

Feature Selection and Fault Diagnosis Method for Switchgear Cabinet

Lifeng Liu^{1,*}, Jing Zhang², Haiyan Yao¹, Xiang Hu², Zhihao Yang¹ and Tiyyin Li¹

¹Power Supply Bureau of Yuhang, Hangzhou, China

²Zhejiang Qunli Electric Co., Ltd, Hangzhou, China

Abstract. In this paper, the scheme to monitor the fault features in switchgear is designed. From the features reflecting different types of faults such as insulation, mechanical, temperature rise, arc and so on, the original feature set for switchgear fault can be achieved which is including some diagnostic indicators. Two kinds of score with local preserving and global separation is weighted by introducing a weight coefficient, then the improved Laplacian score is formed to sort the importance level of fault features, which refines the local preserving for adjacent samples and global separation for non-adjacent samples of the features subset. By using fuzzy support vector machine (SVM) classifier to check feature subset, and then the optimal fault feature subset of switchgear is obtained. Finally, the fault diagnosis of switchgear is implemented by using Mahalanobis distance (MD) to quantify the similarity of fault features and standard samples. According to the instance analysis of monitoring data from a switching station, the correction of the proposed method is verified.

1 Introduction

The Switchgear as a very important equipment is widely used in the power generation, transmission and distribution such as the opening and closing lines and the control and protection of electrical equipment. It's safe operation will directly affect the reliability of user power supply [1]. In recent years, the switchgear fault types in the forefront sorting are insulation, mechanical, temperature, misuse and arc fault in Chinese [2-4]. Therefore, the study of switchgear fault Feature Selection and Fault Diagnosis Method is of great significance for the normal operation of switchgear.

In this paper, the scheme to monitor the fault features in switchgear is designed. Two kinds of score with local preserving and global separation is weighted by introducing a weight coefficient, then the improved Laplacian score is formed to sort the importance level of fault features. By using fuzzy SVM classifier to check feature subset, and then the optimal fault feature subset of switchgear is obtained. Finally, the fault diagnosis of switchgear is implemented by using MD to quantify the similarity of fault features and standard samples.

2 Feature selection based on improved Laplacian score

2.1 Improved Laplacian score

The Laplacian score, in the sorting of features, relies too much on the local structure information of adjacent samples [5], which makes features prominent in the representation ability to adjacent samples similarity, and weakened the distinguish ability to non-adjacent samples. So, this paper introduces Fisher score and weight coefficient to determine the separation degree of features, and which, on the one hand, makes the Laplacian score have priority in preserving local information, on the other hand, increases the separation degree of global space non-adjacent samples. After derivation, the relations between Fisher score and Laplacian score was,

$$F_r = \frac{1}{L_r} - 1 \quad (1)$$

We can see that Laplacian score can be directly calculated Fisher score. And, when a large, the Fisher score of a feature is comparatively large and its global divisibility of non-adjacent samples is comparatively strong, features should be reserved for fault discrimination. The improved Laplacian score is,

$$L'_r = \alpha L_r - (1 - \alpha) F_r = \alpha L_r - (1 - \alpha) \left(\frac{1}{L_r} - 1 \right) \quad (2)$$

Where $\alpha \in [0, 1]$, when $\alpha=1$, the improved Laplacian score L'_r degenerates into a standard Laplacian score. The smaller L_r is and the larger F_r is, the smaller the improved L'_r will be. It will make the preserving ability to local information stronger and the separation ability to global information better.

2.2 Feature selection based on improved Laplacian score

The steps of feature selection are as follows:

1) Input the original sample data of N features, and the initial value of weight coefficient is $\alpha_{k-1}=0$, $k=1$. The vector f_r' is obtained by averaging the sample data of N features. The improved Laplacian value L_r of N features is calculated by the standard Laplacian value L_r' of each feature and the weight coefficient.

2) N features in ascending order by L_r' , and N nested alternative feature subsets order in importance: $S_1 \subset S_2 \subset \dots \subset S_N$, set $i=1$, then go to step 3). Choose the selected feature subset S_i , and use the fuzzy SVM classifier [6] to check feature subset, and calculate the classification accuracy $\eta(S_i)$ of feature subset S_i .

3) If $i < N$, $i=i+1$, and turn to step 2). Otherwise stop checking the selected feature subset, and choose subset S_i which has the largest $\eta(S_i)$ as the optimal feature subset when the weight coefficient is α_{k-1} , then go to step 4).

4) If $\alpha_{k-1} < 1$, $\alpha_k = \alpha_{k-1} + \Delta\alpha$ (take $\Delta\alpha = 0.2$), $k = k + 1$, then turn to step 3); Otherwise, end the feature selection and output the subset S_i which has the largest $\eta(S_i)$ under different weight coefficient as a global optimal feature subset (if there are subsets having the largest $\eta(S_i)$ together, the subset containing the least amount of features will be chosen).

3 The scheme design of switchgear fault diagnosis

3.1. Switchgear fault features monitoring

Switchgear is used for the overhead incoming and outgoing line, the incoming and outgoing cable, the bus bar and other occasions, and its structure can be divided into four independent rooms such as the bus room, the circuit breaker handcart ventricular, the cable room and the relay instrument room [7]. From incoming line to outgoing line, there are bus room, circuit breaker room and cable room in Switchgear. The features characterizing switchgear fault can be divided into electrical quantities (voltage, current, etc.) and environmental quantities (temperature, humidity, etc.) in property, and can be divided into bus room, circuit breaker room cable room and instrument room in space. The structure for switchgear fault features monitoring is shown in Figure 1. Monitoring signals of each room are transmitted to the data acquisition module, then they are analyzed and extracted by the data acquisition module to achieve switchgear fault features diagnosis.

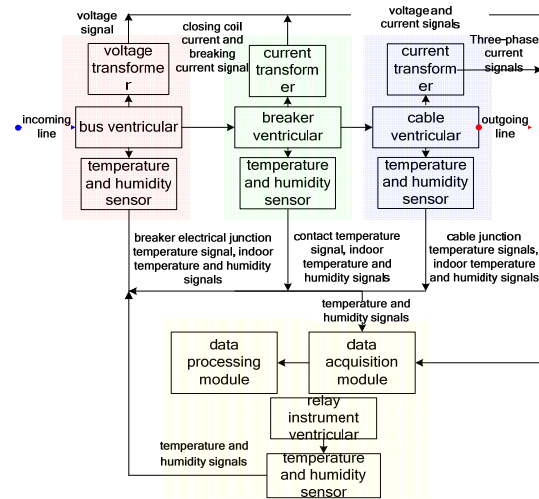


Figure 1. Structure diagram for switchgear fault features monitoring

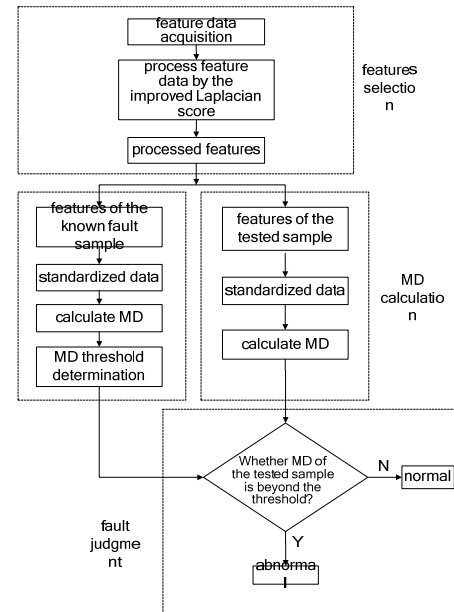


Figure 2. Flow chart for switchgear fault diagnosis

Indoor temperature $Fs1$ and humidity $Fs2$ of bus ventricular reflects the working environment of bus; temperature $Fs3$ of bus electrically connected with the circuit breaker reflects the contact problem of junction temperature rising. Voltage, current, active power, reactive power and power factor $Fs4$ to $Fs8$ reflect the internal and external system fault, and flash signal $Fs9$ is used to determine the arc fault. Indoor temperature $Fs10$ and humidity $Fs11$ of breaker ventricular reflect the working environment of breakers, as well as the changes of temperature and humidity resulting from contact fault; breaking current of breaker $Fs12$ reflects the breaking capability of breaker and the degree of relative electrical wear; temperature of breaker contact $Fs13$ reflects the changes of temperature caused by partial discharge; closing coil current of breaker $Fs14$ reflects the mechanical property of breaker. Indoor temperature $Fs15$ and humidity $Fs16$ of cable ventricular reflect the working environment of cable, as well as the changes of temperature and humidity brought from cable insulation fault; temperature of cable joint $Fs17$ reflects the temperature changes of cable tie-in caused by insulation damage. Indoor temperature $Fs18$ and humidity $Fs19$ of relay instrument ventricular reflect the working environment of switchgear instrumentation.

3.2 Method for switchgear fault diagnosis

Switchgear fault diagnosis can be achieved by quantizing the difference between the fault features in normal operation and the fault features of test samples. The process of method is shown in Figure 2. Switchgear fault diagnosis can be achieved by quantizing the difference between the fault features in normal operation and the fault features of test samples. The switchgear fault diagnosis can be divided into four parts: feature selection, fault threshold determination, MD calculation and fault diagnosis. Due to the introduction of covariance matrix,

which can minimize the influence of redundant information, but also fully consider the correlation between sample data [8]. The MD is determined the similarity of the known fault samples and the tested sampled with the sample, and the former is used to determine the MD threshold.

Set the number of each sample features is m after selecting fault features, and there are n groups sample data in the standard samples. The MD d_m of the tested sample z and the standard sample X is,

$$\begin{cases} d_m = \sqrt{(z - \bar{x})^T C^{-1} (z - \bar{x})} \\ \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i^T \end{cases} \quad (3)$$

Where z is $m \times 1$ orders tested sample vector, \bar{x} is $m \times 1$ orders characteristic mean vector, x_i in the i -th row of sample set X is $1 \times m$ orders vector, and C is $m \times m$ orders covariance matrix of sample set X ,

$$C = \frac{1}{n-1} \sum_{i=1}^n (x_i^T - \bar{x})(x_i^T - \bar{x})^T \quad (4)$$

In order to determine whether the switchgear has fault, there is a need to use a known fault sample data to determine the MD threshold d_t , then when the MD of the tested sample in $[0, d_t]$ within, the switchgear is in the normal operation, otherwise the fault will occur.

4 Case study

4.1 Results for switchgear fault feature selection

Based on the 800 groups original feature set data of normal operation and known fault types (normal sample and fault sample respectively is 300 groups and 500 groups). 400 groups are random selected as accuracy test of feature selection method and the rest are treated as the training sample of feature selection. Use the improved Laplacian score to get the fault feature subset which corresponding to different weight coefficients and ascending sort the 19 features. As shown in Table 1, the order of fault feature subset which corresponding to different weight coefficient is different.

Table 1. Fault feature sequences under different weight coefficients.

Weight coefficient	Sort of fault feature quantity
0	17, 11, 1, 19, 12, 10, 18, 7, 8, 6, 5, 15, 2, 4, 9, 16, 3, 14, 13
0.2	17, 13, 3, 9, 12, 14, 4, 5, 8, 6, 7, 15, 1, 10, 18, 16, 2, 11, 19
0.4	17, 8, 5, 9, 15, 10, 4, 3, 13, 11, 7, 18, 1, 14, 16, 12, 2, 6, 19
0.6	17, 4, 3, 6, 12, 1, 5, 14, 19, 13, 10, 15, 9, 7, 2, 11, 16, 18, 8
0.8	17, 3, 14, 13, 9, 12, 8, 4, 2, 5, 6, 7, 15, 19, 16, 18, 1, 10, 11
1	17, 6, 18, 1, 16, 14, 7, 15, 8, 9, 4, 10, 2, 5, 3, 12, 13, 11, 19

Table 2. Optimal subsets of fault features under different weight coefficients.

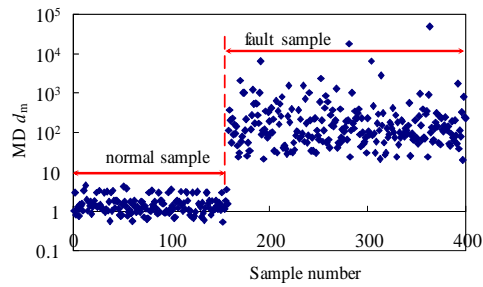
Weight coefficient	Fault feature optimal subset	Accuracy (%)
0	17, 11, 1, 19, 12, 10, 18, 7	98
0.2	17, 13, 3, 9, 12, 14, 4, 5	99
0.4	17, 8, 5, 9, 15, 10, 4, 3, 13, 11, 7	98
0.6	17, 4, 3, 6, 12, 1, 5, 14	98.5
0.8	17, 3, 14, 13, 9, 12, 8, 4	98
1	17, 6, 18, 1, 16, 14, 7, 15, 8, 9, 4	98.5

Fuzzy SVM classification is used to test the accuracy of fault feature subset classification and the accuracy which is changed with number of features under different weight coefficients is obtained. Increasing the number of fault features will not improve classification but increase the complexity of calculation when classification accuracy reaches to the maximum. Fault feature subset which corresponding to the maximum classification accuracy is the optimal subset of the weight coefficient.

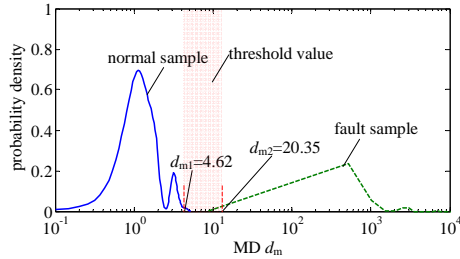
The optimal feature subset which corresponds to the maximum accuracy under different weight coefficients is shown in Table 2. From Table 2, we can see all the optimal feature subset are obtained when weight coefficient $\alpha=0.2$, and corresponding set is $\{Fs17, Fs13, Fs3, Fs9, Fs12, Fs14, Fs4, Fs5\}$.

4.2 Results for switchgear fault feature diagnosis

Use the MD of training data (157 groups normal operation and 243 groups known fault types) to confirm the MD threshold of fault diagnosis.

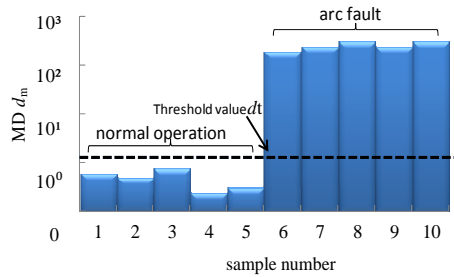


(a) MD

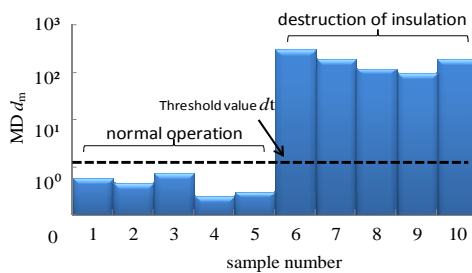


(b) probability distribution

Figure 3. MD and its probability distribution of training samples



(a) arc fault



(b) destruction of insulation

Figure 4. MD of testing samples under different operation conditions

sample are concentrated distribution around 1 and the minimum value and maximum value respectively is 0.55 and 4.62. But the minimum value of MD of fault sample is 20.35. The probability distribution of MD is shown in Figure 3b. The MD threshold of fault diagnosis shall be selected in the border area of curves (the shadow area in Figure 3b).

Calculate the MD of fault sample and normal sample through 5 groups samples of various switchgear fault operating condition, and the results are shown in Figure 4. As the Figure 4 illustrated, the MD of normal sample are lower than the threshold and the MD will bigger than the threshold when arc fault or destruction of insulation happened in switchgear.

5 Conclusion

The analyze result of site test data of 10kV switching cabinet indicate that:

1)The modified Laplace sort can describe the locality preserving ability and overall separating ability of feature quantity, and improve the accuracy of classification through remove the irrelevant redundancy quantity, which is better than normal Laplace sort.

2)Fault diagnosis method which based on MD can identify the feature quantity of normal operating and various fault condition efficiently.

3)The fault feature monitoring data can not only illustrate the each index and operating condition of switching cabinet, but also indicate the service and maintain of switching cabinet.

References

- [1] Huang Xinbo, Fang Shouxian, Wang Xiaokuan, et al. High voltage switchgear cabinet based on IoT technology [J]. Electric Power Automation Equipment, 2013, 33(2): 147-151.
- [2] Du Yanming, Gu Nihong. Present situation and accidents analysis of distribution switchgear in domestic power systems [J]. Power System Technology, 2002, 26(2):70-76.
- [3] Wang Youyuan, Liu Qiang, Lu Guojun, et al. Propagation characteristics of ultra-high frequency electromagnetic wave influenced by the switchgear structure[J]. Transactions of China Electrotechnical Society, 2013, 28(11): 293-300.
- [4] Duan Jiandong, Ye Bing, Zhang Qingshan, et al. Online contact temperature monitoring and control system based on finned radiator and ZigBee for switch cabinet [J]. Electric Power Automation Equipment, 2014, 34(7): 157-162.
- [5] Wang Lei, Liu Yan. Semi-supervised feature selection algorithm based on constraint Laplacian score [J]. Journal of Jilin University (Information Science Edition), 2010, 28(4): 404-409.
- [6] Dong Xiucheng, Tao Jiagui, Wang Haibin, et al. Incremental algorithm of adaptive fuzzy support vector machine in transformer fault diagnosis [J]. Electric Power Automation Equipment, 2010, 30(11): 48-52.
- [7] He Xiaoqun. Application of Multivariate Statistical Analysis [M]. Beijing: China Statistics Press, 2010.
- [8] Shen Zhiheng, Hu Xiande, Zhou Hao, et al. Switchgear voltage-raise reforming from 10kV to 20kV [J]. Electric Power Automation Equipment, 2012, 32(4): 122-128.