Method for Insulation Defect Pattern Recognition of Gas Insulated Switchgear based on Support Vector Machine Algorithm

XIONG Jun^{1, a*}, YANG Sen^{1,b}, LI Guangmao^{1,c}, WU Xiaogui^{1,b}, YANG fengyuan^{2,b}, SHENG Gehao^{2,b}, JIANG Xiuchen^{2,b}

¹ Electric Power Research Institute of Guangzhou Power Supply Bureau Co. Ltd. GUANG Zhou, China;² Department of Electrical Engineering, Shanghai Jiao Tong University, Shanghai City, China

Keywords: partial discharge (PD); gas insulated switchgear (GIS); support vector machine (SVM); pattern recognition.

Abstract. As partial discharge (PD) can reflect the type of insulation defects in a gas insulated switchgear (GIS) and the damage of GIS caused by different types of discharge varies evidently, identifying the type of discharge correctly is of significant value in ensuring the safe and reliable operation, assessing the insulation condition and making a rational maintenance strategy for GIS. In order to study the characteristics of PDs triggered by different defects in GIS, we designed four kinds of typical discharge defects to simulate the insulation defects that may occur in a GIS. To describe the typical characteristics of PDs, eight statistical characteristic parameters were extracted from the ultra-high-frequency signals acquired by experiments. A classifier that can achieve quaternary classification was constructed based on the support vector machine (SVM) algorithm, and then the PD type was identified by voting method. Experimental results show that the proposed method possesses a high recognition accuracy and can effectively identify four typical PDs in GIS.

Introduction

Gas insulated switchgear (GIS) has found increasing applications in power systems owing to several advantages such as compact structure, maintenance-free and low effects on environment ^[1]. Although GIS generally operates with high reliability, the defects in manufacturing or installation often lead to the occurrence of accidents. Therefore, to detect the partial discharge (PD) signals of GIS is of fairly great significance to ensuring the GIS's safe and reliable operation, assessing its insulation condition and thereby making reasonable maintenance policies ^[2]. The insulation defects in a GIS vary in type, and different defects exhibit different PD characteristics and impose different degrees of damages on the GIS's operation. To accurately assess the type and severity of the defects in a GIS, a pattern recognition should be conducted on the acquired PD signals ^[3, 4].

The common PD detection techniques include chemical detection, ultrasonic detection, pulse current method and ultra-high frequency (UHF) method ^[5]. Compared with the other PD detection methods, UHF method (300 MHz~1 GHz) can effectively avoid corona and the radio interference induced by radios and televisions, and has high sensitivity and high signal-to-noise ra¹tio. Moreover, the measured waveform and spectral features of the PD signals using UHF method can be used to distinguish different types of insulation defects and realize the pattern recognition of PDs ^[6, 7].

The PD pattern recognition methods have witnessed a great development for the past years, in which the neural network-based (NN-based) method now is most widely applied ^[8]. Neural network (NN) is a kind of machine learning method following the principle of empirical risk minimization. According to the statistical learning theory proposed by Vapnik and Chervonenkis, if the data follow a certain (fixed but unknown) distribution, the machine should abide by the principle of structural risk minimization to make the deviation between actual output and ideal output as small as possible, i.e., the upper bound of error probability should be minimized. Accordingly, a learning problem easily occurs in the applications of NN: small training errors cannot always lead to the favorable prediction results. In some cases, too small training errors may instead lead to the decline of generalization ability, i.e., the increase of actual risks ^[9]. Support vector machine (SVM) is a new learning method developed based on the statistical learning theory, and realizes the structural risk minimization theory.

It cannot only solve small-sample learning problems just like NN, but also overcome the problems such as local minimum, overfitting and underfitting. By contrast with the traditional NN, SVM has a simple structure, and moreover, can be improved in various performances especially the generalization ability, which have made SVM another research hotspot after NN^[10].

In this article, based on the characteristics of some primary insulation faults in GIS, the models of 4 typical GIS defects were designed, and an UHF-based PD detection system was constructed in the laboratory. Then, a great number of PD signal samples of various types of artificial defects were acquired. By investigating the PD characteristics of various types of defects, 8 characteristic parameters were extracted to recognize the types of PD-related defects. Finally, based on the SVM theory, a classifier for the typical defects of GIS was constructed, and the experimental results demonstrate that this classifier has a favorable recognition performance.

Experimental setting

Models of various PD defects in GIS

The primary defects leading to the PD in GIS include the internal contamination, surface burrs, poor contact, floating potential, and stains on the solid insulated surface. Accordingly, four typical insulation defects of GIS were designed, which will be described in detail below: (1) Conical needlepoint on the surface of high-voltage conductor, with the radius and length of 4 and 20 mm, respectively. The needlepoint is made up of copper and is 5 mm away from the plate electrode; (2) Void in solid insulation, i.e., the epoxy resin plate with void was placed between two plate electrodes, and the void was approximately 4 mm in diameter; (3) Free metal particles, i.e., the free-moving copper particles with the radius of 1 mm were placed on the electrodes; (4) Discharge by floating electrode, i.e., a metal plate was placed between two epoxy resin plates and then sandwiched between two plate electrodes ^[11]. The models of these four insulation defects are shown in Fig. 1 (a)-(d), respectively.



Fig. 1 Schematic diagram of defect models

UHF-based PD testing system for GIS

Fig. 2 shows the UHF-based PD testing system for GIS, from which we can observe that the system consists of a transformer, a 220kV PD-free power supply, a coupling capacitor, a detection impedance, an UHF sensor, a digital PD detector, an UHF detection system, a GIS cavity and the defect models. The above-described four typical defect models were placed into the GIS cavity filled with 0.4 MPa SF6, respectively. The upper end of the cavity was connected to a high-voltage electrode and the lower end was grounded. The voltage was gradually increased until the stable PD phenomenon could be observed in the digital PD detector, and then the UHF signals triggered by PD were collected using the UHF detection system. After coupled into the UHF sensor, the PD signals were simulated in a regulation unit and sent to the UHF PD acquisition system.



Fig. 2 UHF-based PD testing system for GIS

Experimental results

In the UHF detection system, the PD data with 50 cycles were saved as a PD sample, i.e., each PD sample contained the PD information within 1 second. Fig. 3 shows the three-dimensional (3D) spectra of four typical GIS insulation defects.





As shown in Fig. 3, the discharge amplitude and phase and the interval between two successive discharges were significantly different for different defects. To extract the corresponding characteristic parameters from these variables will undoubtedly contribute to the discrimination of these typical discharge types.

Extraction of characteristic parameters [12]

Based on the three basic parameters, namely, amplitude of PD, interval between two successive PDs (Δt) and discharge intermission (ΔT), eight relevant statistical characteristic parameters were extracted. The definitions of these three basic parameters are shown in Fig. 4. It should be noted that the discharge intermission (ΔT) denotes a fundamental frequency cycle ^[13].

The PD characteristics of two half-waves in a fundamental frequency cycle probably vary greatly. Thus, to analyze the PD characteristics of each half-wave, we should divide the whole fundamental frequency cycle into two half-cycles before data processing--positive half-cycle and negative half-cycle. The discharge amplitudes in positive and negative half-cycles constitute two PD amplitude sequences, which would be normalized in accordance with the following Eq. (1).

$$p_k = \frac{1000*(p'_k - p'_{min})}{p'_{max} - p'_{min}} \tag{1}$$

The intervals between two successive PDs in positive and negative half-cycles constitute the sequences Δt_p and Δt_n , and all the discharge intermissions ΔT in the samples constitute the sequence ΔT_w . Then the fundamental frequency cycle was averagely divided into 20 intervals. The interval in which the discharge appears can be regarded as a PD interval.



(3) Discharge intermission;

Fig. 4 Schematic diagram for fundamental parameters

According to Eq. (1), we first calculated the total numbers of PDs, $N_{mag}p$ and $N_{mag}n$, the total numbers of time intervals, $N_{int}p$ and $N_{int}n$, as well as the total numbers of PD intervals, $N_{reg}p$ and $N_{reg}n$.

The average amplitude of two half-waves can be calculated by:

$$E_{mag_p} = \frac{\sum_{k} p_{kp}}{N_{mag_p}} , \quad E_{mag_n} = \frac{\sum_{k} p_{kn}}{N_{mag_n}}$$
(2)

The standard deviation of the amplitudes of two half-waves can be calculated by:

$$\begin{cases} S_{mag_p} = \sqrt{\frac{\sum_{k} (\frac{p_{k_p}}{Emag_p} - 1)^2}{N_{mag_p}}} \\ S_{mag_n} = \sqrt{\frac{\sum_{k} (\frac{p_{k_n}}{Emag_n} - 1)^2}{N_{mag_n}}} \end{cases}$$
(3)

The average time interval of two successive PDs in two half-waves can be calculated by:

$$E_{\text{int}p} = \sum \frac{\Delta t_p}{N_{lnt_p}}, \quad E_{\text{int}n} = \sum \frac{\Delta t_n}{N_{lnt_n}} \tag{4}$$

Given these fundamental parameters, the other parameters can be calculated based on the following Eqs. (5)~(12).

$$E_{\text{mag}} = \max(E_{mag,p}, E_{mag,n}) \tag{5}$$

$$N_{mag} = N_{mag,p} + N_{mag,n} \tag{6}$$

$$S_{\text{mag}} = \max(S_{mag,p}, S_{mag,n}) \tag{8}$$

$$E_{int} = \max(E_{int,p}, E_{int,n})$$
⁽⁹⁾

$$N_{\text{int}} = N_{int,p} + N_{int,n} \tag{10}$$

$$N_{reg} = N_{reg,p} + N_{reg,n} \tag{11}$$

$$\Delta T_{mean} = \frac{\omega \omega^{\mu} w}{N(\Delta T_{w})} \tag{12}$$

where the parameters E_{mag} , S_{mag} , N_{mag} represent the amplitude characteristics of PDs; E_{int} and N_{int} represent the characteristics of time interval of PDs; ΔT_{mean} represents the intermittent characteristic of PDs; r and N_{reg} represent the PD distribution regions.

Defect recognition

SVM classifier

In this article, the original experimental samples were screened to form the model sample database. Some samples in this database were selected as the training sample, while the other samples were adopted as the test sample.

Since the SVM classification algorithm is essentially a binary classifier (i.e., a one-vs-one (1v1) classifier)^[14] and quaternary classification was required in the present experiment, we should extend this algorithm to multi-classification problems. The idea of the construction of a SVM-based

multi-pattern recognizer is to combine multiple binary classifiers. At present, there are primarily two branch algorithms, namely, one-vs-all (OVA) algorithm and 1v1 algorithm.

Compared with an OVA algorithm, each SVM classifier in a 1v1 algorithm only needs to consider two types of samples, with easy training and simpler decision boundary. Moreover, the 1v1 algorithm often has a higher classification accuracy and less computational efforts. Therefore, in this article, a SVM classifier based on 1v1 algorithm was designed for arbitrary two types of samples. For four types of samples, six SVM classifiers are required (4*(4-1)/2=6). For a simplified description, the needlepoint discharge on the surface of high-voltage conductor is denoted as the A-type discharge, the discharge by the internal void in solid insulation is denoted as the B-type discharge, the discharge is denoted as the C-type discharge and the floating electrode discharge is denoted as the D-type discharge. To construct a SVM model that can achieve quaternary classification, the following steps should be performed.

(1) The A-type and B-type training samples are taken as the original data.

(2) The original data are de-noised using db8 wavelet.

(3) Characteristic spectra are constructed and the statistical characteristic parameters are extracted.

(4) The classifier is trained using the extracted statistical characteristic parameters, to obtain a 1v1 classifier model. The normalization and standardization are first conducted on the statistical parameters, and the dimension reduction processing is then conducted using principal component analysis. The parameter optimization is achieved through grid search to search out the optimal SVM parameters. Subsequently, the constructed 1v1 classifier is trained. Fig. 5 shows the specific procedure of the construction of the 1v1 classifier model.



Fig. 5 Flow chart for the construction of a 1V1 classifier model

(5) The data of A-type and C-type training samples, A-type and D-type training samples, B-type and C-type training samples, B-type and D-type training samples, and C-type and D-type training samples are taken as the original data. Repeating the steps (1)~(4), 6 1v1 classifier models can be acquired.

(6) The test samples are to be identified. The statistical characteristic parameters of the test samples are extracted and used to test the 6 1v1 classifier models. Then, a set of results are obtained via voting. The voting procedure is illustrated in Fig. 6.





The specific voting procedure is described below.

1) Initiation, A = B = C = D = 0.

2) The statistical characteristic parameters of the test samples are input into the A-B classifier model. If the identification result is A, A = A+1; else, B = B+1.

- **u** The statistical characteristic parameters of the test samples are input into the A-C classifier model. If the identification result is A, A = A+1; else, C = C+1.
- **u** The statistical characteristic parameters of the test samples are input into the C-D classifier model. If the identification result is C, C = C+1; else, D = D+1.

3) The maximum in A, B, C and D is adopted as the identification result of the SVM model that can achieve quaternary classification.

Identification results

Using the SVM model constructed for quaternary classification, four types of test samples were identified. The identification results are listed in Table 1.

Defect model	Number of test	Number of	Number of correct	Correct
	sample	training sample	recognitions	recognition ratio
Needlepoint on the surface	110	240	110	100%
of high-voltage conductor				
Void in solid insulation	360	500	320	88.89%
Free metal particles	150	360	150	100%
Floating electrode	150	350	140	93.33%

Tab. 1 Recognition result of SVM model

As illustrated in Table 1, the correct recognition ratios for different types of defects were all above 88%. Specifically, for the discharges induced by needlepoints on the surface of high-voltage conduct and free metal particles, the correct recognition ratios were as high as 100%; for the discharges induced by floating electrodes, the correct recognition ratio was up to 93.33%; and for the discharge induced by the void in solid insulation, the correct recognition ratio was 88.89%. In some discharge cases induced by the void in solid insulation, the discharge was falsely identified as the discharge induced by floating electrode. It can be concluded from the recognition results of test

samples that, using the constructed SVM-based defect recognition system, the primary GIS discharge types can be identified in a high accuracy.

Conclusions

(1) In this article, four types of typical GIS defect models were constructed. Using the developed UHF-based PD detection system, a great number of PD sample data were acquired in the laboratory.

(2) Then we conducted appropriate processings on the PD amplitude sequence and time interval sequence, and extracted eight statistical characteristic parameters to represent the types of GIS insulation defects. These eight parameters describe the amplitude, time interval, intermission and distribution of PDs, respectively.

(3) Based on the SVM classification algorithm, arbitrary two types of samples in the four types of PD samples were selected for constructing a 1v1 classifier. Each classifier was trained using sample data, and totally 6 classifier models were constructed. The recognition results of the SVM classifier that can achieve quaternary classification were then acquired through voting. The experimental results demonstrate that, using the proposed SVM-based recognition method, the correct recognition ratios are quite high for the four types of typical GIS defects.

References

[1]HUANG Xing-quan, KANG Shu-ying, LI Hong-zhi. Research on ultra-high-frequency method for detection of partial discharge in GIS[J]. Power System Technology, 2006, 30 (7).

[2]M. D. Judd, O. Farish, and B. F. Hampton. Excitation of UHFsignals by partial discharge in GIS, IEEE Tr ans. Dielect. Elect.Insul., Vol. 3, pp. 213-228, 1996.

[3] QIU Yu-chang. GIS devices and insulation technology[M]. Xi'an, China: Xi'an Jiaotong University Press, 2007.

[4] LI Jisheng, ZHAO Xue-feng, YANG Jing-gang, et al. Measurement and analysis of partial discharge on typical defects in GIS[J]. High Voltage Engineering, 2009, 35(10): 2440-2445.

[5] XIAO Yan, YU Wei-yong, Present status and prospect of research of online partial discharge monitoring system in GIS[J]. High Voltage Engineering, 2005, 31(1): 47-49.

[6] QIAN Yong, HUANG Cheng-jun, JIANG Xiu-chen, XIAO Yan. Present situation and prospect of ultrahigh frequency method based research of on-line monitoring of partial discharge in Gas Insulated Switchgear, 2006, 30(7): 40-43.

[7] Qiu Yuchang. UHF based online insulation monitoring of GIS[J]. High Voltage Apparatus, 1997, (4): 36-40.

[8] N.C. Sahoo, M.M.A. Salama and R. Bartnikas. Trends in partial discharge pattern classification: a survey, IEEE Trans. Dielectr. Electr. Insul., Vol. 12, pp. 248-264, 2005.

[9] M.G. Danikas, N. Gao, M. Aro. Partial discharge recognition using neural networks: a review[J]. Electrical Engineering, 2003, (85): 87-93.

[10] REN Xian-wen, XUE Lei, SONG Yang, GUO Dan-dan, SHEN Zhong. The pattern recognition of partial discharge based on fractal characteristics using LS-SVM[J]. Power System Protection and Control, 2011, 39(14): 143-147.

[11] WANG Hui, ZHENG Wen-dong, WU Xiao-chun, YAO Lin-peng, HUANG Cheng-jun, QIAN Yong, Jiang Xiu-chen. Optimization method of parameter for fuzzy clustering algorithm and application in the PD pattern recognition for GIS. High Voltage Engineering, 2010, 36(12): 3002-3006.

[12] Wensheng Gao, Dengwei Ding, Weidong Liu. Research on the typical partial discharge using the UHF detection method for GIS. IEEE Trans. Power Deliver, Vol.26, pp.2621-2629, 2011.

[13] J.Y. Koo, S.Y. Jung, C.H. Ryu, S.W. Lee, B.W. Lee. Identification of insulation defects in gas-insulated switchgear by chaotic analysis of partial discharge[J]. IET Science, Measurement and Technology, 2010, 4(3):115-124.

[14] GONG Yanpeng, LIU Youwei, WU Liyuan. Identification of partial discharge in gas insulated switchgears with fractal theory and support vector machine[J]. Power System Technology, 2011, 35(3): 135-139.