

Fault Diagnosis for the Power Transformer Based on Multi-feature Fusion algorithm

Chenfei Liu^{1,a}, Haoyang Cui^{1,b,*}, Gaofang Li^{1,c}

¹College of Electronics and Information Engineering, Shanghai University of Electric Power, Shanghai, China

^a313455801@qq.com, ^bcuihy@shiep.edu.cn, ^clioneix_1128@sina.com

*corresponding author

Keywords: fault diagnosis; SVM; multi-feature; D-S evidence theory; transformer

Abstract: To address the low accuracy and low stability of a single algorithm for transformer fault diagnosis, this dissertation is based on multi feature fusion diagnosis algorithm by combing support vector machine (SVM) and D-S evidence theory, The way to construct the basic probability assignment(BPA) of evidence has been improved by calculating the correct recognition rate and misdiagnosis probability of the SVM classification results. Simulation results show that this method can obtain more reliable belief function of the evidence, and further improve the accuracy of multi-feature fusion fault diagnosis.

1. Introduction

With the development of artificial intelligence algorithm and the increasing demand of the safe and stable operation of power system, artificial intelligence algorithm such as decision tree and artificial neural network has been widely researched and applied in the field of electric power equipment condition monitoring. Many achievements has been made in the research and application of in transformer fault diagnosis^[1] by support vector machine with its unique advantages in solving the small sample, nonlinear and high dimensional pattern recognition problems and good generalization ability. However, the relation between the fault mechanism and the fault feature of the transformer is very complicated, and single diagnosis algorithm usually has some limitations.

Based on single SVM, there are some problems in fault diagnosis method such as low reliability and low accuracy of transformer fault diagnosis. Without considering the influence of each characteristic data on the equipment condition recognition is different, SVM maps data from sampling space to higher dimensional characteristic space by the kernel functions to divide the fault classification region. While the features of transformer generally reflect a variety of faults or a fault can be reflected by a variety of features, and here are circumstances that some characteristics mismatch some fault^[2]. Information fusion technology is a method combining and processing information of multi-feature to get a more reasonable decision. Many scholars have achieved fusion diagnosis through D-S evidence theory combined with SVM. Efficiently fusion diagnosis was carried out by obtaining BPA from the voting rates of all types of samples through “1-a-1” SVM^[3]. To solve the problem that samples in the fuzzy region are difficult to classified, the output distance of SVM was converted to the membership degree of the sample to each state, and multisensory information was fused by D-S evidence theory on this basis^[4].

First, this paper established models separately by five characteristic gasses of transformer. And then the BPA of evidence was calculated by the recognition rate modified by misdiagnosis probability parameter that we will introduce in the next section. Finally, the multi-feature information was fused by using D-S combining rule.

2. Fusion Diagnosis Method Based on Multi-feature

2.1. SVM Diagnosis Based on Single Feature

Data used in transformer fault diagnosis oil includes H_2 , CH_4 , C_2H_6 , C_2H_4 , C_2H_2 . We used each characteristic gases to train “1-a-1” SVM model for initial diagnosis. In this way, k types of samples for example, $k(k-1)/2$ sub-classifier should be trained at first. And diagnosis of each SVM model is all determined by sub-classifier vote. The decision function of each sub-classifier is as follows:

$$\hat{y} = \text{sgn}(\sum_{i=1}^N y_i \alpha_i^* \langle x_i \cdot x \rangle + b) \quad (1)$$

where α_i^* —lagrange multiplier; x , \hat{y} —features and label of test sample; x_i , y_i —feature and label of training sample.

2.2. Constructing Basic Probability Assignment

Because the output of SVM is +1 or -1, there is no way to construct the basic probability assignment function directly from the output. Therefore, we used sigmoid function to estimate the probability of pairing classes^[5]. Finally, the BPA function is determined by the posterior probability p_i and the recognition accuracy rate q_i of the training samples:

$$m(A) = p_i q_i \quad (2)$$

In the above equation, q_i is the result of single feature SVM. But due to the simplex information, the diagnosis based on single feature may cause the result a rather serious error. This will further lead to the unsatisfactory results of fusion diagnosis. Take a-SVM (multi-classification SVM based on feature a) for example, there is the possibility that the diagnostic accuracy to samples of label 1 is 100%, while a large number of samples of other labels are incorrectly diagnosed as label 1 by a-SVM. Under these circumstances, we will get a higher BPA for samples of label 1 by equation (2), which means we have more confidence than actual one in the output of a-SVM about samples of label 1.

Considering the influence of misdiagnosis on final decision, probability parameter $c(1 - r_i'/r)$ of misdiagnosis is introduced to amend recognition accuracy rate q_i . The improved equation is as follows:

$$m_i(A) = p_i q_i c(1 - r_i'/r) \quad (3)$$

Where r is total misdiagnosis number of a-SVM, r_i' is the number of samples, which are mis-diagnosis, as label i , and c is normalized parameter.

2.3. Fusion Diagnosis Based on D-S Evidence Theory

First, we established the frame of discernment $\Theta = \{A_1(\text{normal}), A_2(\text{overheat}), A_3(\text{high-energy breakdown}), A_4(\text{low-energy breakdown})\}$. The function $m: 2^\Theta \rightarrow [0, 1]$ is defined as basic probability assignment function which satisfies the following conditions: For empty set Φ , $m(\Phi) = 0$; for $\forall A \in 2^\Theta$, $\sum m(A) = 1$. Function $m(A)$, the BPA of A , represents the precise confidence in proposition A . The BPA m_1, m_2, m_3, m_4, m_5 of different evidence can be obtained by five different features, and the combination process of these BPA is as follows:

$$m(A) = k^{-1} \sum_{\cap A_i = A} \prod_{1 \leq i \leq 5} m_i(A_i), A \neq \Phi \quad (4)$$

where

$$k = 1 - \sum_{\cap A_i = \Phi} \prod_{1 \leq i \leq 5} m_i(A_i) \quad (5)$$

We defined the sum of the BPA of those evidence that support proposition A as belief function, that is $\text{Bel}(A) = \sum_{B \subseteq A} M(B)$. The whole fusion diagnostic model is shown in Figure 1.

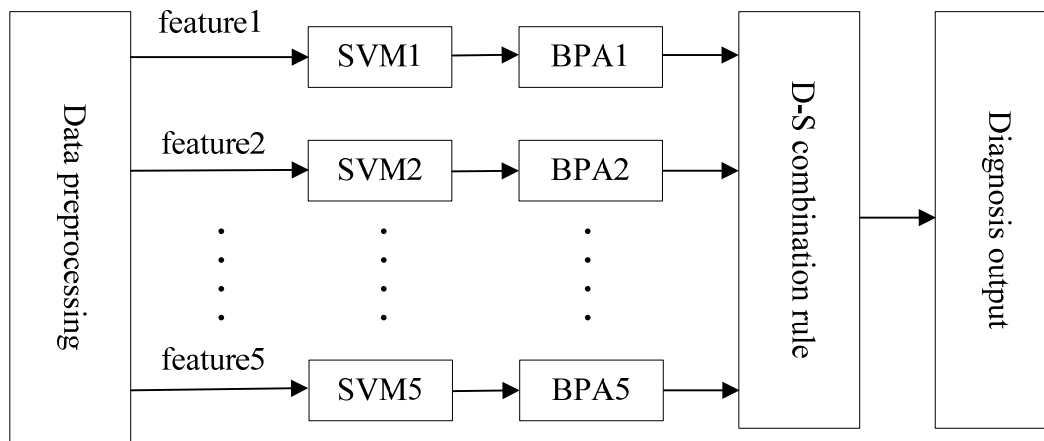


Figure 1 Fusion diagnosis model for transformer

3. Experimental Results and Analysis

This thesis takes the DGA datasets collected by literature [6]. The datasets consists of five characteristic gas contents, with a total sample size of 75. Forty-one cases of the datasets were randomly selected as training datasets, and the rest as testing datasets. Simulation results based on Matlab are shown in table 1 and table 2.

Table 1 diagnosis by different method

Fault Type	Accuracy for Test Datasets		
	SVM	D-S SVM	Proposed Method
High-energy breakdown	2/7	6/7	7/7
Low-energy breakdown	2/5	2/5	3/5
overheat	18/18	18/18	18/18
normal	3/4	3/4	3/4
total	73.5%	85.3%	91.1%

Table 2 diagnosis of partial samples by different method

Test samples No.	Fault Type	Diagnosis Result		
		SVM	D-S SVM	Proposed Method
2	high-energy breakdown	overheat	overheat	high-energy breakdown
3	high-energy breakdown	overheat	high-energy breakdown	high-energy breakdown
4	high-energy breakdown	overheat	high-energy breakdown	high-energy breakdown
6	high-energy breakdown	overheat	high-energy breakdown	high-energy breakdown
7	high-energy breakdown	overheat	high-energy breakdown	high-energy breakdown
11	low-energy breakdown	overheat	overheat	low-energy breakdown
12	low-energy breakdown	overheat	overheat	low-energy breakdown

From the table 1, the accuracy is greatly improved by multi-feature fusion diagnosis based on D-S evidence theory and SVM. On this basis, we improved the method to access the basic probability assignment function by introducing probability parameter of misdiagnosis. Compared with initial fusion diagnosis, it can enhance the reliability of diagnosis model furthermore.

Table 2 presents the diagnostic results that differ greatly in three methods. Because SVM directly maps multidimensional data to higher dimensional characteristic space by the kernel functions to divide the fault classification region. Obviously, this method is unable to consider the different influence of each feature on the fault diagnosis. So, due to the lack of thoroughly analyzing each features and a large number of overheating faults in training datasets, SVM diagnosis model will over fit overheating fault samples. As shown in table 2, many testing samples were misdiagnosed as overheating faults. Multi-feature fusion diagnosis based on D-S evidence may solve this problem to some extent. First the features of datasets are trained and recognized respectively, and then evidence theory is used to fuse those recognition results. Due to considering the different influence of features on each fault type, as shown in table 2, partial misdiagnosis results were revised by D-S and SVM.

As mentioned in section 2, we put forward a method to improve the acquisition approach of BPA. Take sample No.2 in table 2 as an example, we first got five recognition results based on five features. Sample No.2 was misdiagnosed as overheating by SVM model based on H_2 and C_2H_4 , while the recognition results are consistent with the actual fault type by others. As shown in table 2, we obtained a high degree of belief for overheating by equation (2), which the value is about 0.75, and these made the final result turn to overheating. After analyzing the recognition results of initial SVM model, we found that SVM trained on H_2 and C_2H_4 (H_2 -SVM and C_2H_4 -SVM) believed the fault type of the sample is overheating with high BPA, while results of SVM trained on other features are same as the fact. Further, recognition results shows that H_2 -SVM and C_2H_4 -SVM easily misdiagnose samples as overheating. That means when evidence given by H_2 -SVM or C_2H_4 -SVM indicated that transformer is overheating, misdiagnosis may be likely to occur. In light of this situation, this paper proposed a method as described in section 2 to adjust BPA. Thus the degree of belief with relatively high misdiagnosis rate was lowered. As shown in table 2, the diagnostic accuracy has been further improved.

4. Conclusion

Compared with the single algorithm, transformer fault diagnosis based on multi-feature fusion can make a full use of the information on different features sufficiently, and the diagnosis is more reliable. Based on D-S evidence theory and SVM, this paper proposes a method of constructing BPA for transformer fault diagnosis. First the probability parameter of misdiagnosis is introduced to amend recognition accuracy rate to get more accurate BPA. On this basis, we can get more reliable evidence by the combination rule of evidence theory. The simulation results show that, compared with the original method, the diagnostic accuracy of each fault type is improved, and the accuracy of comprehensive diagnosis is improved about 6%.

Acknowledgements

This work was supported by the NSFC (61107081, 11647023), Shanghai Local colleges and universities capacity building Program (14110500900, 15110500900), Shanghai Natural Science Foundation, China (17ZR1411500), and Intelligent Power Grid Center Knowledge Service Platform Achievement Cultivation Project (A-0010-16-029-001).

Reference

- [1] GUO Chuangxin, ZHU Chengzhi, ZHANG Lin, et al. A fault diagnosis method for power transformer based on multiclass multiple-kernel learning support vector machine[J]. Proceedings of the CSEE, 2010, 30(13):128-134.
- [2] WANG Ribin, SHE Caiqi, LIU Xindong, et al. Fault diagnosis of transformer based on D-S evidence theory[J]. Modern Electric Power, 2012, 29(2):6-10.
- [3] JIANG Wanlu, WU Shengqiang. Multi-data fusion fault diagnosis method based on SVM and evidence theory[J]. Chinese Journal of Scientific Instrument, 2010, 31(8):1738-1743.

- [4] LIU Mingliang, ZHEN Jianju, SUN Laijun, et al. Modification of SVM classification fuzzy area based on DS evidence theory[J]. Electric Power Automation Equipment,2012,32(3):71-75.
- [5] HE Dongjian, QIAO Yongliang, LI Pan, et al. Weed recognition based on SVM-DS multi-feature fusion[J]. Transactions of the Chinese Society for Agricultural Machinery,2013,44(2):182-187.
- [6] GanyunLv, Haozhong Chen, Haibao Zhang, et al. Fault diagnosis of power transformer based on multi-layer SVM classifier [J]. Electric Power Systems Research,2005,75:9-15.