2018 17th International Symposium on Distributed Computing and Applications for Business Engineering and Science

Study on Multi-RBF-SVM for Transformer Fault Diagnosis

QU Li-ping[1*], ZHOU Hao-han[2], LIU Chong-jie[2], LU Zhao[2]
[1*]Engineering Training Center of Beihua University, Beihua University
[2]College of Electrical and Information Engineering, Beihua University
Jilin, China
[*]*E-mail: 309637683@qq.com*

***Abstract*—Two improved fault diagnosis algorithms were proposed in this paper. One is New Three-Ratio (NTR) algorithm which introduces the code "000" as normal type code into the Transformer Three-Ratio(TTR).The other is Multi-Radial Basis Function-Support Vector Machine (M-RBF-SVM) algorithm which introduces the multi-RBF kernel function. And, the representative Tradition Three-Ratio algorithm is selected as the simulation comparison object. The results indicate that M-RBF-SVM can achieve higher diagnosis accuracy and excellently generalization ability.**

Keywords-fault diagnosis; new three-ratio; M-RBF-SVM; Kernel function; accuracy

I. INTRODUCTION

Transformer is one of the most important electrical equipment in the power system. It plays an important role of the safe and stable operation in power system. According to normal transformer fault diagnosis, the gas in transformer oil is the key point of analysis. Transformer daily operation and maintenance require the staff to detect the dissolved gas content by using the gas chromatography analyzer regularly and irregularly [1]. The commonly used diagnostic methods include the characteristic gas method[1,2], the classic three-ratio method [3,9], four-ratio method[4], graphics and the artificial intelligence method[5,6,8,10,11,12,13,14], etc. Due to the missing code, the three-ratio method has not been widely used. And the four-ratio method is used to resolve the diagnostic accuracy of each step. Graphics method can achieve a better correction rate, but it needs a lot of actual fault data.

The Tradition Three-Ratio (TTR) and New Three-Ratio (NTR) often need more fault samples to obtain feasible fault diagnosis model, butin fact there are no enough reality data. The Support Vector Machine (SVM) is a class of supervised learning algorithms based on the statistical learning and machine learning [6], and it can solvehigh dimension and local minima problem of neural network algorithm. Especially, it has good processing ability for the limited samples [15,16]. But the Traditional Support Vector Machine (TSVM) is not useful for the transformer fault diagnose. Thus, the other Multi-Radial Basis Function-Support Vector Machine (M-RBF-SVM) is proposed.

II. PRINCIPLE OF THE FAULT DIAGNOSIS

A. Principle of Transformer Fault Diagnosis

Transformer will work for a long time if there is no something wrong with it. Up to right now, the transformer fault diagnosis gas analysis (DGA) is a hot research topic. The common analysis gases are five kinds of gases: hydrogen(H_2), methane(CH_4), ethane(C_2H_6), ethylene (C_2H_4), acetylene(C_2H_2). The transformer fault types were figured out by the corresponding relationship between the different gas content.

B. NTR Algorithm of Fault Diagnosis

TTR is a set of encoding rules, which can be used to analyze the transformer fault data[1]. The three ratio of the five gases respectively are: $x1=C_2H_2/C_2H_4$; $x2=CH_4/H_2$; $x3=C_2H_4/C_2H_6$. At present, the TTR method has a higher accuracy[11], but still lack encode. The new encoding combination 000 is added in the encoding rules to improve the diagnostic accuracy based on the encoding rules. TABLE I lists the corresponding types of new encoding.

TABLE I. NEW ENCODING TYPE FAULT LOOKUP TABLE

Types	Fault Types	Actual test fault	Encoding combination		
			X1	*X2*	*X3*
1	Low temperature overheat (<150℃)	Line overheat	0	0	1
	Low temperature overheat (150~300℃)	Tap-changer poor contact, lead screw loosening or bad welding joint, eddy current fever, iron core magnetic flux leakage, short circuit and multi-point earth		2	0
2	Middle temperature overheat (300~700℃)			2	1
3	High temperature overheat (>700℃)			0,1,2	2
4	Partial discharge	Moisture, gas discharge		1	0
5	Low energy discharge	Spark discharge between the floating potential components	2	0,1	0,1,2
6	Low energy discharge and overheating	Tap and oil discharge gap between different parts		2	0,1,2
7	Arc discharge	Between interterm, interlayer, alternate with, lead or to discharge, tap-changer arc discharge caused by circulation	1	0,1	0,1,2
8	Arc discharge and overheating			2	0,1,2
9	Normal	Safe	0	0	0

C. Support Vector Machine Foundation

SVM is a linear classification method; such the Fig.1 is classified into two parts. The circular and the triangles represent two different kinds, respectively. A suitable straight line $(\omega \cdot \mathrm{x}) + b = 0$ is the key to solve the follow two classification problems.

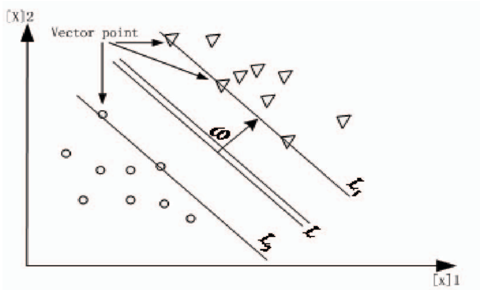

Fig.1 Two classification diagram

First, the two sides are defined as 1 and -1 (representing the positive class and negative class, respectively). The (1) plane is located at the sides of the plane, and (2) of the classification ||ω|| is the biggest, which can be constrained optimization problem by (3).

$$\begin{cases} < \omega \cdot x > +b = 1 & y_i = 1 \\ < \omega \cdot x > +b = -1 & y_i = -1 \end{cases} \tag{1}$$

$$d = \frac{|1-(-1)|}{\sqrt{\omega_1^2 + \omega_2^2 + \cdots + \omega_N^2}} = \frac{2}{\|\omega\|} \tag{2}$$

$$\begin{cases} \Phi(\omega) = \frac{1}{2} \|\omega\|^2 \\ y_i(\omega \cdot x_i + b) \ge 1 \quad i = 1, 2, \cdots n \end{cases} \tag{3}$$

It can be concluded that all the i $(y_i = 1)$ can make the $(\omega \cdot x_i) + b \ge 1$ truly and all the i $(y_i = -1)$ can make the $(\omega \cdot x_i) + b \le -1$ true. Then, the Lagrange function can be constructed to solve the optimization problem. The Lagrange function can be expressed in (4) after introduced the Lagrange :

$$L(\omega, b, \alpha) = \frac{1}{2}(\omega^T \cdot \omega) - \sum_{i=1}^{n} \alpha_i \{y_i[\omega \cdot x_i + b] - 1\} \tag{4}$$

The Partial Differential Equations of (4) on the ω and b can be used to get the ω and b through (5).

$$\begin{cases} \omega = \sum_{i=1}^{l} \alpha_i y_i x_i \\ \sum_{i=1}^{l} \alpha_i y_i = 0 \end{cases} \tag{5}$$

When combining (4) and (5), (6) can be obtained:

$$\begin{aligned} L(\omega, b, \alpha) &= \frac{1}{2}(\omega^T \cdot \omega) - \sum_{i=1}^{n} \alpha_i \{y_i[\omega \cdot x_i + b] - 1\} \\ &= \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \end{aligned} \tag{6}$$

When subject to $\sum_{i=1}^{n} \alpha_i y_i = 0, \alpha_i \ge 0$, finding the L maximum value is important. Here α_i^* is the best value and $\omega^* = \sum_{i=1}^{n} \alpha_i^* y_i x_i$ to get the line $(\omega^* \cdot x) + b = 0$. Lastly, using (7) to find labels.

$$y = \mathrm{sign}(a) = \begin{cases} 1 & a \ge 0 \\ -1 & a < 0 \end{cases} \tag{7}$$

That is shown as (8):

$$\begin{aligned} f(x) &= \mathrm{sgn}[(\omega^* \cdot x) + b^*] \\ &= \mathrm{sgn}(\sum_{svs} \alpha_i^* y_i < x_i \cdot x > + b^*) \end{aligned} \tag{8}$$

Here $(svs) = \{x_i \mid \alpha_i^* > 0, i = 1, 2, \cdots, N_s\}$,Ns is the number of the SVM, and others are $\alpha_i = 0$.

The above process is a linear classification based on SVM, but the transformer fault diagnosis is a process of nonlinear classification.In order to solve the nonlinear classification problem, the kernel function is introduced[5]. Thus, the optimal classification function can be changed into (9).

$$f(x) = \mathrm{sgn}(\sum_{svs} \alpha_i^* y_i k(x_i \cdot x) + b^*) \tag{9}$$

Consequently, using (9) to construct the classification sign function can sole the nonlinear classification problem of the transformer diagnose.

III. M-RBF-SVM IS APPLIED IN TRANSFORMER FAULT DIAGNOSIS

Common kernel functions includes: Linear kernel function, Polynomial kernel function, radial basis function (RBF) kernel function and Sigmoid kernel function[18]. These four basic kernels are as follows:

(1)Linear $K(x_i, x_j) = (x_i \cdot x_j{}')$

(2)Polynomial $K(x_i, x_j) = (r \cdot (x_i \cdot x_j{}') + 1)^d$

(3)RBF $K(x_i, x_j) = \exp(-\gamma \cdot \|x_i - x_j\|^2)$

(4)Sigmoid $K(x_i, x_j) = \tanh(r \cdot (x_i \cdot x_j{}') + r0)$

The construction and selection of kernel function are very important in the transformer fault diagnosis. The best kernel functions will obtain highest accuracy in the transformer diagnosis[19]. The RBF has powerful ability to deal with the nonlinear processing [20]. The RBF kernel function is as follow:

$$K(x, y) = \exp(-g * \|x - y\|^2) \tag{10}$$

From (10) the parameter "g" is fixed in the traditional RBF kernel function. Therefore, the contribution will be the same in the SVM model. But the relationship are different between characteristic gas and fault type, and the fixed parameter "g" will reduce the function of useless feature in the transformer fault diagnosis. Thus, the variable parameter "g_m" which presents vary "g" coefficient to improve the RBF kernel function is added. It is called Multi-RBF in (11).

$$K(x, x_i) = \exp(-g_m * \|x - x_i\|^2) \tag{11}$$

Here m is different labels among all kinds of samples.

Then it invert the final decision function into (12).

$$f(x) = \text{sign}(g(x)) = \sum_{i=1}^{l} y_i \alpha_i^* K(x, x_i) + b^* \\ = sign(\sum_{i=1}^{l} y_i \alpha_i^* e^{(-g_m * \|x - x_i\|^2)} + b^*) \tag{12}$$

A. The Way to Diagnosis

Thus, transformer fault nine types will be classified by the SVM method.In order to avoid being affected by the imbalance of the different types data, those data must be preprocessed. The preprocessed forms are as follows:

Pattern1: $x_s^* = \dfrac{x_s}{\max\limits_{i=1}(x_i)}$ Pattern 2: $x_s^* = \dfrac{x_s - x_{\min}}{x_{\max} - x_{\min}}$

Pattern 3: $x_s^* = \dfrac{x_s}{\sum\limits_{i=1}^{n}(x_i)}$

Here x_{max} and x_{min} represent the maximum value and the minimum value, respectively; x_i($i=1,2,3,4,5$) indicates the concentration of the raw gas data; x_s indicates the actual data and x_s* indicates preprocessed data. The different patterns will obtain different accuracy.

B. The Fault Diagnosis Process

The procedure of the transformer fault diagnosis based on M-RBF-SVM is as follows:

1) Obtain a certain number and select a clear conclusion sample data as the training sample set and test sample set.

2) Classify the sample which is provided in TABLE 1.

3) Set data preprocessing and normalization processing.

4) Select the appropriate kernel function and the parameter optimization.

5) Train model samples and test whether the accuracy requirements, if not, then return to 4) reselection.

6) Determine kernel function and parameter values.

7) Take the test data into the M-RBF-SVM classifier model and receive the diagnosis results.

8) The fault diagnosis flow chart is shown in Fig.2.

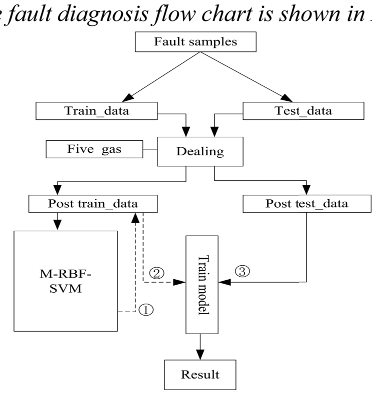

Fig.2 General flow chart of M-RBF-SVM fault diagnosis

IV. SIMULATION RESULT AND ANALYSIS

223 groups of transformer fault types collecting from different transformer substation are used to be examined actual transformer fault. The transformer fault data are divided into training-data (127samples) and testing-data (96samples). In this paper, the five kinds of gas are selected as input features.

The simulation results of above three preprocessed patterns indicate that the pattern3 has highest accuracy in TABLE II. Therefore, the patterns3 is chosen as coping with the fault data.

TABLE II. COMPARISON OF DIAGNOSIS RESULT BETWEEN DIFFERENT PATTERNS

Pattern	Testing-data	Accuracy_SVM
Pattern 0	96	34.375%
Pattern 1	96	89.5833%
Pattern 2	96	89.5833%
Pattern 3	96	94.7917%

Then, in order to illustrate the M-RBF-SVM diagnosis methods valid, this paper uses the same transformer fault data to predict the fault types with the TTR and NTR algorithms.

From the Fig.3 and Fig.4, the TTR algorithms can correctly determine the corresponding types for 85 test samples, the accuracy rate is 88.54%. And, using the NTR algorithms correctly determine the corresponding type for 88 test samples, the accuracy rate result can be achieved up to 91.67%.

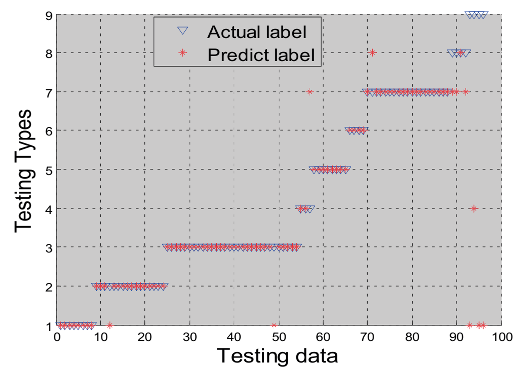

Fig.3 The results of the TTR simulation

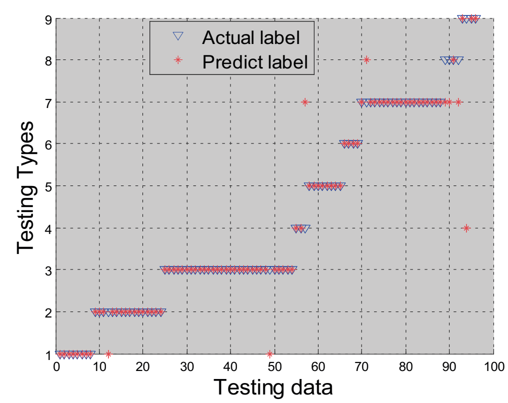

Fig.4 The results of the NTR simulation

From the Fig. 5 and Fig. 6, the TSVM algorithms can correctly determine the corresponding types for 86 test samples, and the accuracy rate result is 89.5833%. And, using the M-RBF-SVM algorithms correctly determine the corresponding type for 91 test samples, and the accuracy rate result can be achieved up to 94.79%.

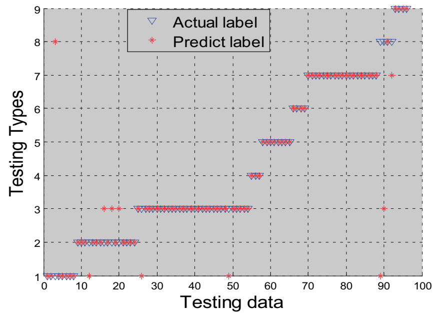

Fig.5 The results of the TSVM simulation

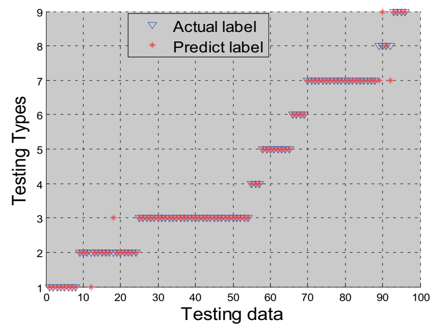

Fig.6 The results of the M-RBF-SVM simulation

Thus, comparing the Fig.5 and Fig.6, the M-RBF-SVM algorithms not only can solve the classification and predict the transformer fault types effectively, but also increases the accuracy. All the simulation results are put in Fig.7.

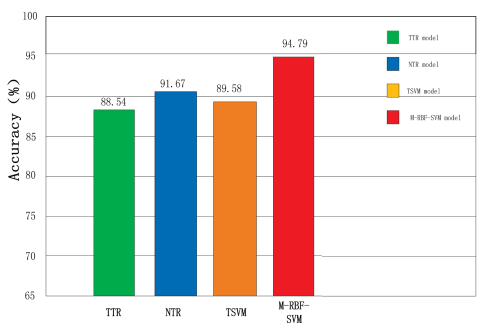

Fig.7 The comparison simulation experimental results

Two improved algorithms which increase the accuracy of fault diagnosis are proposed. One is the NTR algorithm which improves traditional three ratio code. The other is M-RBF-SVM algorithm which improves the traditional RBF kernel function. And, according to the comparing simulations above two kinds of algorithms, the results show that the M-RBF-SVM can improve the correct diagnosis higher rate and the accuracy is up to 94.79%.

* Corresponding author:
QU Li-ping
E-mail: 309637683@qq.com

REFERENCE

[1] A Review of Faults Detectable by Gas-in-Oil Analysis in Transformers [J]. IEEE Electrical Insulation Magazine.2002, 18(3):8-17.
[2] Zhu D H, Yan Z, Tan K X. Electrical equipment status inspection and fault diagnosis technology[M].Beijing: China Electric Power Press, 2009.
[3] Chen Q M, Tang W. Comparative study on three kinds of transformer fault diagnosis method [J].Power System Technology, 2006,10 (30):423-425.
[4] ZOU J, LU J, ZHOU X F. Four ratio method in the application of the transformer overheating fault judgment[J]. Transformer, 2011.48(10):66-67.
[5] Castro A R G, Miranda V. Knowledge discovery in neural networks with application to transformer failure diagnosis[J].IEEE Transaction on Power Systems, 2005,20(2):717-724.
[6] Vapnik V N. The nature of statistical learning theory[M]. New York: Springer,2000.
[7] The state economic and trade commission of the People's Republic of China.DL/T 722—2000 Guide dissolved gas in transformer oil, and judgment [S]. Beijing, China electric power press ,2001.
[8] Xiao Y C, Li C Y, Zhang Q , Han X. Design of Transformer Remote Fault Diagnosis System [J]. Control Engineering of China, 2014,06:982-985.
[9] Sun D G, Jiang L W. Improved the three ratio method in the application of transformer fault diagnosis[J]. Electric safety technology, 2014.4 (16):65-67.
[10] Ni Y P, Zhou J H.Li B H, Zou J H. Method of f ault diagnosis for power transfor- mer based on rough set theory[J]. Control and Decision, 2004, 08:943-946.
[11] Wang X C, Shi F ,Yu L , Li Y. Matlab Neural Network 43 Case Analysis[M]. Beijing: Beijing university of aeronautics and astronautics press, 2014:102-179.
[12] H.-J. Li,Z.-M. Zhao, X.-L. Yu.. Grey theory applied in non-sub sampled Contour let Transform. The institution of engineering and Technology 2012:264-272.
[13] Khmais Bacha, Seifeddine Souahlia. Power transformer fault diagnosis based on dissolved gas analysis by support vector machine. Electric Power Systems Research, 2012(83):73-79.
[14] Zhi L, Hu J Y, Cao S A. Prediction Technique for Transformer Oil Breakdown Voltage via Multi-parameter Correlation Based on Grey Theory and BP Neural Network.2010 International Conference on Information, Networking and Automation:179-183.
[15] Souahlia, Seifeddine. SVM-based decision for power transformers fault diagnosis using Rogers and Doernenburg ratios DGA.10th Internation Multi-Conference on Systems, Signals & Devices, 2013.
[16] Mehta, Amit Kumar,Sharma, R.N.,Chauhan, Sushil,Saho, Satyabrata.Transformer diagnostics under dissolved gas analysis using Support Vector Machine.2013 International Conference on Power, Energy and Control(ICPEC),2013.
[17] Maksim Lapin, Matthias Hein,Bernt Schiele. Learning using privileged information: SVM and weighted SVM[J].Neural Networks,2014,53: 95-108.
[18] M.Bigdeli, M.Vakilian, E.Rahimpour. Transformer winding faults classification based on transfer function analysis by support vector machine.The Institution of Engineering and technology, 2012:268-276.
[19] Eslam Pourbasheer, Siavash Riahi, Application of genetic algorithm-support vector machine (GA-SVM) for prediction of BK-channels activity[J]. European Journal of Medicinal Chemistry,2009,44(12): 5023-5028.
[20] Keskes.H, Braham, A .DAG SVM and pitch synchronous wavelet transform for induction motor diagnosis. Power Electronics, Machines and Drives, 7th IET International Conference on,2014.