

Research Article

Modeling Analysis of Power Transformer Fault Diagnosis Based on Improved Relevance Vector Machine

Lutao Liu¹ and Zujun Ding²

¹ College of Information and Telecommunication, Harbin Engineering University, Harbin 150001, China

² Department of Electronic and Electrical Engineering, Huaiyin Institute of Technology, Huai'an 223001, China

Correspondence should be addressed to Zujun Ding; dzj_king@263.net

Received 13 September 2013; Accepted 15 October 2013

Academic Editor: Zhiguang Feng

Copyright © 2013 L. Liu and Z. Ding. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A new method of transformer fault diagnosis based on relevance vector machine (RVM) is proposed. Bayesian estimation is applied to support vector machine (SVM) in the novel algorithm, which made fault diagnosis system work more effectively. In the paper, the analysis model is presented that the solutions of RVM have the feature of sparsity and RVM can obtain global solutions under finite samples. The process of transformer fault diagnosis for four working statuses is given in experiments and simulations. The results validated that this method has obvious advantages of diagnosis time and accuracy compared with backpropagation (BP) neural networks and general SVM methods.

1. Introduction

Power transformer is one of the key equipment for electric power transmission and distribution, which is a widely distributed, complex, and expensive equipment in the power system. And its safety situation plays a great effect on stability and security level of power system. Therefore, it is of great realistic significance to study the fault diagnosis technology and raise the level of maintenance of power transformers. Dissolved gas analysis (DGA) [1, 2], which provides operation information by detection of certain gases generated in an oil-filled transformer, is the method widely adopted by the utilities. The concentrations of the dissolved gases, their generation rates, ratio of respective gases, or total combustible gases in the oil are the attributes used in the DGA method to interpret a malfunction. Accuracy and reliability of the DGA data are influenced by many factors, so the DGA three-ratio method and its improved ratio methods by classical IEC [3] are considered that limitations are obvious, such that the ratio crosses the coding boundary or codes change sharply [3, 4]. And it cannot offer completely objective, accurate diagnosis for all the faults, such as the low diagnostic accuracy for overheating fault. With the development of artificial intelligence, some solutions [5, 6] based on neural networks are applied to power transformer fault diagnosis process. Though some

improved results in faults diagnosis can be obtained, this kind of methods has some inherent disadvantages in application, such as local optimization and danger of overfitting [7, 8]. Support vector machine (SVM) overcomes the drawbacks of neural networks in aspects of convergence and real-time application [9, 10]. In recent years, SVM methods have been applied to fault diagnosis and identification of power transformers [11–13]. However, the input characteristic information during the process of fault diagnosis is large in order that SVM methods take heavy computation to approximate the optimal solution and spend much time in parameter searching [14, 15]. So they are difficult in real-time monitoring and power transformer fault diagnosis.

According to practical situations that there are many uncertainties in the running of transformers and there are finite samples during fault diagnosis process, relevance vector machine is introduced to fault detection and identification of power transformers. RVM, which applies Bayesian estimation model to SVM algorithms [16, 17], can decrease hyperparameters in diagnosis algorithms and bring good adaptive ability in kernel function and model parameters chosen. RVM is a machine learning methodology based on sparse Bayesian learning theory proposed by Tipping in 2000. It absorbs the advantages of wonderful generalization and precision from SVM; meanwhile, it overcomes some inherent

limitations of SVM and possesses advantageous features such as high degree of sparsity, fewer kernel functions, and low computation load. Particularly, the RVM classification can provide posterior probabilities for class memberships, which is suitable to analyse indeterminate problem in power transformer fault diagnosis. By dissolved gas analysis of power transformer oil, fault diagnosis models are established based on RVM in the paper. Through fault diagnosis for several kinds of power transformers, the proposed algorithm is validated to be better than BP neural networks and general SVM methods in the aspects of finite sample size, diagnosing speed, and diagnosing accuracy. It provides a new way of thinking for resolving the transformer fault diagnosis problem in real time. In the future, the research of transformer diagnosis faults based on RVM is real-time implementation for power transformer systems. Recently, data-driven fault diagnosis methods become more popular in many industry sectors [18–21]. So a combination of the proposed method and data-driven methods is the future work to realize large-scale real-time implementation in fault diagnosis of power transformers.

The paper is organized as follows. In the next section, we introduce RVM model of fault diagnosis. Diagnosis process for power transformers is extended in Section 3, which includes fault information preprocessing, diagnosis model training, and fault identification for power transformers. In Section 4, we give the experiment results in application and offer some benchmark comparison with least squares support vector machine (LS-SVM) and BP neural networks before summarising in Section 5.

2. Analysis of RVM Model for Faults Diagnosis

RVM based on Bayesian learning framework introduces Bayesian theory of Gaussian process in SVM. The results of inference are shown in the form of probability density, and it can be written as

$$y(x) = \sum_{j=1}^n h_j \phi_j(x) + h_0, \quad (1)$$

where $\phi_j(x)$ is a nonlinear kernel function and h_j is the weight of the model. For avoiding overfitting of traditional SVM, based on Bayesian framework [22], maximum likelihood method is applied to training of the model weight. RVM defines the prior probability distribution as follows:

$$p(h_j | a_j) = \left(\frac{a_j}{2\pi}\right)^{1/2} \exp\left(-\frac{1}{2}a_j h_j^2\right), \quad (2)$$

where a_j is hyperparameter of the prior distribution. Assuming that input training sample set is $\{x_j, k_j\}_{j=1}^N$, target value k_j is independent, and input noise is Gaussian distribution with variance σ^2 , the maximum likelihood function of training sampling set correspondingly can be expressed as

$$p(\mathbf{k} | \mathbf{h}, \sigma^2) = (2\pi\sigma^2)^{-N/2} \exp\left(-\frac{1}{2\sigma^2}\|\mathbf{k} - \Phi\mathbf{h}\|^2\right), \quad (3)$$

where $\mathbf{k} = [k_1, \dots, k_N]^T$, $\mathbf{h} = [h_1, \dots, h_N]^T$, and $\Phi = [\Phi_1(x_j), \dots, \Phi_n(x_j)]^T$ is the response of all kernel functions with input x_j . The posterior distribution of the weight value conditioned on the data is given by combining the likelihood function (3) and the prior probability distribution (2) within Bayesian rule:

$$p(\mathbf{h} | \mathbf{k}, \mathbf{a}, \sigma^2) = \frac{p(\mathbf{k} | \mathbf{h}, \sigma^2) p(\mathbf{h} | \mathbf{a})}{p(\mathbf{k} | \mathbf{a}, \sigma^2)}, \quad (4)$$

where $\mathbf{a} = [a_1, \dots, a_n]^T$ denotes the vector of hyperparameters. And the posterior probability distribution over the weights is multivariate Gaussian distribution:

$$p(\mathbf{h} | \mathbf{k}, \mathbf{a}, \sigma^2) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad (5)$$

where $\boldsymbol{\Sigma} = (\sigma^{-2}\Phi^T\Phi + \mathbf{A})^{-1}$ is the posterior covariance of the multivariate Gaussian distribution, \mathbf{A} is defined as diagonal matrix with the elements (a_1, \dots, a_n) , and $\boldsymbol{\mu} = \sigma^{-2}\boldsymbol{\Sigma}\Phi^T$ is the mean of the multivariate Gaussian distribution. Integrating the distribution with respect to weight value, the likelihood distribution of training value can be given by

$$p(h_j | k_j, a_j, \sigma^2) = \int p(k_j | h_j, \sigma^2) p(h_j | a_j) dh_j. \quad (6)$$

Marginalize (6); marginal likelihood distribution of hyperparameters can be written as

$$p(\mathbf{k} | \mathbf{a}, \sigma^2) \sim N(\mathbf{0}, \mathbf{C}), \quad (7)$$

where $\mathbf{C} = \sigma^2\mathbf{I} + \Phi\mathbf{A}^{-1}\Phi^T$. Estimated values of weights in RVM method are given by the means of the posterior probability distribution. Maximum posterior estimation of weights depended on hyperparameter \mathbf{a} and noise variance σ^2 , and estimated values $\bar{\mathbf{a}}$ and $\bar{\sigma}^2$ can be computed by the maximum marginal likelihood distribution. Uncertainty of diagnosis model prediction can be represented by the uncertainty of optimum weight, which is shown in the posterior distribution. If input \mathbf{x}^* is given, the corresponding probability distribution of output can be expressed as

$$\begin{aligned} p(k_j | x_j^*, \bar{a}_j, \bar{\sigma}^2) \\ = \int p(k_j | x_j^*, h_j, \bar{\sigma}^2) p(h_j | k_j, \bar{a}_j, \bar{\sigma}^2) dk_j, \end{aligned} \quad (8)$$

and this probability distribution is Gaussian distribution:

$$p(\mathbf{k}^* | \mathbf{x}^*, \bar{\mathbf{a}}, \bar{\sigma}^2) \sim N(\mathbf{y}^*, \sigma^{2*}). \quad (9)$$

In (9), the predict mean and variance are, respectively,

$$\begin{aligned} \mathbf{y}^* &= \boldsymbol{\mu}^T \Phi(x^*), \\ \sigma^{2*} &= \bar{\sigma}^2 + \Phi^T(x^*) \boldsymbol{\Sigma} \Phi(x^*). \end{aligned} \quad (10)$$

Using Bayesian theory to compute parameters of fault diagnosis model based on RVM, diagnosis parameters choice can be optimized and application range of the faults diagnosis can be extended.

3. Diagnosis Process of the Transformer Based on RVM

Internal faults of transformers are classified into heat fault and electrical fault. In this paper, RVM method was applied to diagnose four types of working situations in power transformers, which include high energy discharge, low energy discharge, overheating, and normal condition operation. After the following three steps: characteristic information extraction by DGA, diagnosis model training, and faults recognition, the fault of transformer can be identified accurately.

3.1. Fault Information Preprocessing. Through DGA, concentrations of dissolved gases H_2 (hydrogen), CH_4 (methane), C_2H_6 (ethane), C_2H_4 (ethylene), and C_2H_2 (acetylene) could be obtained. The information of concentration of the five dissolved gases can be used to verify the four fault types stated above. So the ratio of each gas can be computed firstly by the following method:

$$x^k = \frac{c^k}{\max_{i=1}^5 (c^i)}, \quad (11)$$

where $k = 1, 2, \dots, 5$ is the index of each gas. c^k is the concentration of gas in the unit ppm. Then, we can obtain special character information by computing the maximum concentration of the gas as $x^6 = \log_{10}(\max_{k=1}^5 (x^k))$, which is used for crosswise comparisons for all groups of test data. The six characteristics of data construct a six-dimension vector $\mathbf{x} = [x^1, x^2, x^3, x^4, x^5]^T$. This fault character information vector can be used to determine the working condition of power transformer: high energy discharge, low energy discharge, overheating, or normal condition.

3.2. Diagnosis Model Training. After faults feature extraction, it is necessary for fault diagnosis that RVM classifiers are trained to identify four possible states: high energy discharge, low energy discharge, overheating, and normal condition. So three diagnosis RVMs should have been set up and trained by all training samples. Each RVM corresponds to one of the fault states in power transformers. Training samples are sent into RVMs for training and identification. The output is assigned to one if the data of identification result is larger than zero. On the contrary, the output is zero if the data is less than zero. Firstly, the RVM1 is trained to overheating identification by the training sample set. Output of RVM1 for overheating samples is set to one, and the output for other fault states remains zero. Secondly, the RVM2 is trained to separate high energy discharge from low energy discharge. The output of RVM2 is set to one when the input is a high energy discharge sample; otherwise output of RVM2 is zero. Finally, RVM3 for low energy discharge is trained. In the same way, the output for low energy discharge samples is set to one; otherwise output is zero. Thus, three RVMs could be obtained to realize binary classification to all the samples for faults diagnosis. In the process, the kernel functions implicated in the three RVMs are all Gaussian radial basis functions. The flowchart of diagnosis model training was shown in Figure 1.

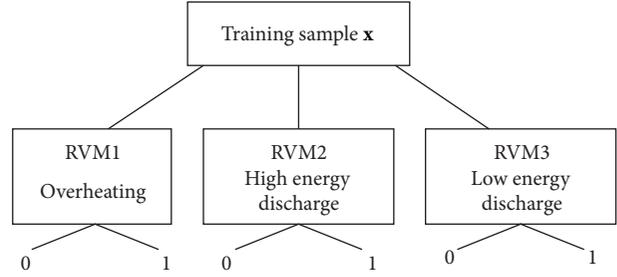


FIGURE 1: The flowchart of diagnosis model training.

3.3. Fault Identification of Power Transformers. After features extraction of DGA in Section 3.1, the test data with character information are considered as identification samples, which would be input to the RVMs established for faults diagnosis. The three RVMs are used to recognize the fault types. If the output of RVM1 is larger than zero, we could consider that the fault of power transformer is overheating. If the output of RVM1 is less than zero, the sample continues to be sent to RVM2 and RVM3 for recognizing the faults of high energy discharge and low energy discharge in the same way. If the outputs of the three RVMs are all less than zero, the power transformer would be judged in normal working condition. Thus, RVM classifier identifies the four states of transformers after three times of identification process.

4. Experimental Results and Analysis of Faults Diagnosis of Power Transformers

Sixty groups of historical data about gas concentrations in certain power transformer were collected as the training set. The data corresponding to actual fault conditions were known. In the training dataset, there were four working states of the power transformer, which were overheating, high energy discharge, low energy discharge, and normal operation. Each state had 15 samples. Three RVMs (RVM1, RVM2, and RVM3) were constructed by the samples related to each state in the training dataset, separately. And other 10 groups of samples of the running transformer were considered as the test data set; the test samples included 3 overheating samples, 2 high power discharge samples, 3 low power discharge samples, and 2 normal operation samples; see Table 1.

The 10 groups of test data were sent to the three RVMs for faults diagnosis after the preprocessing mentioned above. The output of test data through RVM1 was $[4.4361, 0.9235, 3.4215, -8.4312, -7.6557, -9.5236, -4.1404, -2.9146, -6.3237, -8.3624]$; then, we could modify the output result of RVM1 as $[1, 1, 1, 0, 0, 0, 0, 0, 0]$, so the first three samples in Table 1 were adjudged as overheating faults by RVM1. In the same way, the test samples were sent to RVM2 for the identification of high energy discharge fault. The output of RVM2 was $[-3.7431, -17.2762, -4.0026, 12.7902, 9.2563, -12.5385, -3.0571, -6.7612, -37.6247, -26.6382]$. So the result of RVM2 was represented as $[0, 0, 0, 1, 1, 0, 0, 0, 0]$; samples numbers 4 and 5 were recognized as high energy discharge fault. Finally, the test samples were sent to RVM3, and the outputs were $[-10.9261, -0.7481, -7.6892, -3.6297, -2.6589, 1.2792,$

TABLE 1: Sample group of dissolved gases.

No.	Fault type	Data of gas content (unit: ppm)				
		H ₂	CH ₄	C ₂ H ₆	C ₂ H ₄	C ₂ H ₂
1	Overheating	74	523	142	1200	6
2	Overheating	220	382	82	1019	17
3	Overheating	116	352	90	465	5
4	High energy discharge	295	51	12	117	124
5	High energy discharge	449	83	12	104	176
6	Low energy discharge	566	92	35	46	2
7	Low energy discharge	152	54	32	20	3
8	Low energy discharge	172	78	36	28	6
9	No fault	5	3	2	6	6
10	No fault	6	7	6	13	6

TABLE 2: Diagnosis comparison of the algorithms.

Diagnosis model	Target error	Training time (s)	Test times (s)	Accuracy diagnosis
BP neutral net	0.01	116.47	0.25	93%
LS-SVM	0.01	1.25	0.107	96%
Proposed RVM	0.01	0.714	0.05	100%

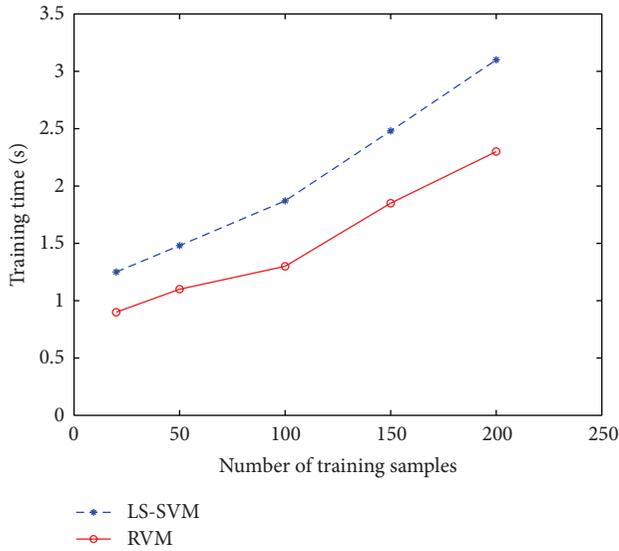


FIGURE 2: Training time via different numbers of training samples.

3.4769, 4.2915, -6.4163, -5.4794]; the result through RVM3 was written as [0, 0, 0, 0, 0, 1, 1, 1, 0, 0]; we could obtain the decision that low energy discharge happened in samples of numbers 6, 7, and 8. From the whole process of fault diagnosis, we found that the output results of three RVMs were all zeros for the samples numbers 9 and 10, which means that the transformer was in normal condition. From the results of the fault diagnosis by the proposed RVM method in experiments, we concluded that it had the ability for fault detection and identification of power transformers and the identification rate was high.

For evaluation of the proposed fault diagnosis model, we compared three learning models: the proposed RVM, BP neutral network and LS-SVM. There were 6 six input nodes,

44 hidden nodes, and 4 output nodes in BP neutral network, in which target error is 0.01. For LS-SVM, target error is 0.01 with the Gaussian radial basis function as the kernel function. Test parameters included training time, test time, and diagnosis accuracy of three methods in the condition of the same training samples. The test samples were 50 groups of historical data samples of the power transformer. The performance of testing was shown in Table 2. It is shown in Table 2 that the faults diagnosis models of LS-SVM and proposed RVM algorithms are superior to those of BP neutral network in the aspects of training time, test time, and diagnosis accuracy under the same target error. Furthermore, because of application of Bayesian probability model in the proposed RVM algorithm, the number of hyperparameters decreases. So test time and diagnosis accuracy of the proposed algorithm are better compared to those of conventional LS-SVM algorithm. It is more suitable to practical application. Next, the performance of training time, diagnose accuracy, and test time of LS-SVM and the proposed RVM are illuminated. Figures 2 and 3 show the performance of training time and diagnose accuracy in different numbers of training samples. The range of training samples number is from 25 to 200. From Figure 2, we can find that training time obviously become, longer for both of fault diagnosis models with the training sample number increasing. However, training time of the proposed optimum RVM model is much less than that of LS-SVM. In Figure 3, the testing diagnosis accuracy reaches 100% for the proposed method when the number of training samples is larger than 25. The proposed method is also superior to LS-SVM in test time see (Figure 4), in which the training sample number is 200. All the performance described above proves that the proposed RVM algorithm has advantages in the aspects of global optimization and sparsity of the solution. It has a wider range of adaptability than that of conventional SVM model.

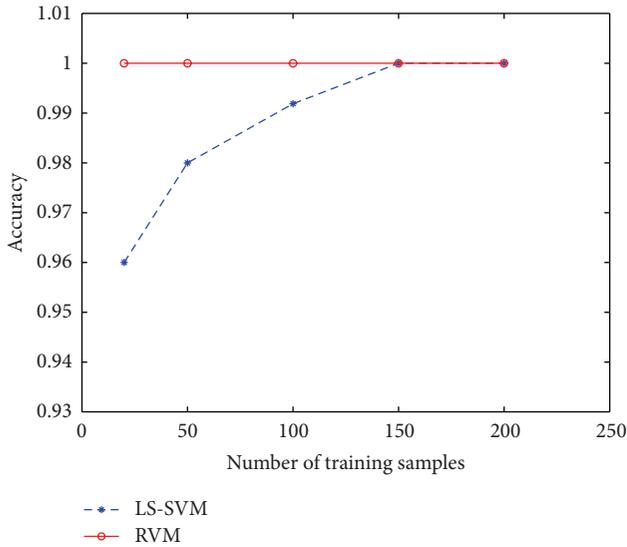


FIGURE 3: Diagnosis accuracy via different numbers of training samples.

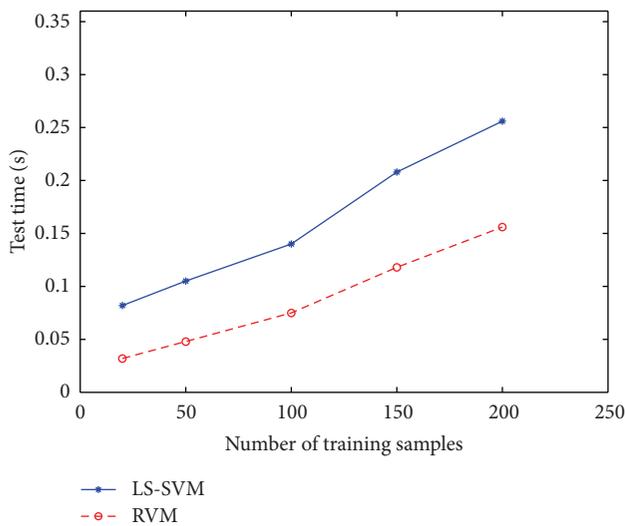


FIGURE 4: Test time via different numbers of test samples.

From the results of the faults diagnosis model by the RVM algorithm in the experiments, we can conclude that the proposed method can output the probability of working states of power transformers, which supply more useful information for transformer overhaul. The method is superior to conventional diagnosis methods in diagnosis accuracy, response time, and sample number.

5. Conclusion

The novel fault diagnosis model based on RVM for power transformers is proposed in the paper. Bayesian learning framework applied in the model makes hyperparameters decrease and improves the sparsity of the solutions. It overcomes the limitations of some traditional algorithms and enhances applicability of fault diagnosis. The proposed

method uses three RVMs to identify the faults of power transforms and brings about good results. The probability statistics of characteristics of the diagnosis model provide the advantages of small sampling, high sparsity, and low computation load. Compared to BP neural network and LS-SVM, the test time of fault diagnosis is shorter and accuracy of fault diagnosis is much higher. The results of experiments and simulations validate that this diagnosis method is suitable for real-time surveillance and identification of power transformers fault. The study of this issue is quite significant for transformers fault diagnosis application. The future work is to realize real-time implementation in power transformer systems by the proposed method.

Acknowledgments

This research has been supported by the National Natural Science Foundation of China (no. 61201410) and the Natural Science Fund of Jiangsu Province (11KJB470002). The authors are grateful to their colleagues from College of Information and Telecommunication in Harbin Engineering University for providing helpful comments.

References

- [1] R. Zhang, R. Guo, H. Li, and Z. Yan, "Feature selection of DGA data based on transformer fault classification," *High Voltage Engineering*, vol. 4, no. 31, pp. 32–33, 2005.
- [2] S. Singh and M. Bandyopadhyay, "Dissolved gas analysis technique for incipient fault diagnosis in power transformers: a bibliographic survey," *IEEE Electrical Insulation Magazine*, vol. 26, no. 6, pp. 41–46, 2010.
- [3] M. Duval and A. dePablo, "Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases," *IEEE Electrical Insulation Magazine*, vol. 17, no. 2, pp. 31–41, 2001.
- [4] "IEEE guide for the interpretation of gases generated in oil-immersed transformers," IEEE Standard C57.104-2008, 2009.
- [5] M. Hui, C. Ekanayake, and T. K. Saha, "Power transformer fault diagnosis under measurement originated uncertainties," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 19, no. 6, 1982.
- [6] A. R. Hooshmand, M. Parastegari, and Z. Forghani, "Adaptive neuro-fuzzy inference system approach for simultaneous diagnosis of the type and location of faults in power transformers," *IEEE Electrical Insulation Magazine*, vol. 28, no. 5, pp. 32–42, 2012.
- [7] B. Li, G. Godi, and M.-Y. Chow, "Detection of common motor bearing faults using frequency domain vibration signals and a neural network based approach," in *Proceedings of the American Control Conference*, pp. 2032–2036, June 1998.
- [8] S. Yin, *Data-Driven Design of Fault Diagnosis Systems*, Fortschritt-Berichte, Dsseldorf, Germany, 2012.
- [9] X. G. Zhang, "Introduction to statistical learning theory and support vector machines," *Acta Automatica Sinica*, vol. 26, no. 1, pp. 32–42, 2000.
- [10] S. Souahlia, K. Bacha, and A. Chaari, "SVM-based decision for power transformers fault diagnosis using Rogers and Doernenburg ratios DGA," in *Proceedings of the 10th IEEE International Multi-Conference on Systems, Signals & Devices (SSD '13)*, vol. 1, pp. 1–6, Hammamet, Tunisia, March 2013.

- [11] V. Miranda, A. R. G. Castro, and S. Lima, "Diagnosing Faults in power transformers with autoassociative neural networks and mean shift," *IEEE Transactions on Power Delivery*, vol. 27, no. 3, pp. 1350–1357, 2012.
- [12] D.-H. Liu, J.-P. Bian, and X.-Y. Sun, "The study of fault diagnosis model of DGA for oil-immersed transformer based on fuzzy means Kernel clustering and SVM multi-class object simplified structure," in *Proceedings of the 7th International Conference on Machine Learning and Cybernetics (ICMLC '08)*, pp. 1505–1509, July 2008.
- [13] L. V. Ganyun, C. Haozhong, Z. Haibao, and D. Lixin, "Fault diagnosis of power transformer based on multi-layer SVM classifier," *Electric Power Systems Research*, vol. 74, no. 1, pp. 1–7, 2005.
- [14] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, NY, USA, 1995.
- [15] S. Wen-gang, "Support vector machines based approach for fault diagnosis of valves in reciprocating pumps," *Journal of Mechanical Strength*, vol. 24, no. 3, pp. 362–364, 2002.
- [16] M. E. Tipping, "Sparse bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, no. 3, pp. 211–244, 2001.
- [17] M. E. Tipping, *The Relevance Vector Machine*, vol. 12 of *Advances in Neural Information Proceeding Systems*, MIT Press, Cambridge, Mass, USA, 2000.
- [18] S. Yin, X. Yang, and H. R. Karimi, "Data-driven adaptive observer for fault diagnosis," *Mathematical Problems in Engineering*, vol. 2012, Article ID 832836, 21 pages, 2012.
- [19] S. Yin, S. X. Ding, A. H. A. Sari, and H. Hao, "Data-driven monitoring for stochastic systems and its application on batch process," *International Journal of Systems Science*, vol. 44, no. 7, pp. 1366–1376, 2013.
- [20] S. Yin, S. Ding, A. Haghani, H. Hao, and P. Zhang, "A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process," *Journal of Process Control*, vol. 22, no. 9, pp. 1567–1581, 2012.
- [21] S. Yin, S. Ding, and H. Luo, "Real-time implementation of fault tolerant control system with performance optimization," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 5, pp. 2402–2411, 2013.
- [22] D. J. MacKay, "Bayesian interpolation," *Neural Computation*, vol. 4, no. 3, pp. 415–447, 1992.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

