Application of SA-SVM Incremental Algorithm in GIS PD Pattern Recognition

Ju Tang[†], Ran Zhuo*, DiBo Wang*, JianRong Wu* and XiaoXing Zhang*

Abstract – With changes in insulated defects, the environment, and so on, new partial discharge (PD) data are highly different from the original samples. It leads to a decrease in on-line recognition rate. The UHF signal and pulse current signal of four kinds of typical artificial defect models in gas insulated switchgear (GIS) are obtained simultaneously by experiment. The relationship map of ultra-high frequency (UHF) cumulative energy and its corresponding apparent discharge of four kinds of typical artificial defect models are plotted. UHF cumulative energy and its corresponding apparent discharge are used as inputs. The support vector machine (SVM) incremental method is constructed. Examples show that the PD SVM incremental method based on simulated annealing (SA) effectively speeds up the data update rate and improves the adaptability of the classifier compared with the original method, in that the total sample is constituted by the old and new data. The PD SVM incremental method is a better pattern recognition technology for PD on-line monitoring.

Keywords: Partial discharge, Characteristic quantity, Simulated annealing, Support vector machine, Pattern recognition

1. Introduction

Recognition of partial discharge (PD) pattern in gas insulated switchgear (GIS) is vital for on-line monitoring and faults diagnosis of electric equipment [1]. PD pattern recognition is divided into three parts, namely, data acquisition, characteristics extraction, and pattern classification. First, data are acquired on-line or off-line by the sensors and data acquisition system. Quality of data is associated with the character of sensors, accuracy of the acquisition system, interference from the environment, and so on. Subsequently, data are processed by artificial intelligence (AI) to obtain the features extracted and to finish the pattern recognition. With different AI, different process methods possess a distinct recognition ability. Thus, improving the recognition ability cannot be achieved by merely ameliorating the hardware. Determining the new characteristic quantities and better classifiers are extremely important as well.

There are two main kinds of characteristic quantities for PD recognition. The first is based on the statistical regularities of the discharge on power cycles, such as figures in two or three dimensions [2-5], while the second is based on the real-time characteristics of PD signals [6]. The former reflects better on PD's physical features. Its characteristic quantities, discharge phase, discharge quantity,

and discharge repetition rate have a tight relationship with PD's severity and an extremely high data dimension, which becomes its main shortage [3]. Considering its high data dimension, reducing dimensions is necessary. However, balancing between redundancy and abandonment of information is difficult.

The second kind emphasizes on the detailed features of PD signals from different defects. Moreover, it can provide a high recognition rate. However, a PD pulse is in nanosecond scale, which provides a strict requirement to the data acquisition system and does no good for the extension of this method. Similarly, both kinds require quantities of data that lead to difficulty in data storage forming, and training the classifier. For the reasons stated above, searching for new characteristic quantities will remain to be the hot point in this research field.

In [7], certain regularities between the apparent charge and the amplitude of ultra-high frequency (UHF) signal from a PD are found to exist. The existence of the regularity is proved in theory [8], which provides us with a good foundation. At present, the main algorithm for pattern recognition includes neural network [4, 9], cluster analysis [10], and Fisher kernel [3], to name a few. Though each offers its own advantages and has been applied to a different extent, the classifier can hardly adapt when the data updated are still the common question. A classifier using support vector machine (SVM) [15] can save memory space and simultaneously retain the information of the original classifier when data are updated. Furthermore, given that the constructed classifier depends only on the support vectors in the samples, abandonment of non-support vectors does not affect the classifier. Considering adaptability to small

[†] Corresponding Author: State Key Laboratory of Power Transmission Equipment & System Security and New Technology, Chongqing University, China. (cqtangju@cqu.edu.cn)

* State Key Laboratory of Power Transmission Equipment & System Security and New Technology, Chongqing University, China. (zrhungry@yahoo.com.cn)

Received: November 6, 2013; Accepted: May 26, 2015

sample, nonlinear, high-dimensional samples, there have been some instances in which SVM has been applied in PD pattern recognition [11, 12]. Nevertheless, SVM's advantage of fast adaptation has not been reflected in PD on-line monitoring. The study on parametric optimization for SVM requires further work as well.

In this article, we aimed at decreasing recognition rate caused by changes in insulated defects, the environment, and so on.(here we mean that as the change of types of insulated detect, the environment and other factors, the recognition rate would always decrease, thus we put our eyes on this phenomenon to reduce the decreasing level.) making the new PD data extremely different from the original sample. Four typical PD models were established to generate PD signals. The PD was detected using the UHF antenna and pulse current method simultaneously. Numerous experiments were conducted on these four models. Acquired data were divided into the following three groups: sample set T_1 for training, T_2 for testing, and incremental data set T_3 . A classifier for the PD pattern recognition was constructed using the SVM incremental method. In this process, training sample set T_1 was used to construct the original classifier. Testing sample set T_2 was used to prove its availability. Moreover, incremental data set T_3 was used to prove the fast adaptability of the classifier.

2. PD Experiment

2.1 Experiment set

To perform the experiments, we placed four types of defect models [13] in a GIS simulator: N type for needle defect, P type for particle defect, M type for metal defect, and G type for gap defect. The experiment circuit is shown in Fig. 1. The power source is a testing transformer (YDYW-25/100) free of corona, the sensor is a homemade microstrip antenna [14], while the oscilloscope is a Tectronix7104 (bandwidth 1GHz, maximum sample rate 20G/s). Environmental temperature was 10 °C. The pulse current method is adopted according to the IEC60270 standard.

2.2 Calibration for pulse current

The pulse voltage amplitude obtained using IEC60270 is

Fig. 1. Experiment circuit

Fig. 2. Calibration curve of pulse current method

Fig. 3. Actually measured signals on needle defect using UHF and IEC60270

proportionate to the apparent discharge. However, the proportionality factor is affected by the characteristics of elements in the circuit and the performance of the instrument. The proportionality factor K_c can only be obtained through calibration. Subsequently, the apparent discharge Q can be calculated. To achieve this without the power source, an adjustable PD calibrator was connected to the ends of the model. The PD calibrator produced a pulse with known discharge quantity; it caused a certain voltage pulse between the ends of the model, through which we obtained a point. Furthermore, by adjusting the PD calibrator to produce different pulses with different discharge quantities, a series of points was obtained. Accordingly, the relationship between the voltage on the detection impedance U (mV) and the apparent discharge Q (pC) was described. Fig. 2. presents their tight linear relation.

2.3 Experiment procedure

The main experiment steps are described as follows:

- 1) Source voltage was slowly increased until breakdown occurred in the test sample. The inception discharge voltage and breakdown voltage were recorded. To generate a stable PD without breakdown, three test

points were properly selected and marked as U_1 , U_2 , and U_3 between these two voltages.

- 2) The oscilloscope was set at two-channel mode to collect the PD UHF signal and pulse current signal simultaneously.
- 3) Five hundred sets of data were collected, respectively, on all three voltages and four defect models.

2.4 PD signal collected

PD signals on all four defect models were collected in the experiments. Considering the limit of length, however, we only provided the signals collected on needle defect (Fig. 3.). In the figure, the upper part represents the UHF signal in 50 mV/scale while the lower part pertains to the pulse current signal in 200 mV/scale. Time was plotted in 1 us/scale.

3. Characteristic Quantities

3.1 UHF cumulative energy W

The PD UHF signals collected in the experiments are voltage signals. Voltage points are read by oscilloscope and UHF cumulative energy is calculated using the following equation [8]:

$$W = \frac{\Delta t}{R_L} \sum_N U_n^2 \quad (1)$$

In the experiments, sample rate was 2.5 GS/s; data were recorded for 10 us. Thus, the sample interval Δt is 400 ps; R_L is a matching impedance of 50 Ω ; N is the total number of the sample points, which is $2.5\text{GS/s} \times 10\mu\text{s} = 25000$; and U_n stands for the value of the n th sample point.

3.2 Apparent discharge Q

Apparent discharge can be acquired by the following steps:

- 1) Voltage pulse signals obtained from IEC60270 are used. Amplitude of the signal H is found using a computer program or an oscilloscope.
- 2) Proportionality factor K_c is determined using the method mentioned in Section 2.2.
- 3) Apparent discharge Q is calculated as shown below:

$$Q = H \times K_c \quad (2)$$

3.3 Selection of characteristic quantities

After data were collected, the UHF cumulative energy W

(a) N type defect

(b) P type defect

(c) M type defect

(d) G type defect

Fig. 4. Relationship between cumulative energy of UHF signal and apparent discharge

and the apparent discharge Q in those four typical defect models was plotted (shown in Fig. 4.) using Equations (1) and (2).

In the figure, changing of the UHF cumulative energies with apparent discharges is observed to obey quadratic functions. Moreover, given that obvious differences exist among the four defects, we decide to use W and Q as the input characteristic quantities to achieve pattern recognition.

4. Support Vector Machine

4.1 Basic principle

To describe it simply, a support vector machine performs a nonlinear transformation, ϕ , on the input vector to map it into a considerably higher-dimensional space F , and conducts a linear separation in this space [15]. For a given sample set $\{(x_i, y_i)\}_{i=1}^N$, $x_i \in R^n$ is the input, N is the total number of samples, and $y_i \in \{\pm 1\}$ is the class number.

In this article, UHF cumulative energy W and apparent discharge Q are used as inputs. Thus, $x_i = \{(W_i, Q_i)\}$ and y_i are the class numbers of the four typical defects. If data are separable in the space F , then SVM will construct a hyperplane in that space to separate the data. The hyperplane can be described as $f(x) = (w \cdot \phi(x)) + b$, where w is the normal vector of the hyperplane and b is the offset. w and b are optimized by the following expressions:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{j=1}^N \xi_j \quad (3)$$

$$s.t. \begin{cases} y_j (w \cdot \phi(x) + b) \geq 1 - \xi_j \\ \xi_j \geq 0, j = 1, 2, \dots, N. \end{cases} \quad (4)$$

Here, ξ pertains to the relaxation factor, which is regarded as the distance between the training samples and the separate hyperplane. $C > 0$ refers to the penalty factor, whose discrimination function is as follows:

$$y = \text{sgn} \left(\sum_{x_j \in SV} (y_j a_j k(x_j \cdot x) + b) \right) \quad (5)$$

In the equation, SV stands for the support vector while $k(x_j, x)$ is the kernel function. In our research, Gaussian radial basis function (RBF) kernel, expressed in the following equation, was selected.

$$k(x_i, x_j) = \exp \left(\frac{-\|x_i - x_j\|^2}{\sigma^2} \right) \quad (6)$$

4.2 Dual-tree classification for insulation defects

Given that GIS PD pattern recognition is a multiclass question while SVM is a binary classifier, algorithms such as one-to-one, one-to-many, or dual-tree should be used to solve the multi-classification question [16]. Among these algorithms, a dual-tree requires the least SVMs. Thus, we chose dual-tree to extend SVM to a multi-classifier and perform the recognition. Its basic thinking is shown in Fig. 5.

In a training sample set with C classes, $C-1$ SVM should be trained. The first SVM, signed as SVM_1 , yields positive results to Class 1 and negative results to Classes

Fig. 5. Dual-tree SVM of insulation defects

$2, 3 \dots C$. In the same vein, SVM_i provides positive results to Class i and negative results to Classes $i+1, i+2 \dots C$. Until SVM_{C-1} is achieved, all classes are separated.

In our research, PD defects include the N, P, M, and G types. Thus, the constructed dual-tree will be similar to Fig. 5(b), where the order of N, P, M, and G is variable. SVM_N uses UHF cumulative energy W and apparent discharge Q as inputs. Furthermore, it is trained to separate type N from the others. In the same way, we can determine the usage of SVM_P and SVM_M . By these three SVMs, all four kinds of defects can be separated.

4.3 Parameter optimization in SVM by SA

Apart from kernel function and input $\{(W, Q)\}$, two parameters must be determined before the construction of SVM: penalty factor C and parameter of the kernel function σ . Determination of these two parameters has a serious effect on SVM's recognition performance. In this article, C and σ are optimized by simulated annealing (SA) [17] algorithm.

4.3.1 SA settings

The object of optimization is to achieve the maximum recognition rate, according to which the following shows several settings on SA optimization.

1) Objective function

$$F(p) = 1 - T_{train} \quad (7)$$

In the function, T_{train} is the recognition rate of the training sample set that is equal to $(N_{correct} / N_{total}) \times 100\%$, where $N_{correct}$ is the number of correct determination among the training samples and N_{total} is the total number of the training samples. To increase sensitivity in parameter searching, parameters C and σ are described as $C = 2^{p'}$, $\sigma = 2^{p''}$.

2) Renewal function for parameter p

$$p_{k+1} = p_k + \sqrt{t} * \text{randn}(2, 1) / \alpha \quad (8)$$

In the function, $\sqrt{t} * randn(2,1)/\alpha$ is a set searching step length, where α is set manually according to the amount of data, t is the temperature parameter [18] in the original SA physical model, and k refers to the iterations. In our problem, we call t as the step length controller. Given that the original solutions $[p', p'']_0$ are random, a larger step length is desired at the beginning. Thus, the initial step length controller t_0 was set to 100 according to the experience.

3) Renewal function for step length controller t

According to the needs of accuracy and searching speed, and considering the procedure of the PD SVM incremental method, two kinds of renewal functions for step length controller were designed: slow and fast. The slow kind is aimed at extending the searching range, and it is appropriate for the generation of initial SVMs from the original samples. Meanwhile, the fast kind aims at detail searching, and it is appropriate for the construction of incremental method based on the original SVM method and associates with the incremental data.

Slow renewal function for step length controller:

$$t = t_0 / \ln(k) \quad (9)$$

Fast renewal function for step length controller:

$$t = t_0 \times 0.95^k \quad (10)$$

Here, $k \geq 3$.

4) Acceptance criteria

$$\begin{cases} \text{Accept} & delE < 0 \\ \text{Accept} & h > rand \\ \text{Refuse} & \text{Others} \end{cases} \quad (11)$$

Here, $delE = F(\hat{p}) - F(p^*)$, where p^* is the present solution and $\hat{p}$ is a new feasible solution, which resulted from equation 8. If $delE < 0$, which indicates the objective function value of $\hat{p}$ is smaller than the objective function value of p^* , then the new feasible solution $\hat{p}$ is accepted. $h = 1 / (1 + \exp(delE / \max(t)))$, $h > rand$ indicate that when the step controller is a certain t , and within a set time range, $delE < 0$ has never occurred. Subsequently, a random number $rand$ (has been defined as a random number) will be generated. Moreover, if $h > rand$, the new feasible solution $\hat{p}$ is accepted.

5) Loop termination criteria

$$\begin{cases} \text{Tolerance of objective function} < 10^{-4} \\ \text{Iterations with same tolerance} > 30 \\ \text{Objective function value} = 0 \end{cases} \quad (12)$$

When any of these situations occurs, the loop will be terminated.

4.3.2 Main steps of SA algorithm

- 1) The initial step length controller t_0 is determined; let p_0 be the initial solution and p^* be the present solution; the objective function $F(p^*)$ is calculated. Subsequently, let iterations $k=0$.
- 2) A new feasible solution, $\hat{p}$ by p^* , is generated using the renewal function for parameter p , and $F(\hat{p})$ is calculated.
- 3) Whether $\hat{p}$ is accepted by the acceptance criteria is decided. If accepted, p^* is replaced by $\hat{p}$, described as $p^* = \hat{p}$. Let $k = k + 1$, and move forward. If it does not keep p^* , then we revert back to Step 2.
- 4) The step length controller t is updated using the renewal function. The loop termination criteria are checked if they are satisfied. If they are satisfied, the loop and output are terminated in the optimal solution $p_{best} = p^*$. If they fail to be satisfied, we revert back to Step 2.

5. PD SVM Incremental Recognition Method

With changes in insulated defects, the environment, and so on, the new PD data are highly different from the original samples. Subsequently, they lead to a decrease in on-line recognition rate. In this article, a PD SVM incremental recognition method [19] is presented. It can create accurate prediction toward new data samples while keeping the pre-existing information intact. It can obtain superior adaptability while saving a sizeable storage space.

Taking those four typical defects as examples the dual-tree SVM classifier, denoted by Γ , constructed with the training sample set TI , includes SVM_N , SVM_P , and SVM_M . The incremental data sets are $T3_i = \{x_i | i = 1, 2, \dots, l\}$ and $T3'_i = \{y_i | i = 1, 2, \dots, l\} = \{+1, -1\}$, where $x_i \in R^n$ are inputs, y_i are class numbers, and l is the total number of the samples. The concrete steps are described as follows:

- 1) The support vector set TI_{sv} , which is used to construct the classifier Γ is extracted from TI : $TI_{sv} = \{SV_N, SV_P, SV_M\}$. SV_N , SV_P , and SV_M are support vectors of SVM_N , SVM_P and SVM_M .
- 2) The combination of TI_{sv} and incremental data set $T3$ is taken as the new incremental sample set, expressed as follows: $TI_{sv} + T3$. Let B be the non-support vectors in TI . B is used to store the non-support vectors.
- 3) Let $T1 = TI_{sv} + T3$ and train the new classifier Γ^* . Extract the support vector set TI_{sv} and non-support vector set B^* in the new $T1$.
- 4) Let $B = B + B^*$, which is used to back up samples

that have no contribution to the classifier. Steps 1, 2, and 3 are continually repeated to learn the new samples.

- 5) Iterate for many times until the classifier converges to Γ^* .

Equations should be placed at the center of the line and provided consecutively with equation numbers in parentheses flushed to the right margin, as in (1).

$$A^T X + XA + XRX + Q = 0 \quad (1)$$

Be sure that the symbols used in your equation have been defined before the equation appears or immediately following.

It is strongly encouraged that the authors may use SI units only.

6. Instance Simulation

The 1989 PD samples obtained in our experiments were divided into three groups. Training sample set $T1$ was used to construct the original classifier; testing sample set $T2$ was used to prove its availability; and incremental data set $T3$, collected after a slight change in defect models, was used to prove the fast adaptability of the classifier. Samples in each group are shown in Table 1.

In the process of constructing the original classifier, $T1$ was used as a training sample set. The initial values of C and σ in SA algorithm were generated randomly. The parameter in searching step length α was set to 10 to improve the ability to jump out from a local optimum and to lengthen the step. The parameters in SVMs were updated using Equation 9, the slow renewal function. After 55 iterations, the acquired SVM parameters are $C = 79.26$ and $\sigma = 0.863$, and the number of support vectors is $N_{SV} = 465$. The constructed classifier was denoted as Γ , whose support vector set is $T1_{SV}$. The changing of the objective function value is shown in Fig. 6. Using classifier

Γ to recognize $T2$, we obtained 88.67% correct solutions, as demonstrated in Table 2.

In actual situation, the collected PD samples may change considerably because of the changing defects or the environment. To test the fast adaptability of the classifier, several slight changes were made on the defect models before the incremental data set $T3$ was collected.

After using the classifier Γ – trained by $T1$ only – to recognize $T3$, we obtained the result showing that the accuracy rate is extremely low. To improve this, the training sample set was replaced by combining $T1$ and $T3$. The initial values of C and σ in SA algorithm are still generated randomly. Moreover, the searching step length parameter α was set to 10 to lengthen the step and to improve the ability to jump out from a local optimum. The parameters in SVMs were updated using Equation 9, the slow renewal function. After being fully trained, that is, being subjected to 64 iterations lasting 15 hours, the objective function value reached stability. The obtained parameters are $C = 97.68$ and $\sigma = 0.832$, and the number of support vectors is $N_{SV} = 672$. Testing it on $T2$, the obtained recognition rate is 89.63%, as seen in Table 2. Noticeably, once there is a new sample set, the entire processes of construction and optimization must be repeated; thus, a large amount of time is wasted.

Adopting the PD SVM incremental method mentioned in the previous section, we replaced the training sample set as the combination of support vector set $T1_{SV}$ and $T3$. $C = 79.26$ and $\sigma = 0.863$ were used as the initial values of C and σ . Given that these values had been optimized by $T1$, performing a global optimization became unnecessary. However, a local particular searching was conducted using the fast update function, Equation 10. With $\alpha = 200$, after 46 iterations, the objective function value became stable. The changing of the objective function value is shown in Fig. 6.

In the results, $C = 96.85$, $\sigma = 0.837$, and $N_{SV} = 673$. The process lasted six hours. The acquired classifier is

Table 1. Properties of the sample

Defect type \ Groups	$T1$	$T2$	$T3$	Total
N	150	200	150	500
P	150	200	150	500
M	150	200	150	500
G	150	189	150	489
Total	600	789	600	1989

Table 2. Recognition results of incremental PD SVM

Training sample	Accuracy of training	Accuracy of recognition	N_{SV}	Training time	Iterations	C	σ
$T1$	90.6%	88.67%	465	5 hours	55	79.26	0.863
$T1+T3$	89.63%	89.63%	672	15 hours	64	97.68	0.832
$T1_{SV}+T3$	89.61%	89.63%	673	6 hours	46	96.85	0.837

PS: C and σ are the result of slow optimization

Fig. 6. The value of optimizing objective function

denoted by Γ^* , whose support vector set is T_{SV}^* . Testing on T_2 , we obtained the recognition rate of 89.63%, as seen in Table 2. Assuming that there is a new sample set T_4 , the combination of T_4 and T_{SV}^* will be the new incremental sample set. Using $C=96.85$ and $\sigma=0.837$ as the initial values, and repeating the former process, the classifier Γ^* can be updated within a short time while information on the original classifier can be kept intact.

By comparing the PD SVM incremental method based on SA optimization and the method using the combination of training sample set and incremental data set as the new training samples set, it can be observed that though both have the same recognition rate and similar results of N_{SV} , C , and σ , the incremental method decreases the samples to a large extent. Thus, training time is significantly reduced. As a result, the memory space could be saved a lot, and the higher order of magnitude of data could be proceed. Moreover, the incremental method keeps the information on the original classifier intact, and it can be updated quickly with the entry of new data.

Finally, a comparison was made among the PD SVM incremental method based on SA optimization, the neutral network based on fuzzy C-means clustering algorithm (FCM) and K-means clustering algorithm. The results we used the combination of support vector set $T1_{SV}$ and T_3 as the training sample set are shown in Table 3. The SA-SVM method has obvious advantages in all three items, which are training accuracy, recognition accuracy, and training time. More important, both FCM algorithm and K-means clustering algorithm are difficult to use in dealing with the incremental data.

Table 3. Comparison of recognition results

Method	Training accuracy	Recognition accuracy	Training time
FCM	81.2%	78.6%	180 seconds
K-means	80.3%	77.9%	125 seconds
SA-SVM	89.61%	88.32%	115 seconds

5. Conclusion

- 1) PD UHF signals and pulse current signals of all four defect models were collected simultaneously. The relationship between the UHF cumulative energy W and the apparent discharge Q was mapped out. The method of using W and Q as the characteristic quantity to recognize the four types of defects was proposed.
- 2) PD SVM incremental method based on SA optimization was constructed. It solves the problem of the drop in the PD recognition rate, which results from the relative difference between the newly acquired PD data and the original data. The said difference between the data sets results from the changing insulation defects and the environment.
- 3) The slow and fast update functions of the step length

controller, used in the optimization of SVM parameters based on SA, were provided. When parameters were unknown, the slow one should be adopted. Meanwhile, when parameters were known, the fast one was more appropriate. Thus, the classifier can be updated quickly and adaptability of the classifier is improved.

- 4) By comparison, the PD SVM incremental method was determined to have the capacity to save storage space and shorten training time to a large extent. Furthermore, the PD SVM incremental method was shown to have the capacity to deal with incremental data while keeping the original information intact, which both FCM method and K-means algorithm cannot achieve at the same time.

Even though a conclusion may review the main results or contributions of the paper, do not duplicate the abstract or the introduction. For a conclusion, you might elaborate on the importance of the work or suggest the potential applications and extensions.

References

- [1] Kranz, H. -G., "Fundamentals in computer aided PD processing, PD pattern recognition and automated diagnosis in GIS", IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 7, pp. 12-20, 2002.
- [2] Jian Li, Caixin Sun, "Partial Discharge Image Recognition Using a New Group of Features", IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 13, No. 6, pp. 1245-1253, 2006.
- [3] Xiaoxing Zhang, Jiangbo Ren, Ju Tan, Caixin Sun, "Kernel Statistical Uncorrelated Optimum Discriminant Vectors Algorithm for GIS PD Recognition", IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 16, No. 1, pp. 206-213, 2009
- [4] M. M. A. Salama, R. Bartnikas, "Determination of Neural-Network Topology for Partial Discharge Pulse Pattern Recognition", IEEE Transactions on Networks, Vol. 13, No. 2, pp. 446-456, 2002.
- [5] N. C. Sahoo, M. M. A. Salama, "Trends in Partial Discharge Pattern Classification: A Survey", IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 12, No. 2, pp. 248-264, 2005.
- [6] Yanbin Xie, Ju Tang, Qian Zhou, "Feature extraction and recognition of UHF partial discharge signals in GIS based on dual-tree complex wavelet transform", European Transactions on Electrical Power, Vol. 20, pp. 639-649, 2009.
- [7] Jin Lijun, Zhang Mingrui, Liu Weidong, "Examination and Research on the Partial Discharge Fault Diagnostics in Gas Insulated Switchgear", Transactions of China Electrotechnical Society, Vol. 20, No. 11, pp. 88-92, 2005.

- [8] TANG Ju, WU Jianrong, ZHUO Ran, XIE Yanbin, "Relationship Between VHF Signals and Discharge Magnitude of Partial Discharge from Needle Plate Electrode", High Voltage Engineering, Vol. 36, No. 5, pp. 1083-1089, 2010.
- [9] Kranz, H. -G., "Diagnosis of partial discharge signals using neural networks and minimum distance classification", IEEE Transactions on Electrical Insulation, Vol. 28, pp. 1016-1024, 1993.
- [10] Contin A., Cavallini A., Montanari G. C., Pasini G., Puletti F., "Digital detection and fuzzy classification of partial discharge signals", IEEE Trans. On Dielectrics and Electrical Insulation, Vol. 9, No. 2, pp. 335-348, 2002.
- [11] Sharkawy, R. M., Mangoubi, R.S., Abdel-Galil, T. K., Salama, M. M. A., Bartnikas, R., "SVM Classification of Contaminating Particles in Liquid Dielectrics Using Higher Order Statistics of Electrical and Acoustic PD Measurements", IEEE Transactions on Dielectrics and Electrical Insulation Vol. 14, No. 3, pp. 669-678, 2007.
- [12] L. Hao, P. L. Lewin, "Partial Discharge Source Discrimination using a Support Vector Machine", IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 17, No. 1, pp. 189-197, 2010.
- [13] Tang Ju, Zhou Qian, Tang Ming, Xie Yanbin, "Study on Mathematical Model for VHF Partial Discharge of Typical Insulated Defects in GIS", IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 14, No. 1, pp. 30-38, 2007.
- [14] Zhang Xiaoxing Tang Ju Peng Wenxiong Meng Yanhui Sun Caixin, "Study on the Outer UHF Microstrip Patch Antenna for Partial Discharge Detection in GIS", Chinese Journal of Scientific Instrument, Vol. 27, No. 12, pp. 1595-1599, 2006.
- [15] V. N. Vapnik, The Nature of Statistical Learning Theory, Springer-Verlag Inc., 1995.
- [16] Nello Cristianini, John Shawe-Taylor, Royal Holloway, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press, Inc., Cambridge, 2000.
- [17] Scott Kirkpatrick, "Optimization by Simulated Annealing: Quantitative Studies", Journal of Statistical Physics, Vol. 34, pp. 975-986, 1984.
- [18] Longhan Cao, Shanquan Zhou, Rui Li, Fan Wu, Tao Liu, "Application of optimizing the parameters of SVM using genetic simulated annealing algorithm", WCICA 2008. 7th World Congress on Intelligent Control and Automation, pp. 5381-5385, 2008.
- [19] BAI Dong-ying, WANG Xiao-dan, MA Fei, "SVM-based Incremental Learning Method and its Application", Aeronautical Computing Technique, Vol. 37 No. 4, pp. 23-26, 2007.
- [20] Joseph Vardi and Benjamin Avi-Itzhak, Electric Energy Generation; Economics, Reliability and Rates: MIT, 1981, p.75-94.

as a chief scientist he is involved in high voltage electric equipment insulation on-line monitoring and fault diagnosis.

Ju Tang was born in Pengxi, Si Chuan Province, China, in 1960. He received a Bachelor's degree from Xi'an Jiaotong University, and Master's and Doctoral degrees from Chongqing University. Dr. Tang is a professor in the Key Laboratory of High Voltage Engineering, Chongqing University. Currently,

involved in high voltage electric equipment insulation on-line monitoring and fault diagnosis.

Ran Zhuo was born in Guiyang, Gui Zhou Province, China, in 1986. He received his Bachelor's degree in Electrical Engineering at Chongqing University, China. He is now working toward his Doctor's degree in the Key Laboratory of High Voltage Engineering, Chongqing University. He is

involved in high voltage electric equipment insulation on-line monitoring and fault diagnosis.

Di-bo Wang was born in Liuzhou, Guangxi Province, China, in 1988. He received his Bachelor's degree in Electrical Engineering at Chongqing University, China. He is now working toward his Master's degree in the Key Laboratory of High Voltage Engineering, Chongqing University. He is

involved in high voltage electric equipment insulation on-line monitoring and fault diagnosis.

Jian-rong Wu was born in Pixian, Sichuan Province, China, in 1986. She received her Bachelor's degree in Electrical Engineering at Chongqing University, China. She is now working toward her Master's degree in the Key Laboratory of High Voltage Engineering, Chongqing University. She is

is involved in high voltage electric equipment insulation on-line monitoring and fault diagnosis.

Xiao-xing Zhang was born in Qianjiang City, Hubei Province, China, in 1972. He received Bachelor's and Master's degrees at the Hubei Institute of Technology and Doctoral degree at Chongqing University. Dr. Zhang is a professor in the College of Electric Engineering, Chongqing University. He