Research on Early Warning Technology of Lightning Arrester Defects Based on Multi-stage Information and Bayesian Inference

Ying Pei¹, Lin Niu¹, Haifeng Li¹, Yajin Li²*, Dayang Yu²

¹State Grid of China Technology College, Jinan 250002, China
²School of electrical engineering, Shandong University, Jinan 250061, China

*Corresponding author e-mail: liuyingnan@sdu.edu.cn

Abstract. The differences in the probability of occurrence of different equipment and defects lead to the small sample characteristics of the defect of the arrester, which makes it difficult to train an accurate prediction model. It is difficult to identify the abnormal state when the arrester monitoring data does not exceed the limit and increase steadily relying on the arrester monitoring index and threshold to judge the defect. Therefore, a lightning arrester defect early warning method based on multi-stage information and Bayesian inference is proposed. The Bayesian inference algorithm is used to calculate the probability of defect cause categories under different feature quantities. According to the new test evidence, the probability of the defect cause category under different feature quantities is updated to identify the defect cause. The algorithm automatically adjusts the prior probability indicators of equipment defects and causes in the model based on the new detection data and annotation conclusions to ensure the accuracy of defect cause classification. The lightning arrester operation and maintenance data and online monitoring system of a power company is used to analyze and verify the effectiveness and correctness of the method proposed in this paper, which provides effective support for the lightning arrester operation and maintenance.

Keywords: Bayesian classification; lightning arrester; Classification of defects; Multi-stage information.

1. Introduction

In order to ensure the safe operation of arresters, the online monitoring technology of arresters in substations has been widely used, and multi-source heterogeneous condition monitoring data has been obtained. [1] As an important equipment in the station, lightning arrester has high reliability, low probability of occurrence of defects, and few sample data of defects. The replacement time period of equipment is short, and many lightning arrester equipment have not failed during the life cycle, which result in fewer defective data samples and affect the application effect of artificial intelligence technology based on a large number of labeled sample data in the evaluation of the status of the arrester.

In view of the monitoring and evaluation of the deterioration and aging state of the arrester, the existing research has proposed multiple arrester evaluation indicators such as leakage current, resistive
current, resistive current third harmonic, fifth harmonic, etc. [2], and formulated the corresponding maintenance test procedures [3] and trend analysis method [4] to achieve the status evaluation of the arrester combined with the test and actual operation experience. But the external temperature and humidity [5] and power frequency electromagnetic field interference [6] will increase the measurement error of resistive current, which will affect the analysis results. In response to the above problems, some studies have proposed online parameter correction methods [7] and resistive current harmonic compensation methods [8]. However, the established temperature, humidity, and full current three-dimensional surface model has certain complexity, and the harmonic compensation method can only measure the resistive current under the fundamental voltage. Considering the operating environment of the arrester and other random uncertain factors, intelligent diagnosis method for arrester status based on particle swarms [9], genetic algorithm [10] and improved evidence theory [11] are proposed. The artificial intelligence methods also be applied to equipment fault diagnosis, but its application requires a certain amount of data to be collected to achieve accurate evaluation.

With a small sample of data, Bayesian inference can effectively combine the prior information from a small sample of multiple sources to obtain the posterior probability, which provides an idea for defect classification. Therefore, based on the analysis of arrester defect occurrence mechanism and various state detection methods, and using Bayesian principle, a method of arrester defect classification based on multi-stage information is proposed. In the first stage, reasoning is made under incomplete information. Perform on-site inspection under the guidance of Bayesian classification model, second-stage information is obtained, diagnostic data is supplemented, and posterior probability is calculated using new inspection data on the basis of prior probability to identify the cause of the defect and realize online diagnosis and early warning of the arrester. The prior knowledge in the model is obtained using a two-layer knowledge structure [13], which reduces the dependence of existing methods on data samples, and can be effectively applied to the classification of small sample data arrester defects.

2. Classification of Defects of Lightning Arrester and Measurement Of Features

Metal oxide arresters are mostly used in substations, and the low current area of zinc oxide resistors mainly presents dielectric characteristics. The arrester withstands lightning overvoltage and operating overvoltage for a long time during the operation of the arrester. Different types of faults or defects such as valve aging, internal moisture, surface cracking, pollution, heating, and discharge are prone to occur due to the long service life and the humid and hot environment. According to the different detection methods of lightning arrester defects, a multi-source multi-dimensional data model is established. The data sources of the arrester information model include monitoring, testing, operation and maintenance, etc. The structure of the information model is shown in Table 1

| Source                      | Type                                      |
|-----------------------------|-------------------------------------------|
| Monitoring data             | Leakage current; Number of operations    |
| Detection data              | Infrared and partial discharge detection data |
| Test data                   | Voltage of DC 1mA condition (kV); Current of 75%U1mA; Insulation resistance |
| Operation and maintenance data | surface cracking, surface contamination and other types; The characteristic data of the defect; The severity of the defect |
| Ledger data                 | The manufacturer and model of the equipment, the number of years of operation |
| Weather data                | Temperature, Humidity, Season             |

3. Early Bwarning Model of Lightning Arrester Defects Based on Bayesian Inference

Bayesian classification provides a method of calculating the posterior probability and making new judgments based on the new evidence obtained in each test on the basis of the prior probability. It is a technology mean of learning expert experience through artificial intelligence and judging the defect of arrester.
3.1. Basic principles of Naïve Bayes

The algorithm of Naive Bayes classification is defined as follows. Let the sample set be \(D = \{d_1, d_2, ..., d_m\}\), the feature attribute set corresponding to the sample data is \(X = \{x_1, x_2, ..., x_d\}\), wherein, \(m\) represents the number of sample sets in the time series, and each \(x\) is a feature in the sum of \(X\) sets. Classification set is \(Y = \{y_1, y_2, ..., y_n\}\). Calculate the conditional probability of each feature \(x_j\) of each category \(y_i\), denoted as \(P(x_j|y_i)\), according to Bayes' theorem,

\[
P(y|X) = \frac{P(X|y)P(y)}{P(X)}
\]

(1)

Because the conditions of each feature attribute are independent of each other,

\[
P(X) = \prod_{j=1}^{d} P(x_j|y_i)
\]

(2)

\(P(X)\) is the same value for all probability calculations, so formula (1) can be simplified to the following formula.

\[
P(y|X) = \prod_{j=1}^{d} P(x_j|y_i)
\]

(3)

If

\[
P(y_i|X) = \max \{P(y_1|X), P(y_2|X), ..., P(y_n|X)\}
\]

(4)

\(X \in Y_1\).

According to the data sample set for classification and feature attribute modeling, calculate the prior probability of different categories and the conditional probability of each feature attribute. Enter a new sample set, we can get the mapping relationship between the items to be classified and the category.

3.2. Algorithm for early warning of lightning arrester defects based on multi-stage information

This paper builds a model for the diagnosis of lightning arrester defects based on Naive Bayes is built, and the most probable arrester status type \(y_k\) from the operating status classification set is obtained based on the characteristics of the new sample.

\[
y_k = \max \{P(y_1|X), P(y_2|X), ..., P(y_n|X)\}
\]

(5)

The specific implementation process is as follows,

- Divide different operating status categories and key characteristics according to the data to be classified to form a data sample set \(\{(x_1, x_2, ..., x_m, y_n)\}\), wherein, the operating status of the equipment is divided into normal and abnormal. Construct the classification set of the operating state of the equipment as \(Y = \{y_1, y_2, ..., y_n\}\). The set of feature quantity of the diagnostic model \(X = \{x_1, x_2, ..., x_m\}\) is established according to the multi-source feature quantity of the arrester in Table 1. The calculation formula of the change of the leakage current trend is \(x = \sum \frac{x_{i+1} - x_i}{x_i} \times 100\%\). The difference of the three-phase leakage current is quantified by the correlation coefficient of the leakage current between the three phases, which realization method is to express the two state monitoring quantities of resistive fundamental wave current and resistive third harmonic current in vector form, and use the distance correlation method [14] to calculate the correlation coefficient between the full current of the three-phase arrester

- Calculate the proportion of different categories in the sample set According to the sample set, which is the prior probability \(P(y_i)\)

- When calculating the conditional probability \(P(X|y_i)\) of different feature quantities, for discrete feature quantities, calculate by counting the frequency of the feature quantity in the sample data set.
For continuous feature quantity, assuming it obeys a normal distribution, characterize the conditional probability use its respective probability density function. Its probability density function is as follows,

$$P(X = x_j \mid y_i) = \frac{\text{sum}(X = x_j \mid Y = y_i)}{\text{Total number of samples}}$$

(6)

Wherein, $\mu_{y_i}$ and $\sigma^2_{y_i}$ are the mean and variance of feature quantities under different classifications.

- For $X = \{x_1, x_2, \ldots, x_d\}$, use formula (3) to calculate $P'(y_i | X)$ for different categories, and rank the probability values of the mapping relationship between the quantified features to be classified and the categories, the largest is the category corresponding to the sample feature $X$. At the same time, calculate the corresponding classification probability.

Add new test evidence $z = \{x_k, x_2, \ldots, x_p\}$, and $P'(y_i | X)$ is transformed into prior knowledge. The conditional probability of new test evidence is $P(z | y_i)$, so $P'(y_i | X) = P'(y_i | X) \cdot P(z | y_i)$. Sort the classification probabilities $P'(y_i | X)$ to get the category with the largest probability, and calculate the posterior probability of the corresponding category.

- The operation and maintenance personnel mark the result according to the on-site confirmation result, and store the marked sample in the sample information database, update the sample set and the probability knowledge in the above steps. The sample information in different regions is summarized based on the two-layer knowledge structure [13], the prior probability and conditional probability knowledge under different categories are calculated, and the knowledge is automatically updated to the analysis and diagnosis model.

![Defect classification algorithm flow](image)
With the accumulation of sample data, the algorithm model parameters for the diagnosis of lightning arrester defects are continuously optimized, and the accuracy of classification and recognition is continuously improved. The specific implementation process is shown in Figure 1.

4. **Case Analysis**

4.1. *Sample data set and calculation of conditional probability*

Count 1,500 operation and maintenance data of a region with a voltage level of 1,000 kV to form a sample set D. The category \( C = \{ \text{normal, abnormal-aging, abnormal-wetting, abnormal-surface contamination, abnormal-other reasons} \} \) included in the sample, in which the number of lightning arrester abnormal-aging samples is 11, the number of abnormal-wetting samples is 15, the number of abnormal-surface contamination samples is 15, the number of abnormal-other reasons is 23. Due to the sparseness of the data, if the sample is classified as 0, the number of all samples will be increased by 1. Characteristic variable \( X = \{ \text{Full current of leakage current, Trend change of leakage current, Infrared abnormal heat, Partial Discharge or not} \} \). Calculate the probability \( P(C) \) of the different state types of arresters in the sample set, as shown in Table 2.

| State type                          | Priori probability/% |
|------------------------------------|----------------------|
| normal                             | 0.956                |
| abnormal-aging                     | 0.007                |
| abnormal-wetting                   | 0.011                |
| abnormal-surface contamination     | 0.009                |
| abnormal-other reasons             | 0.017                |

The calculation results of the conditional probability \( P(x|y_i) \) of the discrete feature variable are shown in Table 3.

| State type                          | Probability of infrared heating of arrester/% | Probability of non-heating infrared of arrester |
|------------------------------------|----------------------------------------------|-----------------------------------------------|
| normal                             | 0.005                                        | 0.995                                         |
| abnormal-aging                     | 0.820                                        | 0.180                                         |
| abnormal-wetting                   | 0.880                                        | 0.120                                         |
| abnormal-surface contamination     | 0.900                                        | 0.100                                         |
| abnormal-other reasons             | 0.999                                        | 0.001                                         |

The Gaussian distribution is used to represent the conditional probability distribution of continuous feature variables under different fault characteristics, and the sample mean and standard deviation of the conditional probability distribution of continuous attributes are calculated through sample sets, as shown in Table 4 and Table 5.

| State type                          | Leakage current/mA | Number of actions |
|------------------------------------|--------------------|-------------------|
|                                   | mean   | standard deviation | mean   | standard deviation |
| normal                             | 3.52   | 1.60               | 40     | 1.06               |
| abnormal-aging                     | 4.20   | 2.40               | 45     | 1.30               |
| abnormal-wetting                   | 4.53   | 1.90               | 46     | 1.02               |
| abnormal-surface contamination     | 4.60   | 1.79               | 46     | 1.40               |
| abnormal-other reasons             | 4.00   | 1.60               | 46     | 1.70               |
Table 5. Calculation results of mean and standard deviation

|                      | Growth rate of leakage current | Ratio of resistive current to full current |
|----------------------|-------------------------------|------------------------------------------|
|                      | mean %                        | standard deviation | mean % | standard deviation |
| normal               | 48                            | 12.00                      | 32     | 1.80               |
| abnormal-aging       | 80                            | 4.50                       | 40     | 2.10               |
| abnormal-wetting     | 60                            | 3.06                       | 43     | 1.90               |
| abnormal-surface contamination | 90                        | 2.45                       | 42     | 1.07               |
| abnormal-other reasons | 86                         | 2.30                       | 40     | 1.47               |

4.2. Case analysis

The above-mentioned lightning arrester defect warning algorithm based on multi-stage information is applied to the online monitoring and analysis system of lightning arresters in UHV substations [1]. During the operation of the online monitoring system, the full current and resistive current of phase B of a 1000 kV section II bus arrester continued to increase. The online monitoring data of the arrester for a continuous period of time is shown in Table 6.

Table 6. Online monitoring data

| Monitoring time | Leakage current/uA | Resistive current /uA |
|-----------------|---------------------|------------------------|
|                 | Phase A  | Phase B  | Phase C  | Phase A  | Phase B  | Phase C  |
| 2019/4/10 11:56 | 5455.1   | 5450.2   | 5448.2   | 461.4    | 471.2    | 447.1    |
| 2019/8/10 11:56 | 5478.0   | 5526.1   | 5471.2   | 749.6    | 1163.6   | 587.7    |
| 2019/8/11 11:56 | 5478.0   | 7170.9   | 5470.3   | 753.9    | 1568.9   | 595.7    |
| 2019/8/12 11:56 | 5475.0   | 9680.0   | 5468.9   | 741.3    | 259.9    | 688.4    |

According to the monitoring data of 2019/8/12, it can be seen that the ratio of the C-phase full current is larger than the other phase. The Bayesian inference model is used to calculate the C-phase classification probability based on the leakage current and the resistive current to the full current ratio, which as shown in Table 7.

Table 7. Online monitoring data

| State type                  | Probability of classification/ % |
|-----------------------------|----------------------------------|
| normal                      | 5.48×10^{-24}                    |
| abnormal-aging              | 9.73×10^{-31}                    |
| abnormal-wetting            | 7.29×10^{-25}                    |
| abnormal-surface contamination | 7.47×10^{-119}             |
| abnormal-other reasons      | 2.01×10^{-117}                   |

The classification probability calculated based on the Bayesian inference model shows that the state of the arrester is normal, but the probability of abnormal-wetting is not much different from the normal probability. The corresponding posterior probability of abnormal-wetting is as follows.

$$P(y_i = \text{abnormal-wetting} | X) = \frac{P(y_i = \text{abnormal-wetting} \cap X)}{P(X)}$$

$$= \frac{P(X | y_i) P(y_i)}{P(X)} = 0.117$$

It was found that the temperature of the phase B arrester was abnormal, the hot spot temperature was 36.6 °C, and the temperature difference between phases was 1.7 °C through the infrared temperature...
measurement of lightning arrester. Further calculate the reason for the defect of the arrester, the results are shown in Table 8.

Table 8. Probability calculation of defect type added with new evidence

| State type               | Probability/% |
|-------------------------|--------------|
| normal                  | 2.74×10⁻²⁶   |
| abnormal-aging          | 7.98×10⁻³¹   |
| abnormal-wetting        | 6.42×10⁻²⁵   |
| abnormal-surface contamination | 6.72×10⁻¹¹⁹ |
| abnormal-other reasons  | 2.01×10⁻¹¹⁷  |

The posterior probability of abnormal-moisture is calculated after the detected hot spot temperature is added.

\[ P(y_i = \text{abnormal-wetting} | X) = \frac{P(X | y_i)P(y_i)}{P(X)} = 0.959 \]

The comparison with the result of the abnormal-wetting reasoning probability calculated in the previous stage is shown in Figure 2.

![Fig. 2 Probability value in different stages](image)

Further use the lightning arrester power failure test to verify, the experimental data is shown in Table 9. The body insulation of the B phase arrester has been significantly reduced, and the arrester was damp due to the loosening of the arrester insulation base through the on-site inspection, the effectiveness of the diagnostic method was verified.

Table 9. Experimental data

| Phase       | Phase A          | Phase B          | Phase C          |
|-------------|------------------|------------------|------------------|
| Insulation resistance /MΩ | 7800/10000       | 21.6/10000       | 8100/10000       |
| This time/Handover | U1mA/kV          |                  |                  |
| U1mA/kV     | 764/775          | 331/774          | 767/778          |
| This time/Last time | ΔU1mA/%         |                  |                  |
| ΔU1mA/%     | -1.42            | -57.2            | -1.41            |

5. Conclusion

The online monitoring system can effectively monitor the change trend of the insulation state of the equipment, which provides a basis for the insulation state evaluation through the first stage of reasoning. The priority to the inspection strategy of infrared temperature measurement is given. And the second-stage probabilistic reasoning is carried out based on the new detection data through jointly applying a variety of detection methods, which could determine the reason for the possible defects of the arrester, and provide a more effective basis for decision-making on the spot. Improve the effectiveness and reliability of the algorithm through system application. The algorithm runs online, and automatically adjusts the prior probability indicators of equipment defects and causes in the model according to the monitoring data and conclusions, which could ensure the accuracy of the classification of the causes of
defects, and provide effective artificial intelligence diagnosis support for the operation and maintenance site.

6. References

[1] Wang Hongwei, Zhang Limin, Jiang Jianping, et al. On-Line monitoring and analysis system of lightning arrester leakage current in UHV substation [J]. Insulators and Surge Arresters, 2019, (6): 67-72.

[2] Khodsuz M, Mirzaie M, Seyyedbarzegar S. Metal oxide surge arrester condition monitoring based on analysis of leakage current components [J]. International Journal of Electrical Power & Energy Systems, 2015, 66 (1):188-193.

[3] Q/GDW 1168--2013, et al. Regulations of Condition-Based Maintenance & Test for Electric Equipment [S].

[4] Zhou Shuibin, Liang Wumin, Yong Mingchao, et al. A resistive current trend analysis and fault warning method for lightning arrester [J]. Insulators and Surge Arresters, 2017, (2):56-60.

[5] Ding Guocheng, Li Wei, Liu Yunpeng, et al. Experimental study of influence of environmental conditions on on-line monitoring data of metal oxide surge arrester [J]. High Voltage Engineering, 2008, 34(6): 1283-1287.

[6] Zhang Lu, Wang Yongqin, He Jiahui, et al. Accuracy examination of linkage current online measurement of surge arresters under power frequency interference of electric field and magnetic field in the ultra-high voltage substations [J]. Power System Protection and Control, 2018, (3): 164-169.

[7] Li Haifeng, Liu Junyong, Xu Bin, et al. The Modified Method of MOA On-line Monitoring Parameters by Eliminating External Environmental Factors Interference [J]. High Voltage Engineering, 2018, 44(08): 2580-2586.

[8] Chen Jingliang, Yao Xueling. Investigation of Harmonic Compensation Technique for Measuring MOV Resistive Current [J]. High Voltage Engineering, 2007(03): 53-57.

[9] Zuo Yingzhi, Cao Hongliang, Han Tong, et al. On-line aging monitoring of metal oxide arrester using particle swarm optimization [J]. High Voltage Apparatus, 2016, 52(11): 113-118.

[10] Yang Zhongjiang, Cao Hongliang, Li Pengfei, et al. On-line monitoring of metal oxide arrester using genetic algorithm [J]. High Voltage Engineering, 2015, 41(9):3104-3109.

[11] Wei Dongliang, Jiang Yiwen, Peng Hao, et al. Comprehensive state assessment model for arrester based on cloud theory and improved evidence theory [J]. Guangdong Electric Power, 2019, 32(2): 134-143.

[12] Zhang Congcong, Wang Gang, Gao Dong, et al. Partial discharge pattern recognition based on convolutional neural network [J]. Advanced Technology of Electrical Engineering and Energy, 2021, 40(3):72-80.

[13] Peng Hongxia, Wen Yan, Wang Lei, et al. Differential Operating Maintenance Technology of Power Equipment Based on Two-layer Knowledge Architecture [J]. High Voltage Apparatus, 2019, 55(07): 221-226.

[14] Williams C K I, Barber D. Bayesian classification with Gaussian processes [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(12): 1342-1351.