

# Fault Diagnosis of Wind Turbines' Bearing Based on PSO-AdaBoostSVM

Hao Gan\* and Bin Jiao

Shanghai Dianji University, Shanghai, China

\*Corresponding author

**Abstract**—Bearing is an important part of wind turbine. In the past few years, there have been many intelligent fault diagnosis algorithms for it. Support vector machine is one of them. But it also has many shortcomings for fault diagnosis. For example, how to select kernel and parameters makes the classifier more accurate. In this paper, in order to find the best global parameters, we choose the PSO algorithm, and we also use the adaboost algorithm to improve the classification accuracy. By comparing the experimental results of other classifiers, the proposed PSO-AdaBoostSVM algorithm is superior to SVM and AdaBoostSVM in classification accuracy. So the proposed algorithm can be used in the fault diagnosis of wind Turbines' Bearing.

**Keywords**—fault diagnosis; PSO; bearing; PSO-AdaBoostSVM

## I. INTRODUCTION

Over the past few years, there are many methods for fault diagnosis of wind turbines. Such as BP SVM ELM. Support Vector Machine [1] was proposed by Cortes and Vapnik in 1995 which is a kind of generalized linear classifier. The main idea of SVM can be summarized as two points: (1) It is for linearly separable cases classifier. (2) For linearly inseparable cases, the linearly separable sample is transformed into a high-dimensional feature space to be linearly separable using a nonlinear mapping algorithm. As a classifier, it can be well use in classifying. However, it also have some Shortcomings. The RBF kernel is one of the popular kernels used by SVM, which has a parameter known as Gaussian width,  $\sigma$ . In contrast to the RBF networks, SVM with the RBF kernel can automatically determine the number and location of the centers and the weight values [2].

AdaBoost is a machine learning algorithm based on Boosting idea. It can enhance a weak classifier with slightly higher prediction accuracy, and ensemble it to a strong classifier with high prediction accuracy. AdaBoost creates a collection of component classifiers by maintaining a set of weights over training samples and adaptively adjusting these weights after each Boosting iteration [3]. There have been many ways to proposed in weight updating. Many studies that use Decision Trees or Neural Networks as component classifiers for AdaBoost have been reported. The article [4] used SVM as a component classifier of AdaBoost. We know that it important to choose a kernel for SVM. RBF kernel is more popular these days. But it has really big influence on selecting proper values of C and  $\sigma$ . [4] proposed that  $\sigma$  value adaptively adjusted to obtain a set of moderately accurate RBF SVM component classifiers. Although the AdaBoostSVM

algorithm have a good classification performance, it also needs to set the different value of C and  $\sigma$  step in advance. Liu X [5] has proved that SVM can be a component classifiers for AdaBoost. He also proposed a GA method to enhance the accuracy of AdaBoostSVM. Kapoor N [6] use PSO optimized SVM controller. Some researchers also used AdaBoost algorithm to improve other algorithm. Such as, Liu X [7] use adaboost to enhance BP algorithm in hydraulic system's Fault diagnosis.

In this paper, it uses RBF SVM as the component classifier of AdaBoost. We know that the value of parameters have really significant influence of the component classifier during the AdaBoost iterations. How to choose the right and accurate parameters (C and  $\sigma$ ) is crucial for the classification accuracy of the model. So, a global optimization algorithm can be chosen to find the appropriate parameters. This paper choose the PSO algorithm to iterative optimize by updating the speed and location of the particle. And then it proposed the multi-classified strategies to diagnosis the fault of wind turbine's gearbox. At last, after AdaBoost has enhanced the couples of wear classifier, it has proved that making diagnosis more accurate.

## II. PROPOSED ALGORITHM

### A. Basic Support Vector Machine

SVM is used to solve the problem of data classification in the field of pattern recognition and belongs to a type of supervised learning algorithm. SVM is design for linearly separable cases. But in non-linear conditions, it can transform the original input set to a high-dimensional feature space by using a kernel function, and then achieve optimum classification in feature space [8]. SVM tries to identify a hyper-plane in a multi-dimensional space. Its decision function can be expressed as follows:

$$f(x) = \langle w, \phi(x) \rangle + b \quad (1)$$

The problem of the classifier is that to optimal the w and b. So, it can be solved as the following formula:

$$\min[ g(w, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i ] \quad (2)$$

$$\text{S.t } y_i(\langle w, \phi(x_i) \rangle + b) \geq 1 - \xi_i, \xi_i \geq 0 \quad (3)$$

$\|w\|$  is minimized with Eq. (2). the distance between two support vector points on two sides of this hyper-plane is maximal. It also can be changed into:

$$\min_a \frac{1}{2} \sum_{i=1}^j \sum_{j=1}^l y_i y_j a_i a_j K(x_i, x_j) - \sum_{j=1}^l a_j \quad (4)$$

$$\text{S.t } \sum_{i=1}^l y_i a_i = 0, 0 \leq a_i \leq C, i = 1, \dots, l \quad (5)$$

Decision function:

$$f(x) = \text{sign} \left( \sum_{i=1}^N a_i^0 y_i x_i x + b_0 \right) \quad (6)$$

### B. Particle Swarm Optimization

PSO seeks optimal solutions through collaboration and information sharing among individuals in a group. Particles' characteristics are represented by three parameters, such as position, velocity and fitness[9]. It has been widely used in function optimization, neural network training, fuzzy system control and other genetic algorithms. It can be expressed as updating of the positions and velocities in each iteration. As shown in the following formula:

$$V_{id}^{k+1} = \omega V_{id}^k + c_1 r_1 (P_{id}^k - X_{id}^k) + c_2 r_2 (P_{gd}^k - X_{id}^k) \quad (7)$$

$$X_{id}^{k+1} = X_{id}^k + V_{id}^{k+1} \quad (8)$$

### C. AdaBoost Algorithm

Boosting, also known as reinforcement learning or upgrade process, is an important ensemble learning technology. In the AdaBoost algorithm, each training sample is given a weight indicating the probability that it is selected into the training set by a classification classifier. It can be briefly described as fellow steps (See Table I):

TABLE I. ADABOOST ALGORITHM

<b>AdaBoost</b>
Input: training samples $[(x_1, y_1) \dots (x_n, y_n)]$ where $x_i \in X, y_i \in Y = [1, -1]$ , training times T,;
Initialization: the weight of training samples are the same value $D_1(i) = 1/N$
Iterative update: For t=1...T Training weak classifier using distribution $D_t$ , component classifier prediction is $h_t(x_i)$ ;
Training error $\mathcal{E}_t$ is calculated by the formula: $\mathcal{E}_t = \sum_1^N D_t^i, y_i \neq h_t(x_i)$
Set weight for component classifier: $a_t = \frac{1}{2} \ln \left( \frac{1 - \mathcal{E}_t}{\mathcal{E}_t} \right)$
Update the weight of training samples: $D_{t+1} = \frac{D_t(i) \exp(-a_t y_i h_t(x_i))}{Z_t}$
Output: strong classifier prediction: $H(x) = \text{sign} \left( \sum_{t=1}^T a_t h_t(x) \right)$

### D. Proposed Algorithm: PSO-AdaBoostSVM

In this paper, the problem of gearbox's diagnosis is always a complex issue. Its vibration signal is nonlinear and non-stationary. So RBF is a good choice for a kernel of SVM. How to set the suitable  $\sigma$  and  $C$  values for the component classifier of RBF SVM is critical at the beginning of adaboost iterations. Even more, the parameters of different base classifiers have different results for the classification of ensemble classifiers. PSO algorithm is a global optimization algorithm. In this paper, PSO is used to search the optimal set of parameter values of the ensemble classifier. And then, making the classification more accuracy. The PSO-AdaBoostSVM is showed in Table II:

TABLE II. PSO-ADABOOSTSVM ALGORITHM

<b>PSO-AdaBoostSVM</b>
Input: training samples $[(x_1, y_1) \dots (x_n, y_n)]$ where $x_i \in X, y_i \in Y = [1, -1]$ , training error Threshold $\mathcal{E}_{th}$ ,
the $i$ th particle $X_i$ , velocity is $V_i$ , individual and globe optimal value is $P_i, P_g$ ;
Initialization: initialize the individual by defining random velocities and positions
Iterative update: IF ( $loop < n_{\max}$ (maximum iterations))
Evaluate the fitness value of each AdaBoostSVM
Update the position and velocity.
Output: the best globe optimal value

**E. Proposed Algorithm: PSO-AdaBoostSVM**

SVM is designed for binary problems, so how to solve the multiple classification problems in fault diagnosis? This paper use the one-versus-rest method. In briefly, It classifies a sample of one category into one category and the rest of the samples is classified into another category. The number of classifiers is  $k(k - 1)/2$ .

**III. SIMULATION EXPERIMENT**

In this paper, it designed a simulation experiment for the gearbox’s fault diagnosis. The system installs vibration sensors in the gearbox to get vibration data for the operation of the gearbox. And then extract features of vibration signals. It collected the operating data in the four conditions of the gearbox through experiments, which is normal, out ring damage, inner ring damage and ball damage. Fault diagnosis is Mainly divided into the following steps:

1) Firstly, it selected 8192 sample points, and the sensor sampling frequency is 93.7 kHz to get the high-speed input horizontal vibration signal.

2) In order to increase the speed of computer operation, we randomly intercept 300 points into a group, a total of 200 groups of vibration data. 150 of them are training data and 50 of them are test data.

3) Separately extract feature values of training data and test data, which is included Time-Domain and Frequency-Domain Features: Kurtosis index, Margin, Skewed, Power spectrum center of gravity, Power spectral variance, Harmonic factor, Peak factor, Compressed energy, Separately extract feature values of training data and test data, which is included Time-Domain and Frequency-Domain Features: Kurtosis index, Margin, Skewed, Power spectrum center of gravity, Power spectral variance, Harmonic factor, Peak factor, Compressed energy, Compressed Information entropy. Showed in Table III:

4) The feature parameters extracted from all the groups were substituted into the classification model (such as SVM AdaBoostSVM, PSO-AdaBoostSVM) for training, and the advantages of the proposed algorithm were compared with other algorithm.

TABLE III. CHARACTERISTIC PARAMETERS IN DIFFERENT CONDITIONS

Practical fault type	Margin factor	Skewed index	Power spectral variance	Harmonic factor	Kurtosis index	Peak factor	Compressed energy	Compressed Information entropy	label
normal	-0.1075	0.0661	0.0225	-0.5680	2.9055	2.5961	0.4624	0.3154	1
Bearing out ring damage	0.0705	1.7954	0.0118	-0.2011	14.1678	7.3196	0.2483	0.1425	2
Bearing boll damage	0.0696	2.1895	0.0129	-0.4760	14.9366	7.4929	0.2153	0.2166	3
Bearing inner ring damage	-0.8880	0.3574	-0.0339	-0.4908	4.1899	4.3002	0.5361	0.1752	4

The paper use different classifiers to test the effect of diagnosis. The result is showed in the following figures (Figure I-III):

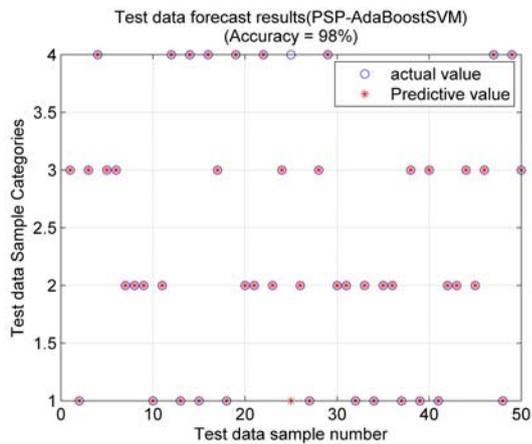


FIGURE I. RESULT OF PSO-ADABOOSTSVM

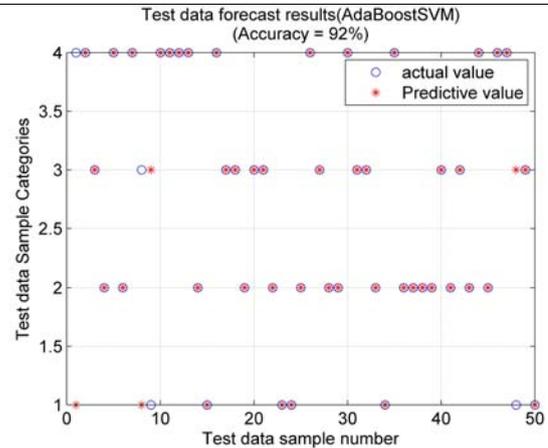


FIGURE II. RESULT OF ADABOOSTSVM

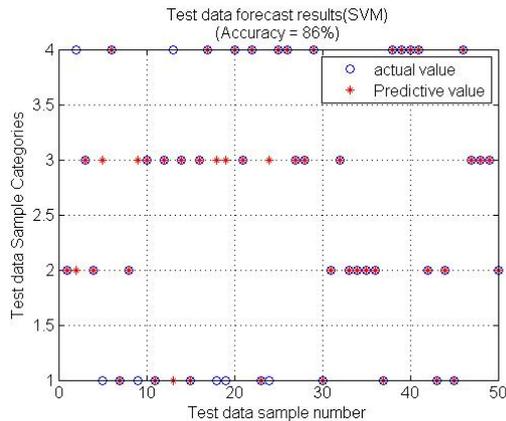


FIGURE III. RESULT OF SVM

From the above figure, we can see that the proposed algorithm has higher accuracy than other algorithms, which is 98%. And AdaBoostSVM is 92%, SVM is 86%.

#### IV. CONCLUSION

SVM is a good classifier, but it has many shortcomings. The adaboost algorithm improves its performance to some extent. But it is not enough. The algorithm proposed in this paper further enhances the performance of the classifier, making it well applicable to the fault diagnosis of wind turbine bearings.

#### ACKNOWLEDGMENT

In this paper, I would like to thank my tutor for giving me guidance. Thanks the headmaster for helping me during the experiment. This experimental study taught me a lot.

#### REFERENCE

- [1] Islam M M, Kim J, Khan S A, et al. Reliable bearing fault diagnosis using Bayesian inference-based multi-class support vector machines[J]. Journal of the Acoustical Society of America, 2017, 141(2):EL89.
- [2] Ring M, Eskofier B M. An approximation of the Gaussian RBF kernel for efficient classification with SVMs[J]. Pattern Recognition Letters, 2016, 84:107-113.
- [3] Yao M, Zhu C. SVM and adaboost-based classifiers with fast PCA for face recognition[C]// IEEE International Conference on Consumer Electronics-China. IEEE, 2017:1-5.
- [4] Li X, Wang L, Sung E. AdaBoost with SVM-based component classifiers[J]. Engineering Applications of Artificial Intelligence, 2008, 21(5):785-795.
- [5] Liu X, Gang C, Gao W, et al. GA-AdaBoostSVM classifier empowered wireless network diagnosis[J]. Eurasip Journal on Wireless Communications & Networking, 2018, 2018(1):77.
- [6] Kapoor N, Ohri J. GA and PSO optimised SVM controller for manipulator[J]. 2016, 2(3):121.
- [7] Liu X, Hu Y, Xu Z, et al. Fault diagnosis for hydraulic system of naval gun based on BP-Adaboost model[C]// Second International Conference on Reliability Systems Engineering. IEEE, 2017:1-6.
- [8] Lin Y L, Wei G. Speech emotion recognition based on HMM and SVM[C]. Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on. IEEE, 2005:4898-4901 Vol. 8.
- [9] Zhou Z, Jiao B. The improvement of particle swarm optimization[C]// International Conference on Systems and Informatics. IEEE, 2017:373-377.
- [10] Baig M M, Awais M M, El-Alfy E S M. AdaBoost-based artificial neural network learning[J]. Neurocomputing, 2017, 248.

- [11] Wang J, Xiong X, Zhou N, et al. Early warning method for transmission line galloping based on SVM and AdaBoost bi-level classifiers[J]. Iet Generation Transmission & Distribution, 2016, 10(14):3499-3507.
- [12] Sun J, Fujita H, Chen P, et al. Dynamic financial distress prediction with concept drift based on time weighting combined with Adaboost support vector machine ensemble[J]. Knowledge-Based Systems, 2016, 120(C):4-14.
- [13] Yao M, Zhu C. SVM and adaboost-based classifiers with fast PCA for face recognition[C]// IEEE International Conference on Consumer Electronics-China. IEEE, 2017:1-5.