Fault Diagnosis Based on Dynamic SVM *

Hongpeng Meng, Haiyan Xu, Qingyan Tan

Abstract—**General support vector machine (SVM) always uses limited amount of sampling data to train models, and then obtains relatively fixed optimal hyperplane to conduct classification, which could seriously influence the judgment of dynamic system failures on the contrary. This paper proposes an improved dynamic SVM based on data flow for fault diagnosis. With regard to the testing data sampled from complex system, the proposed dynamic SVM method firstly generates hyperplane with small scale of initial training data, and then predicts the probable state using initial optimal hyperplane. According to the fault type and its spatial distribution of testing data, the training data varies along with the monitoring process. In a word, the proposed method updates the training data in terms of time sequence, continually achieves the present optimal hyperplane, and diagnoses the failure approximately in real time. Simulation shows this method can decrease the scale of training data, and meantime improves the accuracy.**

I. Introduction

With rapid development of technology, the structure of industrial equipment gradually trends to be large-sized, complicated and of high accuracy, which brings more and more kinds of equipment failures. Meanwhile, it brings in quite complex characteristic information for fault diagnosis, directly leading to the high cost of time-economy for analysis and maintenance, especially for some new equipment. Therefore, the intelligent fault diagnosis based on knowledge or data-driven[1,2] has been rapidly developed and so far, these research has achieved amazing achievements.

Many intelligent algorithms such as neural network, genetic algorithm, wavelet analysis and support vector machine, have been deeply researched or applied in project cases[3~6]. Shengfeng Cheng[7] improved the Particle Swarm Optimization (PSO) algorithm to expedite the training speed of wavelet neural network and result showed it was well-performed in power transformer fault diagnosis. Kaiqin Li[8] established the aero-engine fault diagnosis expert system by means of improved genetic algorithm, and improved the recognition accuracy of faults. However, there exists two problems for fault diagnosis and detection of some equipment. One is that for the majority of machine-learning methods, the more sampling data to learn, the high accuracy will be, the other one is that the sampling data of normal state can be obtained easily, while the failure information is difficult to achieve and sometimes the cost will be rather hard to bear[9,10].

Considering this condition, the machine-learning method SVM, which based on the statistic theory, shows its great superiority than other methods for its advantages in dealing with the small sample, high dimensional and nonlinear problems. Since the appearance of SVM in 1963, it has achieved great development in theory and applications. Shaolei Zhou[11] proposed an algorithm for parameter selection based on data maximum variance-joint criterion and constituted the optimal Radial Basis Function (RBF) combining with PSO; Huimin Zhao[12] put forward a combined intelligent algorithm for fault diagnosis, using the genetic algorithm and PSO to find the best parameter for SVMs, and the model analysis showed that both ways have good effect for failure forecasting. Nevertheless, for any equipment especially for fast consumables, its working condition and state will change rapidly along with the working time, and then the parameters between the initial training data and the present sampling data may differ a lot in spatial distribution. Thus, the training model and the optimal hyperplanes obtained cannot truly reflect the actual state of equipment, which could express the false alert and mistake the failure type. Some methods dealing with the dynamic process always depend on the accurate mathematical model, which is not easy to build for all.

Therefore, intelligent fault diagnosis on dynamic system has been studied. Qinghua Wang[13] built a fault diagnosis expert system for a large, dynamic system, power unit; while its knowledge base and fault rules need to be improved. Shan He[14] studied the artificial neural network in generator fault diagnosis, while it's not a dynamic process. Donghua Zhou[15] introduced the fault diagnosis techniques of dynamic system, presented the current situation of it. Chuanyang Du[16] set up a dynamic SVM-Markov chain model by combing SVM and Markov chain for dam deformation, and results indicated it improved precision and generalization ability. However, when the sampling data accumulate more and more, the quadratic optimization scale of SVM model will become very large, which will seriously weaken the efficiency of fault diagnosis [15].

Considering the advantages of SVM method and in order to avoid the complexity of quadratic optimization scale, this paper presents the dynamic support vector machine based on data flow, in term of time sequence. By means of adjusting the training data set along with the time sequence and using PSO algorithm to find the best parameters for training, the presented dynamic SVM can better monitor the state of some complex system and precisely warn the fault type once failure

* Research supported by National Natural Science Foundation under grant #71471087; Postgraduate Innovative Base (Laboratory) for Opening Foundations under # kfjj20150323 and #kfjj20150320.

Hongpeng Meng, is with the College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, 201106 China(e-mail: hpmeng2013@126.com).

Haiyan Xu, is with College of Economic and Management, Nanjing University of Aeronautics and Astronautics, Nanjing, 211106 China. (e-mail: xuhaiyan@nuaa.edu.cn).

Qingyan Tan, is with College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, 211106 China. (e-mail: 1552063038@qq.com).

978-1-4673-8318-9/16/$31.00©2016 IEEE

occurs.

The rest of this paper is organized as follows: Section 2 firstly presents the basic SVM theory for classification, then dynamic SVM methodology for fault diagnosis is given in the following. Experiments validation and application are shown in Section 3. Section 4 is the conclusions.

II. SVM AND DYNAMIC SVM

A. Optimal Hyperplane of Linear SVM

SVM is a novel machine learning method based on the statistical theory, which contains the VC dimensional theory and the structural risk minimization principle of statistics. On account of the statistics established on the solid mathematics, the SVM can find a balance between the limited sampling information of complex model and learning ability, so as to achieve the best generalization. Different from others, SVM transforms the learning and training process into a quadratic programming problem, effectively avoiding the dimensional disaster and locally optimal solution. Therefore, the SVM proposed by Vlasimir Vapnik[18] has been applied in many fields, intelligent fault diagnosis included.

At the beginning, SVM was introduced for the optimal classification in linear condition. For the binary classification method[18], assume that a training set of n samples: $\{x, y \mid i = 1, 2, \cdots$, $x_i \in R^n$, $y_i \in \{-1, 1\}$; if this training set can be separated by a hyperplane H without error and the margin is the maximum, then this hyperplane is called optimal hyperplane, as shown in Fig.1.

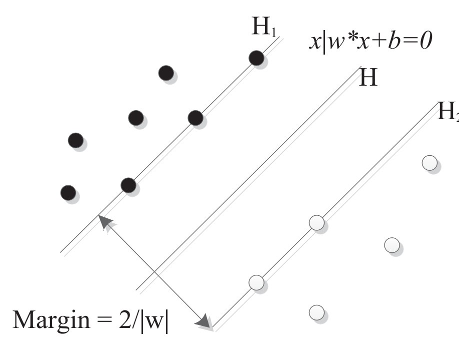

Figure 1. Diagram of SVM classification in linear condition

In Fig.1, H_1 and H_2 are canonical hyperplanes, the distance between them is margin: $2/\|\omega\|$. To achieve a good classification effect, the object turns to maximize the margin; then the aim of SVM is to construct the optimal hyperplane as follows:

$$\min_{\omega,b} 0.5\|\omega\|^2 = \min_{\omega,b} 0.5\omega^T\omega \tag{1}$$

subject to. $y_i = [\omega \cdot x_i + b] - 1 \geq 0$, $i = 1, 2, 3 \cdots$.

That's a convex quadratic optimization problem, and by means of lagrange multipliers this problem can be transformed into a dual form as follows:

$$\max_{\alpha} \sum_{i=1}^{n} \alpha_i - 0.5 \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \tag{2}$$

subject to. $\sum_{i=1}^{n} y_i \alpha_i = 0$, $\alpha_i \geq 0$, $i = 1, 2, 3 \cdots$, where α_i is the lagrange Multipliers.

According to equation (2), the optimal solution α_i^* is gained. Then the optimal hyperplane can be obtained as:

$$f(x) = \text{sgn}\left[\sum_{i=1}^{n} \omega^* \cdot x + b^*\right] \tag{3}$$

where $\omega^* = \sum_{i=1}^{n} \alpha_i^* y_i x_i$.

B. Optimal Hyperplane of Nonlinear SVM

Linear SVM mainly deals with the linear problem, while most of the problems in real are nonlinearity. Therefore, nonlinear SVM is put forward.

Based on the nonlinear mapping algorithm in kernel feature space, input vectors $x_i \in X$ are mapping into a high dimensional space Z through a predetermined nonlinear mapping[19], i.e. $\varphi : R^N \to Z, x \to \varphi(x)$; then construct optimal hyperplane in this high-dimensional space, as Fig.2 shows below.

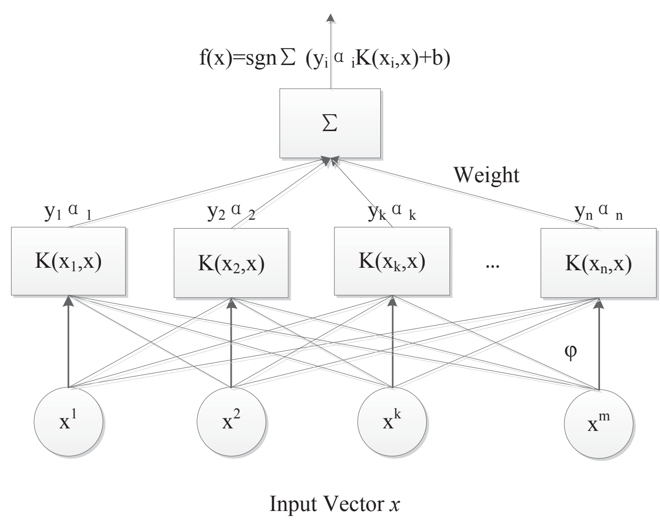

Figure 2. Topological structure diagram of SVM for nonlinear system

According to the equation (2) in Part *A* and Mercer condition, the inner product of $\varphi(x_i) \cdot \varphi(x_j)$ can be replaced by kernel function $K(x_i, x_j)$. Then the quadratic optimization problem and the optimal hyperplane turn to the following:

$$\max_{\alpha} \sum_{i=1}^{n} \alpha_i - 0.5 \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i \cdot x_j) \tag{4}$$

$$f(x) = \operatorname{sgn}\left[\sum_{i=1}^{n} \omega^{*} \bullet \quad x \quad b\right] \tag{5}$$

Generally, there are four forms of Kernel function to conduct this mapping, and they are:

Linear Kernel: $K(x_i, x) = x_i^T \cdot x$;

d-order Polynomial Kernel: $K(x_i, x) = (x_i \cdot x + 1)^d$;

RBF Kernel: $K(x_i, x) = \exp\left\{-\|x - x_i\|^2 / \sigma^2\right\}$;

Sigmoid Kernel: $K(x_i, x) = \tanh(ax^T + c)$.

Experience shows that SVM based on the RBF kernel has better effect in classification and prediction; therefore, RBF kernel is being used in this paper.

C. Feasible Moving Hyperplane of Dynamic SVM

Generally, fault diagnosis based on SVM obtains a model and a fixed optimal hyperplane for prediction by training historical sampling data. While in reality, the equipment state varies along with the time, and a fixed model will lead to the rising of false alert. Therefore, this paper proposes this dynamic SVM.

Obviously, normal state data is more than the failure data, and equipment always works in the normal state. So in the analysis process, it's more hoped that the present optimal hyperplane is concerned with the new sampling data so as to decrease probability of false alert.

Assume an initial training data set:

$$D = \{x_1, x_2, \cdots \tag{6}$$

By means of SVM training, the support vectors and optimal hyperplane can be obtained in the model. At the same time, there are also many nonsupport vectors, which make no contribution to form the hyperplane. Thus as time going on, SVM predicts the fault state by testing data and updates the previous training data as follows:

$$\tilde{D} = \begin{cases} D, d \geq (1+\theta)/\|\omega\| \\ D_1, d < (1+\theta)/\|\omega\| \end{cases} \tag{7}$$

s.t.

$$D_1 = \{x_2, x_3, \cdots \qquad \} \tag{8}$$

$$\theta \in (0, 0.2) \tag{9}$$

Then, the testing data sampled is being tested continually while training data set updates. With this flow of dynamic data, the optimal hyperplane at the close moment can be achieved each time, which can reflect the approximate real state and promote the fault diagnosis accuracy, as Fig.3 shows.

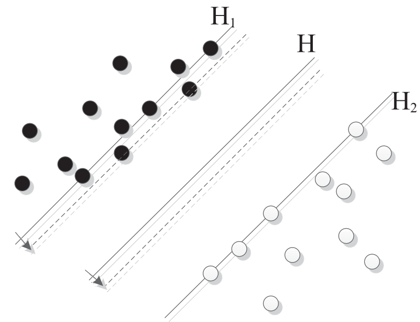

Figure 3. Diagram of dynamic SVM with moving optimal hyperplane

Based on the discussion on structure of best optimal hyperplane, the dynamic SVM algorithm is displayed in the following. This method firstly selects limited historical data that includes typical normal data and failure data, trains which to build a model and forms the initial optimal hyperplane. Next, analysis the distance and update training data according to the spatial distribution of testing data as equation (7) ~ (9). If the new testing data is near the canonical hyperplanes, then replace a nonsupport vector with this data for it may influence the new hyperplane; while this new testing data is far away from the canonical hyperplanes, ignore this data and test the next data. Then the training data updates each time and the optimal hyperplane at present continually forms, which could better validate the sampling data in terms of time sequence.

The diagram of fault diagnosis by dynamic SVM is shown in Fig.4. This figure shows that sampling data is being dynamically diagnosed; meanwhile, the useful and important data comes into being part of the training data to realize the supervision function.

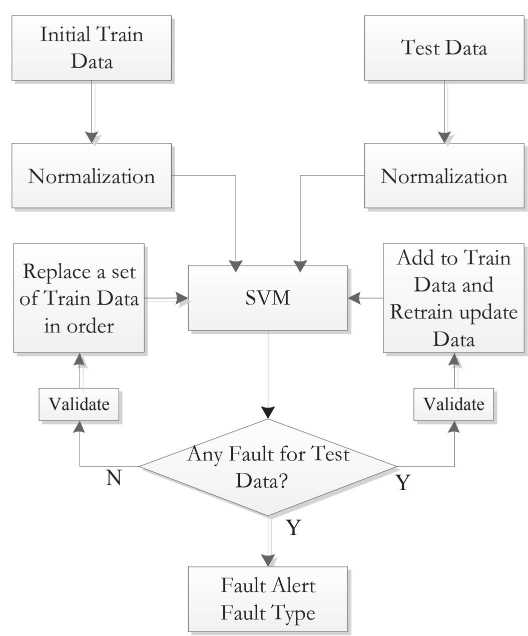

Figure 4. Diagram of fault diagnosis based on dynamic SVM

III. Dynamic SVM Application on Fault Diagnosis of Gear Case

To validate the efficiency of dynamic SVM, this paper uses the sampling data from the simulative gear case in QPZZ-II system to conduct the analysis.

A. Extracting Features

For the gear case system, this paper selects 9 dimensional data sampled from different sensors which measure the rotate speed, the X shift and Y shift of input shaft, the Y acceleration of input shaft's motor side bearing, the Y acceleration of output shaft's motor side bearing, the Y acceleration of input shaft's load side bearing, the X and Y acceleration of output shaft's load side bearing, and the X magneto-electric speed of output shaft's load side bearing, respectively.

Meantime, this experiment lists a list of data for training and testing, which contains 20 sets of normal data, 10 sets of pitting failure data and 10 sets of snaggletooth failure data. Of all data, 10 sets of the normal data, 5 sets of the pitting failure data and 5 sets of the snaggletooth data are used for the initial training. The rest are used for testing. For the sampling data, the real rotate speed of gear case is 1470 r/min, and the frequency is 20 kHz.

To start the diagnosis, normalize all data into interval [0 1] according to (10) as follows.

$$x_{nrom} = \left(x - x_{\min}\right) / \left(x_{\max} - x_{\min}\right) \qquad (10)$$

Some normalized data is as shown in TABLE I. It reflects the information of 3 kind states of gear case.

TABLE I. Part of the Normalization Training Data Set

Failure Type	Atr1	Atr2	Atr3	Atr4	Atr5	Atr6	Atr7	Atr8	Atr9
0	0.9771	0.9957	0.2408	0.6086	0.5170	0.7539	0.6556	0.9964	0.0014
0	0.9824	0.9945	0.1275	0.6348	0.5510	0.8753	0.5166	0.9912	0.0054
0	……	……	……	……	……	……	……	……	……
1	0	0.4519	0.2210	0.6985	0.3719	0.7487	0.5596	0.9896	0.0045
1	0.0127	0.4556	0.0708	0.6816	0.2948	0.7955	0.4868	0.9940	0.0050
0	……	……	……	……	……	……	……	……	……
2	0.5655	0.0079	0.3201	0.7285	0.3787	0.7799	0.5464	0.4386	0.0032
2	0.5570	0.0329	0.0312	0.3483	0.5011	0.7972	0.6821	0	0.0050

In which "0" represents the normal state, "1" represents the pitting failure and "2" represents the snaggletooth failure.

B. Training the Networks

Based on the characteristics of different fault types, the binary-tree SVM are used to analyze and identify the training data, including 10 sets of the normal state, 5 sets of the pitting state and 5 sets of the snaggletooth state. With all training samples, the SVM firstly separates the normal state from the other abnormal states; then the SVM separates the fault 1 from the fault 2. Thus, the initial multiple SVM classifiers are obtained.

C. Testing with Networks

In this process, the testing data are validated in SVM and dynamic SVM methods, respectively. For the SVM, all the fault state of testing data are predicted by the initial model in Part *B;* while for the dynamic SVM, the fault is being predicted along with the testing data updates according to its algorithm.

In this process, the suitable parameters seriously affect the accuracy of diagnosis. For the dynamic SVM, its optimal parameters may differ each time. Therefore, this paper adopts the PSO algorithm to find the most suitable parameters C and σ. Simulation shows that this process will take a lot of time and slow down the efficiency, thus set a relatively small maxi generation and scale to decrease time cost.

D. Experimental Results

Finally, diagnose the fault types by means of SVM and DSVM, respectively. Result shows that the SVM has an accuracy of 75% while the dynamic SVM has an accuracy of 90%, indicating its suitability in dealing with the running machines that wear down quickly and less false alert to the random data. The comparison of the two methods in prediction is shown as follows in Fig.5.

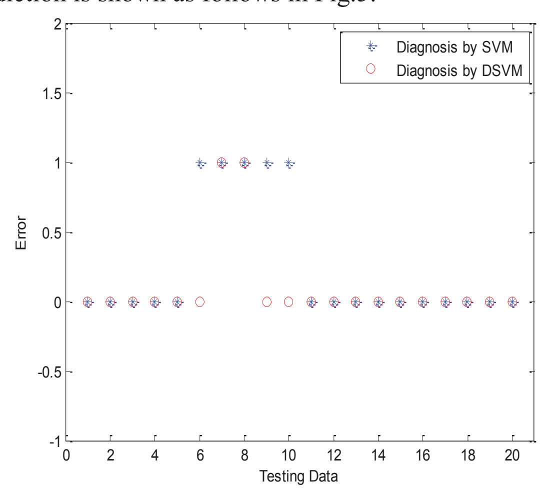

Figure 5. Comparison of SVM method and DSVM method

The above results show that the dynamic SVM can better identify the fault type of gear case. However, the accuracy of SVM is seriously affected by the data while in reality the noisy cannot be avoided. Therefore, this paper takes noise into consideration and tests its robustness.

In order to overcome lack of testing samples, 5%, 10%, 15%, 20% random noise is respectively added to the testing data. The value of feature vector is change to:

$$X' = \left[x_1', x_2', \cdots \tag{11}$$

where $x_i' = x_i \cdot \left(1 + t \cdot randn\right)$, x_i is the feature extracted from testing data and t is the percentage of random noise.

In order to ensure the accuracy not disturbed by the randomness of noise, each method use 500 group of random data to compute the average value. Finally, the results are shown in TABLE II.

TABLE II. RECOGNITION ACCURACY OF SVM AND DSVM

t (%)	SVM	DSVM
5	30.72	50.45
10	28.85	49.99
15	26.81	48.42
20	25.04	47.88

From TABLE II, it's obviously that diagnosis accuracy of the two methods is reduced very much for the existence of noise. However, compared to SVM method, the dynamic SVM seems more robust. With more intense noise, the latter shows better diagnosis capacity than the former.

IV. CONCLUSIONS

SVM is a useful machine-learning method dealing with small sample problem, and possesses good generalization ability, which indicates its good quality for intelligent fault diagnosis. The proposed methos is being improved for the dynamic system based on the binary-tree SVM classification method. It realizes the replacement of training data set and its adjustment of the best optimal hyperplane along with the testing data sampled from present state, which helps to better predict the fault type. Meanwhile, the new method shows some robustness in fault prediction and provides a good idea for the real-time diagnosis. For the fault diagnosis cannot work well with disturbance, there still remain many problems, such as the structure SVMs and the time cost in finding best parameter of SVM, which all needs to be studied in the future.

REFERENCES

[1] Hongwei Li, Dongsheng Yang, Yilan Sun. Study review and prospect of intelligent fault diagnosis technique[J]. Computer Engineering and Design, 2013, 34(2): 632-637.

[2] Dali Qin. Research on intelligent diagnosis for equipment fault based on knowledge management[D]. Hunan University, 2014.

[3] Tao Jiang, Shengfa Yuan. Rolling bearing faults diagnosis based on the improved wavelet neural network[J]. Journal of Huazhong Agricultural University, 2014, 33(1): 131-136.

[4] Hao Pan, Huawei Zhang, Meiling Gao, Main reduction gear quality diagnosis research based on SA-BP algorithm[J]. Journal of Wuhan University of Technology, 2011, 33(1): 161-164.

[5] Yao shi, Qingyang Ren. Research of fault diagnosis based on SVM for mine ventilator[J]. Automation and Instrumentation, 2013, 5:18-20.

[6] Guobiao Wang, Zhengjia He, Xuefeng Chen, Basic Research on Machinery Fault Diagnosis – What is the Prescription[J]. Journal of Mechanical Engineering, 2013, 49(1): 63-72.

[7] Shengfeng Cheng, Xiaohua Cheng, Lu Yang. Application of wavelet neural network with improved partical swarm optimization algorithm in power transformer fault diagnosis[J]. Power System Protection and Control, 2014, 42(19): 37-42.

[8] Kaixin Li. Based on an improved Genetic Algorithm for Aero Engine fault Diagnosis Expert System[D], Nanchang HangKong University, 2012.

[9] Jianguo Cui, Ming Li, Xicheng Chen, Health diagnosis method of the aero-craft based on SVM[J]. Piezoelectrics and Acoustooptics, 2009, 31(2): 266-269.

[10] Tiebin Zhu. Development of fault diagnosis application system for turbofan engine based on data[D]. Nanjing University of Aeronautics and Astronautics, 2013.

[11] Shaolei Zhou, Jian Liao, XianJun Shi. Kernel parameter selection of RBF-SVM and its application in fault diagnosis[J]. Journal of Electronic Measurement and Instrumentation, 2014, 28(3): 240-246.

[12] Huimin Zhao, Caihua Fang, Wu Deng. Research on motor diagnosis model for support machine based on intelligent optimization methods[J]. Journal of Dalian University, 2016, 37(1): 92-96.

[13] Qinghua Wang, Tianhai Yang, Runjie Shen, et al. Fault caused by vibration diagnosis expert system for a pump storage group[J]. Journal of vibration and shock, 2012, 31(7): 158-161+170.

[14] Shan He, Weiqing Wang, Xinyan Zhang, et al. Short circuit fault intelligent diagnosis of MW permanent magnet wind power generator based on artificial neural network[J]. Electric machines & control applications, 2011, 38(9): 24-29.

[15] Donghua Zhou, JianTao Zhou, Xiao He. Review of intermittent fault diagnosis techniques for dynamic system[J]. Acta Automatica Sinica, 2014, 40(2): 161-171.

[16] Chuanyang Du, Dongjian Zheng, Yi Zhang. Dam deformation monitoring model based on dynamic SVM-MC and its application[J], Water Resource and Power. 2015, 33(1): 71-74.

[17] Heng Zhang, Lingjun Li, Chuang Wang. Research of fault diagnosis based on DSVDD[J]. Machinery Design & Manufacture, 2011, 10: 187-189.

[18] Hui Yi. A study of Support Vector Machines based on fault diagnosis and its application[D]. Nanjing University of Aeronautics and Astronautics, 2011.

[19] Liangmou Hu, Keqiang Cao, Haojun Xu, Xinmin Dong, Fault diagnosis and Control technology based on Support Vector Machine, National Defense Industry Press, 2011.