TELKOMNIKA, Vol. 11, No. 10, October 2013, pp. 5842 ~ 5850
ISSN: 2302-4046

Transformer Fault Diagnosis Based On Hierarchical Fuzzy Support Vector Machines

Xiao Yan-Cai*[1], Nan Gui-Qing[1], Zhang Qing[1], Han Xiao[2]
[1]School of Mechanical ,Electronic and Control Engineering, Beijing Jiaotong University
Beijing 100044, China
[2]Grid Control Branch, NARI Technology Development Co., Ltd., Nanjing 210061, China
*Corresponding author, e-mail: ycxiao@bjtu.edu.cn

Abstract

Large power transformer, as the key equipment of power system, plays an influential role to ensure the safe operation of power system. In this paper, transformer fault diagnosis model is built based on Fuzzy Support Vector Machines (FSVM) which combines Support Vector Machines (SVM) with fuzzy degree of membership. Hierarchical classification algorithm for multi-class classification is applied to diagnose the transformer fault. The membership value of the FSVM is obtained by Fuzzy C Means. Parameters of the FSVM model are optimized with Genetic Algorithm (GA). The transformer states are divided into trouble-free(normal), low temperature overheating T1, medium temperature overheating T2, high temperature overheating T3, low-energy discharge D1, high-energy discharge D2, and partial discharge PD. A mass of fault samples are analyzed and the results are compared with those obtained by the methods of Back-Propagation Neural Network (BPNN) and SVM, which shows that the proposed model is more effective and accurate. So the given method of transformer fault diagnosis based on Fuzzy Support Vector Machines is feasible.

Keywords: *fault diagnosis; Fuzzy Support Vector Machines; Genetic Algorithm; transformer*

1. Introduction

Transformer is an important equipment of power system. The research on transformer fault diagnosis has great realistic and practical value [1] [2].

There are multiple methods on transformer fault diagnosis. Dissolved Gas Analysis (DGA) is an effective method in detecting transformer's early fault, which is studied by many experts at home and abroad, having formed three-ratio method and improved three-ratio method in IEC standard. In recent years, many artificial intelligence methods have been applied to transformer fault diagnosis based on DGA, which has better results [3] [4] [5] [6]. The research based on Support Vector Machines (SVM) gradually becomes an important branch of transformer intelligence fault diagnosis.

Support Vector Machines, which was proposed by Vapnik in the mid-1990s, is a machine learning algorithm based on statistical learning theory. It has received increasing attention in recent years. However, it is not considered for SVM that different samples may have different impacts on the formation of optimal separating hyperplane. Therefore, the classification results based on SVM may be worse. Then the Fuzzy theory was introduced in SVM by the scholars [7]. As a consequence, Fuzzy Support Vector Machines (FSVM) is formed, which can improve SVM classification accuracy to the samples doping of fuzzy information.

In this paper, Fuzzy Support Vector Machines (FSVM) is adopted to diagnosis transformer fault. Hierarchical decision method is effectively applied to distinguish transformer's seven states. In addition, the membership value of the FSVM is obtained by Fuzzy *C*-Means algorithm (FCM). Related parameters of the FSVM model are optimized with Genetic Algorithm (GA). The result indicates that the model has higher accuracy.

2. Fuzzy Support Vector Machines

The Fuzzy Support Vector Machines' core idea is to combine SVM with fuzzy degree of membership. According to the contribution of different samples to classification, a corresponding

Received April 22, 2013; Revised July 3, 2013; Accepted July 17, 2013

degree of membership is endowed with so that the impact of outliers is reduced and the classification performance is improved. In this ideology, training samples no longer strictly belong to a class in the two categories. Such the circumstance exists: a training sample 80% possibly belongs to a class, 20% possibly does not fall into this category. That is, for each sample, there is a corresponding fuzzy value $\mu_i, 0 \le \mu_i \le 1$, which is usually called the sample's fuzzy degree of membership.

Assume that the training samples are (y_i, x_i, μ_i) , $i = 1, 2, \cdots n$. Here, $x_i \in \mathbf{R}^m$ is the input; $y_i \in \{-1, 1\}$ is the output; μ_i is fuzzy degree of membership of the training sample.

The optimization problem of FSVM optimal hyperplane is:

$$\min \frac{1}{2} \omega \cdot \omega + C \sum_{i=1}^{n} \mu_i \xi_i \tag{1}$$

Here, ω is hyperplane's normal; *C* is penalty factor; ξ_i is slack variable; $\mu_i \xi_i$ is weight control of misclassified samples. The constraint condition is:

$$y_i(\omega \cdot x_i + b) \ge 1 - \xi_i, i = 1, 2, \cdots n, \xi_i \ge 0$$

Here, *b* is the intercept.

Optimal solution of the problem can be given by the Lagrange functional:

$$L(\omega, b, \alpha, \beta) = \frac{1}{2}(\omega \cdot \omega) + C \sum_{i=1}^{n} \mu_i \xi_i - \sum_{i=1}^{n} \alpha_i \{ y_i [(\omega \cdot x_i) + b] - 1 + \xi_i \} - \sum_{i=1}^{n} \beta_i \tag{2}$$

Here, $\alpha_i and \beta_i$ are Lagrange factors.

$$\overline{\alpha} = \arg \min_{\alpha} \frac{1}{2} \sum_{i=1}^{n} \sum_{i=1}^{n} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) - \sum_{i=1}^{n} \alpha_i \qquad i = 1, 2, \cdots n , \quad j = 1, 2, \cdots, n \tag{3}$$

The constraints are:

$$0 \le \alpha_i \le \mu_i C, \sum_{i=1}^{n} \alpha_i y_i = 0$$

$$\alpha_i \{ [(\omega \cdot x_i) + b] y_i - 1 + \xi_i \} = 0, i = 1, 2, \cdots, n$$

$$(u_i C - \alpha_i) \xi_i = 0, i = 1, 2, \cdots, n$$

During the computation, the inner product operation $(x_i \cdot x_j)$ is generally replaced with the kernel function $K(x_i \cdot x_j)$.

3. Fuzzy Support Vector Machines for Transformer Fault Diagnosis

Characteristic gases dissolved in oil for transformer fault diagnosis are mainly composed of H_2、 CH_4、 C_2H_6、 C_2H_4 and C_2H_2. In this paper, volume fractions of these five kinds of gases are taken as inputs. In order to reduce the impact of value difference between the various gases, DGA initial data (including the training and testing samples) are normalized. The built-in function– mapminmax – in MATLAB2009a is used for normalization.

3.1. The Establishment of Hierarchical Fault Classification Model

According to IEC60599, transformer fault can be divided into six kinds: low temperature overheating, high temperature overheating, medium temperature overheating, low-energy discharge and high-energy discharge and partial discharge. Coupled with the normal state, there are seven states. Transformer fault is identified with the hierarchical structure step by step [8]. And the general classification of FSVM model is shown in Figure 1.

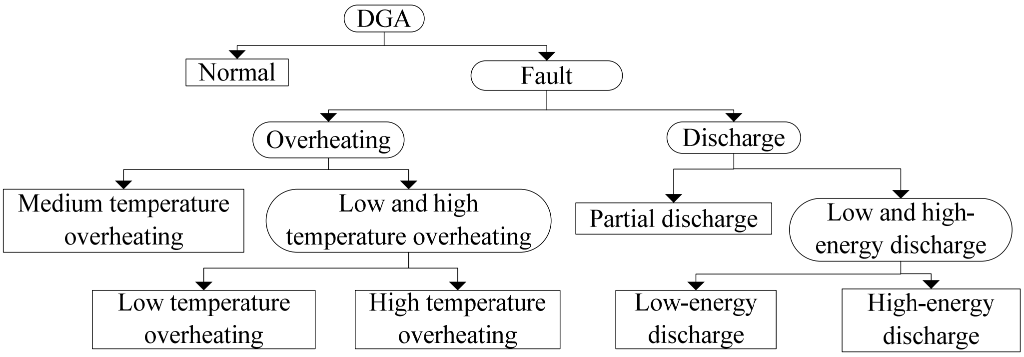

Figure 1. The general classification of FSVM model

3.2. The Process of Transformer Fault Diagnosis

The process of transformer fault diagnosis is as follows.

(1) Determine Training and Testing Samples

Divide the collected DGA gas data into training samples and testing samples, and number the fault types of samples with the following format: 1: trouble-free, 2: low temperature overheating, 3: high temperature overheating, 4: medium temperature overheating, 5: low-energy discharge, 6: high-energy discharge, 7: partial discharge.

(2) Pretreat Data

Data pretreatment mainly makes the data normalized. The purpose of normalization when establish the sorter is to enhance the training convergence and computation speed, as well as guarantee the sample's characteristic information as far as possible.

(3) Calculate Fuzzy Degree of Membership

There are many ways to calculate the membership degree of Fuzzy Support Vector Machines, Such as membership degree method based on distance, Fuzzy *C*-Means algorithm (FCM) [9] and *S*-type membership function. Among these methods, Fuzzy *C*-Means algorithm with simple, rapid superior performance both classifies the data with quite high accuracy and determines each sample's membership degree value under various categories. Therefore Fuzzy *C*-Means algorithm can be used to obtain the degree of membership μ_i in FSVM.

(4) Select Parameters of FSVM

RBF kernel function is selected in FSVM. Its formula is:

$$K(x_i \cdot x_j) = \exp\left(-\frac{(x_i - x_j)^2}{2\sigma^2}\right) \tag{4}$$

Where σ is width parameter. $g = 1/2\sigma^2$.

The selection of penalty factor *C* and RBF kernel function parameter *g* has a major impact on the classification accuracy in the use of Fuzzy Support Vector Machines. The value of *C* is smaller, the generalization ability of the algorithm is better. That the value of *g* is too small will have overfitting influence on the samples.

Now parameter optimization algorithm which has been commonly used includes Grid Search algorithm, Genetic Algorithm and so on [10] [11]. Generally, the computation amount of Grid Search is large. And when sample types are more and training set is bigger, the amount of training and testing model will be quite large, and computation time will be long. Moreover, parameters obtained through Genetic Algorithm correspond with higher diagnostic accuracy

[12]. For these reasons, Genetic Algorithm is used to optimize the parameters of FSVM. Specific process is shown in Figure 2.

(5) Training

a) In training samples, the fault samples will be unified numbering as -1, and trouble-free samples as 1. We use Genetic Algorithm (GA) to optimize parameters to define appropriate *C* and *g*, and then obtain the first level model (FSVM1) through training.

b) Extract fault samples, and renumber them. Discharge fault samples are marked as -1, and overheating fault as 1. Determine the appropriate *C* and *g*, and then get the second level model (FSVM2) through training.

c) Extract overheating fault samples, and renumber them. Medium temperature overheating samples are marked as -1, high and low temperature overheating samples as 1. Determine the appropriate *C* and *g*, and then get the third level model (FSVM3) through training.

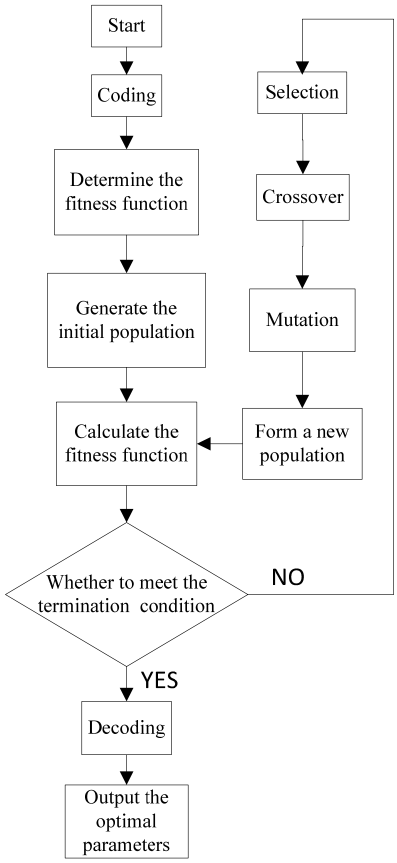

Figure 2. The process of Genetic Algorithm

d) Extract high and low temperature overheating fault samples, and renumber them. Low temperature overheating samples are marked as -1, high temperature overheating samples as 1. Determine the appropriate *C* and *g*, and then get the fourth level model (FSVM4) through training.

e) Extract discharge fault samples, and renumber them. Partial discharge fault samples are marked as -1, other discharge fault samples as 1. Determine the appropriate *C* and *g*, and then get the fifth level model (FSVM5) through training.

f) Extract other discharge fault samples, and renumber them. Low-energy discharge fault samples are marked as -1, and high-energy discharge fault samples are marked as 1. Determine the appropriate *C* and *g*, and then get the sixth level model (FSVM6) through training.

So far, six classification models based on Fuzzy Support Vector Machines have been established. Meanwhile, the learning machine acquires classification knowledge implied in the characteristic data of the key nodes. The trained FSVM model can be used to get the classification of seven states of the transformer.

(6) Testing

Firstly, the testing samples should be normalized in accordance with the method in step 2. These treated samples will be input into first level FSVM model to test whether they are breakdown, and then record the testing results. Secondly, carry on overheating and discharge fault category testing to the samples determined as fault. Input the testing samples determined as overheating or discharge fault into classifier to distinguish overheating fault and discharge fault until obtain the final fault type. The training and testing concrete process of transformer fault diagnosis model is shown in Figure 3.

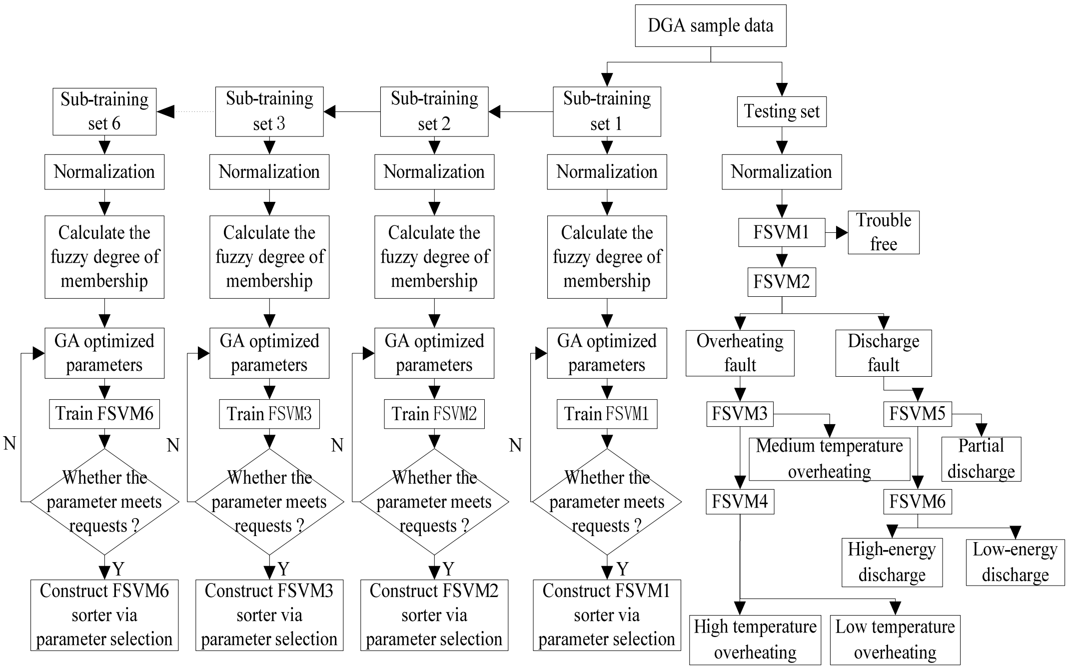

Figure 3. General Realization of the transformer fault diagnosis based on FSVM

4. Examples of Transformer Fault Diagnosis

Through consulting the relevant literature in recent years, and collecting a large number of transformer fault data having clear conclusions from historical data of Beijing Electric Power Corporation, 722 samples are selected. Among these samples, there are 615 groups taken as training set, 107 groups as testing set. 300 groups of samples in training set are normal data, and other 315 groups are fault data. 43 groups of samples in testing set are normal data, and other 64 groups are fault data. The volume fractions of these five kinds of gases – H_2、CH_4、C_2H_6、C_2H_4 and C_2H_2, are used to be the input. The seven states of the transformer – low temperature overheating (T<300℃, marked as T1), medium temperature overheating (300℃≤T<700℃, marked as T2), high temperature overheating (T≥700℃, marked as T3), low-energy discharge (marked as D1), high-energy discharge (marked as D2), partial discharge (marked as PD) and normal state – are taken as output. Distributions of various transformer states in sample sets are shown in Table 1.

Table 1. Fault distributions in sample sets for training and testing

	Normal	Low temperature overheating	Medium temperature overheating	High temperature overheating	Low-energy discharge	High-energy discharge	Partial discharge
Training set	300	45	89	54	39	45	43
Testing set	43	14	10	14	7	8	11

(1) Any Fault Diagnosis Model (FSVM1)

The collected 615 groups of sample data are taken as training set, 300 groups of normal data are marked as 1, and other 315 groups of fault data are marked as -1. Normalize training set to 0 to 1, and calculate membership degree of the normalized sample data, obtaining membership matrix. The matrix is:

$$U = \begin{bmatrix} 0.05057\ 0.33029\ ...0.55289\ 0.84353...0.93833\ ...0.16504... \\ 0.94943\ 0.66971\ ...0.44711\ 0.15647...0.06167\ ...0.83496... \end{bmatrix}$$

The first line represents membership degrees that belong to the first category; the second line represents membership degrees that belong to the second category. It can be seen from the membership matrix that there are outliers in the two types of samples.

The optimal parameters can be obtained by GA method. In this step, maximum evolution generation: maxgen=100. The number of individuals is taken as 20. *C* is in the range of [0.01, 1000], and *g* is in the range of [0.001, 100]. The Generation Gap is 0.9. Crossover rate is 0.4. Mutation rate is 0.01. The final optimal parameters are: *C*= 907.3964, *g*= 0.0313. And now the diagnostic accuracy rate of training set is 98.81%. The diagnostic accuracy rate of testing set is 96.26%.

In order to evaluate the effectiveness of fault diagnosis model based on Fuzzy Support Vector Machine, a BP Neural Network and SVM diagnostic model are established. The BP Neural Network is in the structure of 5-12-7(Number of neurons of input layer is 5, indicating the 5 kinds of gases of the samples. Number of neurons of hidden layer is 12. Number of neurons of Output layer is 7, indicating 7 kinds of states). Carry on diagnosis to the same data using different diagnosis method. The comparison of the results is shown in Table 2.

Table 2. The comparison of diagnosis results of FSVM1

	BP Neural Network		SVM		FSVM	
	Time	Accuracy rate	Time	Accuracy rate	Time	Accuracy rate
Training set	6.77s	99.57%	2.8667s	94.76%	4.0174s	98.81%
Testing set	1.54s	92.52%	0.7598s	91.59%	0.7829s	96.26%

It can be seen from Table 2 that compared with SVM method, training the model consumes more time when FSVM method is used. The reason for this is that each sample's membership degree has been calculated in FSVM model. However, time-consuming difference between two methods is not large when testing the same data. Yet the accuracy rate is improved significantly using FSVM method.

(2) Fault Type Identification Model (FSVM2)

The normalized values of five kinds of gases are as input values. The optimal parameters are obtained with GA, whose result is *C*=320.5748, *g*= 0.0779. Meanwhile, the training set diagnosis accuracy is 96.19%, and the diagnosis accuracy rate of testing set reaches 95.31%. The same data is diagnosed with BP Neural Network (hidden layer neurons for 12) and Support Vector Machines. Accuracy comparison of various diagnosis methods is shown in Table 3.

Table 3. The comparison of diagnosis results of FSVM2

	Three-ratio		BP Neural Network		SVM		FSVM	
	Time	Accuracy rate	Time	Accuracy rate	Time	Accuracy rate	Time	Accuracy rate
Training set	0.0382s	80.00%	10.2173s	98.10%	1.7911s	94.29%	2.9073s	96.19%
Testing set	0.0231s	82.81%	1.0531s	92.19%	0.9052s	93.75%	0.9221s	95.31%

(3) Overheating Fault Identification Diagnosis Model (FSVM3/ FSVM4)

Overheating fault contains low temperature overheating, medium temperature overheating and high temperature overheating. Hierarchical classification algorithm for multi-class classification is applied to separate the three faults correctly. Therefore, two Fuzzy Support Vector Machines models (FSVM3/FSVM4) need to be constructed.

Firstly, all the overheating fault samples are extracted from training samples. Medium temperature overheating samples are marked as -1, high and low temperature overheating samples as 1. Binary Tree classifier (FSVM3) is structured to identify medium temperature overheating fault. Appropriate *C* and *g* are obtained with GA, whose optimal result is *C*=74.8815, *g*=0.0313. Classification results are shown in Table 4.

Secondly, low temperature overheating and high temperature overheating are separated. Another Binary Tree classifier (FSVM4) is structured. High temperature overheating samples are marked as 1, and low temperature overheating samples as -1. After training and testing, table 5 is obtained.

(4) Discharge Fault Identification Diagnosis Model (FSVM5/ FSVM6)

Discharge fault contains partial discharge, low-energy discharge and high-energy discharge. Being similar to overheating fault, two Binary Tree classifiers (FSVM5/ FSVM6) need to be structured.

Firstly, all discharge fault samples are extracted from training samples. Partial discharge samples are marked as -1, other discharge fault samples as 1. Train the first Fuzzy Support Vector Machines model for discharge fault identification. Diagnosis results are shown in Table 6.

Table 4. The comparison of diagnosis results of FSVM3

	Three-ratio		BP Neural Network		SVM		FSVM	
	Time	Accuracy rate	Time	Accuracy rate	Time	Accuracy rate	Time	Accuracy rate
Training set	0.0314s	78.19%	7.7121s	98.94%	0.9538s	92.02%	2.1278s	93.09%
Testing set	0.0211s	78.94%	0.8426s	86.34%	0.6408s	89.47%	0.6017s	92.11%

Table 5. The comparison of diagnosis results of FSVM4

	Three-ratio		BP Neural Network		SVM		FSVM	
	Time	Accuracy rate	Time	Accuracy rate	Time	Accuracy rate	Time	Accuracy rate
Training set	0.0237s	78.79%	4.1480s	100%	0.8669s	96.94%	1.3620s	96.97%
Testing set	0.0175s	78.57%	0.5475s	82.14%	0.5159s	96.42%	0.8541s	100%

Table 6. The comparison of diagnosis results of FSVM5

	Three-ratio		BP Neural Network		SVM		FSVM	
	Time	Accuracy rate	Time	Accuracy rate	Time	Accuracy rate	Time	Accuracy rate
Training set	0.0295s	81.10%	7.0438s	98.43%	1.4904s	95.27%	1.9420s	96.85%
Testing set	0.0276s	80.76%	0.7897s	84.62%	0.8122s	92.31%	0.9017s	96.15%

The second Fuzzy Support Vector Machines model based on low-energy discharge and high-energy discharge is constructed. High-energy discharge fault samples are marked as 1, low-energy discharge fault samples as -1. Classification results are shown in Table 7.

From the above, when distinguishing discharge or overheating fault, three-ratio method shows well, whose diagnostic accuracy rate is 80% or so. The results indicate that when carrying on diagnosis to training set with BP Neural Network, the accuracy rate can achieve a higher value. But when diagnosing testing set with BP Neural Network, the accuracy rate will decline significantly. So its generalization performance is poor. That BP Neural Network is gradually approaching model with the slower convergence speed results in the longer training time. From a certain extent, the deficiency of BP Neural Network is overcome by Support Vector Machine (SVM) model. Diagnosis effect of the SVM model is improved slightly and its training time is reduced significantly. The accuracy rate of the FSVM model surpasses that of the SVM slightly. Only when outliers are existent in training samples or the data is too scattered, Fuzzy Support Vector Machines will show superiority obviously.

Table 7. The comparison of diagnosis results of FSVM6

	Three-ratio		BP Neural Network		SVM		FSVM	
	Time	Accuracy rate	Time	Accuracy rate	Time	Accuracy rate	Time	Accuracy rate
Training set	0.0197s	78.59%	2.1523s	97.57%	0.9469s	89.89%	1.0752s	95.24%
Testing set	0.0132s	80.00%	0.5270s	82.24%	0.6241s	92.11%	0.5748s	100%

To summarize, input data requirements are reduced with the FSVM than with the SVM. Due to the existence of error accumulation, misclassification will keep till final results. It is difficult for the diagnostic accuracy rate of hierarchical classification method for multi-class classification based on Fuzzy Support Vector Machines to reach 100%. However, the number of classifier is greatly reduced and the diagnosis accuracy is improved with this method.

5. Conclusion

In the paper, hierarchical classification algorithm for multi-class classification is used to establish transformer fault diagnosis model based on Fuzzy Support Vector Machines. The membership degree of the FSVM is obtained by Fuzzy *C*-Means algorithm; the parameters of the FSVM are optimized with Genetic Algorithm. The transformer seven states classification is realized. Transformer fault diagnosis examples demonstrate the advantages of the model based on Fuzzy Support Vector Machines. However, the research on FSVM has only just started. There are many issues that need further exploring, and the application on transformer fault diagnosis also needs to be examined thoroughly.

Acknowledgement

This work was supported by the Fundamental Research Funds for the Central Universities (2013JBM078).

References

[1] YANG Qi-ping, XUE Wu-de, LAN Zhi-da. *Development of an Artificial Intelligent Diagnosis System for Transformer Fault*. IEEE/PES Conference and Exhibition: Asia and Pacific, Dalian, China. 2005; 1-5.
[2] SONG Bin. Study on Fault Diagnosis of Transformer based on Dissolved Gas-in-oil Analysis. PhD Thesis. Wuhan: Wuhan University; 2003.
[3] CAO Jian, QIAN Su-xiang, HU Hong-sheng *et al*. *Development of Fault Diagnosis System for Transformer Based on Multi-class Support Vector Machines*. International Conference on Mechatronics and Information Technology. Gifu, Japan. 2007; 6794: 1-6.
[4] Dae-Jong Lee, Jong-Pil Lee, Pyeong-Shik Ji *et al*. *Fault Diagnosis of Power Transformer Using SVM and FCM*. IEEE International Symposium on Electrical Insulation. Vancouver, Canada. 2008; 112-115.
[5] XIAO Yan-cai, ZHU Heng-jun. Power transformer fault diagnosis based on least square support vector machine. *Electric power automation equipment.* 2007; 27(9): 48-50.
[6] Castro ARG, Miranda V. Knowledge discovery inneural networks with application to transformer failure diagnosis. *IEEE Transactions on Power Systems*. 2005; 20(2): 717-724.
[7] Fengqing Han, Hongmei Li, Cheng Wen, Wenjuan Zhao. A new incremental support vector machine algorithm. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2012; 10(6): 1171-1178.
[8] CHEN Wei, LIU Jun-yao, QI Dongke *et al*. Application of fault diagnosis for an equipment based on binary tree support vector machines. *Avionics technology.* 2010; 41(2): 19-23.
[9] YANG Yi-fang. Research of Fault Diagnosis Method Combined by SVM and FCM. PhD Thesis. Xi'an: Xi'an University of Science and Technology; 2008.
[10] XIANG Chang-sheng, ZHOU Zi-ying, XIA Yan-jun. Study on parameters optimization of support vector regression based on heuristic breadth-first search algorithm. *China Journal of Bioinformatics*. 2010; 8(3): 219-222.
[11] LIU Zhen, HU Yun-an, SHI Jian-guo. Estimation of Distribution Immune Genetic Algorithm and Convergence Analysis. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(1): 123-129.
[12] XIAO Yan-cai, CHEN Xiu-hai, ZHU Heng-jun. The application of genetic support vector machine in power transformer fault diagnosis*. Journal of Shanghai Jiaotong University.* 2007; 41(11): 1878-1881.