2020@sgcc.com.cn

Abstract. With the continuous development of the national economy, in order to meet the needs of industrial and agricultural production and people’s daily use of electricity, the power system has gradually developed in the direction of large capacity and high voltage. The thesis established a high-voltage cable line fault detection and lean management system based on big data technology, and gave a basic introduction to the high-voltage cable line fault diagnosis and lean management system modules. Analysed the system function modules and judgment logic based on big data technology, and the importance of the system to the cable operation and maintenance management unit to improve the management level.

Key words. Big data, power cable, operation fault alarm, fault alarm system.

1. Introduction
The safe operation of transmission lines is not only directly related to the reliable operation of the entire power system, but also related to the safety and quality of power users. The fault diagnosis method of transmission line has always attracted the attention of various academic circles and industry circles, and has become a hot spot of their joint research. With the continuous development of big data technology, in addition to being widely used in the Internet, finance, logistics and other fields, big data technology in the energy and power industry has also entered a stage of gradual development. Fault diagnosis is a common method of inspection and testing, which can discover whether there are certain hidden safety problems in the operation of the system and its construction equipment. Fault diagnosis belongs to the category of network survivability prevention and is a powerful means to ensure the system is in a stable operation state. The main tasks of power system fault diagnosis include fault detection, basic type judgment, fault location, error recovery and other stages. Among them, fault detection means that after the central computer establishes a connection with the power system, it can transmit periodic detection signals to the lower computer, and judge whether there is a certain operating fault in the power system by judging the response frame of the received data. Basic type judgment means that after the power system detects the fault problem, it determines the attribute type of the type of fault through steps such as cause analysis.

The operating environment of power cables is generally harsh. Factors such as temperature, humidity, harmful gases, and microorganisms in cable trenches, tunnels, and pipes all adversely affect cable operation, leading to deterioration of cable insulation and cable failure [1]. Once the cable fails,
it is very difficult to find and eliminate the fault point due to the particularity of the laying environment, which will seriously affect the reliability of the power grid. This article will integrate the above-mentioned multi-source information related to cable maintenance, model the maintenance tasks as a whole based on big data mining and analysis algorithms, and flexibly arrange and push maintenance tasks, forming a big data-based intelligent power cable maintenance technology framework. The framework will be able to solve the problem of "massive data and lack of information", improve the pertinence and flexibility of maintenance, and improve the efficiency of maintenance.

2. Faulty big data processing algorithm

Big data analysis algorithms usually have certain requirements for the standardization of the input data format. In order to facilitate the unified big data analysis process, it is first necessary to pre-process the three types of data to have a unified and standardized format. The original data format is different and the range is different (for example, the voltage is a continuous value of 110kV, and the severity of the accident recorded in the cable maintenance record is a discrete value of high, medium, low, etc.). Before the analysis, the data is first continuous, Normalized pre-processing and self-annotation. Continuous and normalized input data are the basic requirements of general big data analysis algorithms, and self-labelling is to facilitate the big data analysis algorithm to automatically learn features and rules related to maintenance from the labelled information [2]. In the framework proposed in this article, continuity is to use integer substitution to convert discrete values into continuous values, and normalization is to normalize different numerical ranges to a fixed range such as [0,1] interval. The general formula for normalization is:

$$x_{new} = \frac{x_{old} - x_{min}}{x_{max} - x_{min}}$$  \hfill (1)

In the formula, $x_{new}$ is the normalized value of the current attribute, $x_{old}$ is the original value of the current attribute, and $x_{max}, x_{min}$ is the maximum and minimum value of the current attribute respectively. For example, the cable maintenance personnel may record the cable running status as "good, normal, low-risk, high-risk" 4 levels, in the continuous process, they can use the 4 integers of 4, 3, 2, and 1 to correspond to it. For a certain value such as 3, the corresponding value after using the above normalization formula is:

$$x_{new} = \frac{3-1}{4-1} \approx 0.667$$  \hfill (2)

The mining problem can be divided into two sub-problems: (1) Find all frequent sets that meet the support conditions; (2) Use frequent sets to generate association rules. Among them, step (1) needs to scan the transaction database multiple times, and the consumption of time and space is the key to restricting the efficiency of mining. The frequent pattern FP algorithm first compresses the database into a frequent pattern tree, which is equivalent to grouping the database and can reduce the number of database scans, while the associated information is still stored in the nodes of the tree. No matter how to improve, there is a common basic definition, namely

$$I = \{i_1, i_2, \ldots, i_n\}$$ \hfill (3)

$I$ is a collection of m different data items, among which: elements are called items, and the collection of items is called itemset. Given a transaction database
$$D = \{T_1, T_2, ..., T_n\} \quad (4)$$

Where: each transaction $T$ is a subset of itemset $I$, $|D|$ is the total number of transactions in $D$, $X$ and $Y$ are items or itemset in $T$, $X \cap Y = \Phi$. If the transaction $T$ contains both $X$ and $Y$, then the association rules can be obtained:

$$X \Rightarrow Y (S\%, C\%, I\%) \quad (5)$$

Where: $S\%$ is the proportion of transaction $T$ that meets the conditions in transaction database $D$. Figure 1 reflects the detailed structure of the big data operation and maintenance framework.

![Big data operation and maintenance framework structure diagram](image)

**Figure 1.** Big data operation and maintenance framework structure diagram

### 3. Common faults of power cables

#### 3.1. The influence of natural factors

In our country, the laying of transmission lines is generally outdoors. They are exposed to the outdoors for a long-time during operation and are easily affected by extreme weather. These influencing factors are not only difficult to predict, but also bring extremes to the transmission line. Strong destruction.

**3.1.1. Bird damage.** The flying behaviour of the birds themselves will not have a great impact on the transmission lines, but the birds often carry debris when they fly over the transmission line. The scattered or attached debris may cause the line [3]. Transmission line failure. In addition, bird nesting and bird droppings flashovers can cause transmission line failures.

**3.1.2. Lightning strikes.** According to the statistical analysis of transmission line failures of the State Grid System, from 2011 to 2019, more than 50% of high-voltage line trips were caused by lightning strikes, which shows that lightning strikes are still the main cause of transmission line flashover. Lightning strikes have a great impact on transmission lines and power equipment.

#### 3.2. The influence of human factors

At present, transmission line tripping accidents caused by human factors are on the rise year by year. The repeated prohibitions of illegal activities such as private construction of houses, extensive road repairs, blasting of mountains, digging of cofferdams, and incineration in the transmission line
protection zone have caused serious threats to the safe and stable operation of transmission lines and grid systems [4]. Therefore, damage to transmission lines caused by human factors cannot be ignored.

4. Design of power cable fault alarm system

4.1. Overall design
The processing framework designed in this paper is divided into 4 parts: data collection, message queue, real-time processing and data visualization. Data collection is carried out through various sensors. The sensor transmits the monitored data information to the message queue through the network. The main function is to classify and aggregate the received information and data and distribute them to the Storm platform [5]. Then, the Storm data processing platform will analyse the data in real time, and finally submit the analysis results to the power monitoring centre in a visual form.

4.2. Introduction to functional modules
The system is interconnected with the production management system, operation management system, online monitoring and other system data, and can automatically obtain the basic ledger, operation and maintenance records, status monitoring and other data of the line during fault diagnosis. Use mobile AP to collect data at the business site for on-site surveys, fault anatomy, etc., to avoid secondary entry. After all the data and information related to fault diagnosis are collected, the system automatically checks the design of the cable line to check whether there are electrical design defects, system overvoltage’s and other design selection reasons that cause cable faults. According to the abnormalities found in the operation records, site survey, fault anatomy, physical and chemical tests, design verification, etc., the system automatically establishes the characteristic fingerprint of the cable fault. On the one hand, the system uses the fault diagnosis expert knowledge base to reason and judge the possible cause of the fault [6]. On the other hand, the fault intelligent diagnosis algorithm based on machine learning is compared with the massive cases in the fault case library, and the fault cause is intelligently diagnosed and similar Failure cases are for experts' reference. Experts combine the results of knowledge base logical reasoning and intelligent analysis of diagnosis algorithms, and refer to the breakdown mechanism model given by the system, finally determine the cause of the failure and automatically generate a failure analysis report. The design principle of the specific warning automatic processing module is shown in Figure 2.

![Figure 2. Schematic diagram of the design of the automatic warning processing module](image)

4.3. Alarm system database design
The alarm database of the new system includes multiple structures such as historical alarm event table, alarm event table, alarm level table, alarm rule table, policy information table, migration policy table, filter rule table, and alarm contact table, and these tables are kept between Closer data exchange
relationship. Among them, the power operation and maintenance fault data as the transmission information circulating in the system can be step by step through the form structure in the alarm database according to the connection sequence. While mastering the detailed alarm rules, indicate the alarm timetable and the associated conditions between the alarm strategy table. When the total amount of power operation and maintenance fault diagnosis data is relatively large, some alarm rules cannot be clearly reflected in the alarm rule form, which may easily cause problems such as lowering the transmission efficiency of alarm instructions. In order to avoid the occurrence of the above situation, the alarm database of the new system establishes an alarm contact table related to power operation and maintenance failure data, and clearly contains reference information such as the alarm type to which these data belong [7]. When there is an obvious accumulation of power operation and maintenance failure data, the alarm contact table can refer to the migration relationship between the data and the data, and allocate and process it, thereby achieving a distributed storage alarm database operation mode and alleviating the transmission crisis of the system. The specific alarm database design principle is shown in Figure 3.

5. Experimental test analysis

For power cables, long-term monitoring of cable status data can be regarded as streaming data, such as partial discharge signals, voltage, leakage current, etc. This article uses partial discharge signals as monitoring data to test the performance of the online real-time diagnosis method based on Storm. The Storm cluster is built in a cluster of 11 virtual machines, with 1 virtual machine as the master node, and the remaining 10 as working nodes. Each working node has 4 process ports, so this platform can provide up to 40 processes. At the same time, set up Kafka and Zookeeper on the cluster.

5.1. Throughput test

Throughput is one of the important indicators for testing big data processing systems. In this test, set the number of working nodes of Storm unchanged, change the concurrency of Kafka components and variable prediction model components, make different combinations, and calculate the variable prediction model the throughput of each thread of the classification component in the same time. The result is shown in Figure 4.
5.2. Processing delay test

This article mainly uses the time stamp method to test the processing delay of each storage tuple in the storage structure. While the simulated monitoring source sends the data tuple to Kafka, it also sends the current operation time. After the processing is completed, the variable prediction model classification component the processing time delay can be obtained by subtracting the time stamp in the tuple from the completion time stored in. In this test, the work process is set to 5, the number of components is fixed, and the sending rate is controlled mainly by changing the time interval of the tuple data in the simulated monitoring source. Set different time intervals of 1 millisecond, 5 milliseconds, 10 milliseconds, and 50 milliseconds, 100 milliseconds. Calculate the delay of the first 500 data tuples.

6. Conclusion

With the continuous improvement of the degree of power informatization, the amount of data related to power cables has grown rapidly. Supporting cable maintenance based on big data analysis has become the inevitable development direction of maintenance work in the future. Based on big data related to cable maintenance, this paper designs a basic framework for cable maintenance based on big data analysis. The framework consists of three modules: source data acquisition and pre-processing, big data analysis and mining, and intelligent maintenance support. Excavating the analysis and processing flow will help the design and implementation of intelligent maintenance, and improve the intelligent level of cable maintenance.

References

[1] Xu, L., Zhu, P., Gu, H., & Xu, W. Early warning of electric equipment current-carrying faults based on variable-scale principal component analysis. Electric Power Automation Equipment, 32(5) (2012) 147-151.
[2] Allen, R. M., Gasparini, P., Kamigaichi, O., & Bose, M. The status of earthquake early warning around the world: An introductory overview. Seismological Research Letters, 80(5) (2009) 682-693.
[3] Zhang, Y., Xu, Y., Dong, Z. Y., Xu, Z., & Wong, K. P. Intelligent early warning of power system dynamic insecurity risk: Toward optimal accuracy-earliness tradeoff. IEEE Transactions on Industrial Informatics, 13(5) (2017) 2544-2554.
[4] Shaodi, H., & Li, X. Early warning of electric equipment current-carrying fault based on equivalent resistance analysis. Chinese Journal of Scientific Instrument, 34(3) (2016) 541-546.
[5] He, M., Wang, Y., & Tao, Z. A new early-warning prediction system for monitoring shear force of fault plane in the active fault. Journal of Rock Mechanics and Geotechnical Engineering, 2(3) (2010) 223-231.

[6] Nippes, P. I. Early warning of developing problems in rotating machinery as provided by monitoring shaft voltages and grounding currents. IEEE Transactions on energy conversion, 19(2) (2004) 340-345.

[7] Wenzel, F., Oncescu, M. C., Baur, M., Fiedrich, F., & Ionescu, C. An early warning system for Bucharest. Seismological Research Letters, 70(2) (1999) 161-169.