TELKOMNIKA, Vol. 11, No. 6, June 2013, pp. 3451 ~ 3462
e-ISSN: 2087-278X

Fan Fault Diagnosis Based on Wavelet Packet and Sample Entropy

Xiaogang XU*, Songling Wang, Jinlian Liu, Zhengren Wu
School of Energy Power and Mechanical Engineering, North China Electric Power University, Baoding, 071003, Hebei Province
*Corresponding author, e-mail: hdxuxiaogang@163.com

Abstract

To accurately diagnose the mechanical failure of the fan, two diagnostic methods based on the wavelet packet energy feature and sample entropy feature are proposed. Vibration signals acquisition of 13 kinds of running states are achieved on the 4-73 No. 8D centrifugal fan test bench. The wavelet packet energy feature vector of each vibration signal is rapidly extracted through the wavelet packet denoising, decomposition and reconstruction. The vibration signal wavelet packet energy feature vector of the five measuring points in the same instantaneous running state are fused into the fan fault feature vector. Finally, the fault diagnosis of the fan is achieved by using improved SVM (Support Vector Machine) classifier, and the accuracy rate is 94.6%. A new fan fault feature vector is put forward, which is the integration of the vibration signal sample entropy of the five measuring points in the same instantaneous running state, and then the fault diagnosis of the fan is achieved by using improved BP (Back Propagation) neural network, and the accuracy rate is 99.23%. The diagnostic results show that these two methods are able to effectively diagnose the category, severity and site of the fan mechanical failures, and suitable for online diagnosis.

Keywords*: fault diagnosis, wavelet packet, support vector machine, sample entropy, neural network;*

1. Introduction

The running state of the fan has the direct bearing on the safe and economic operation of the power plant, and the reliability, security and economy of the fan depends on the efficient operation, the real-time status tracking evaluation, the accurate diagnosis and repair, thus it has great significance to launch the fault diagnosis research of the fan. The rotor unbalance, the rotor misalignment, the bearing looseness and the static and kinetic friction are the common mechanical failures of the fan. There are lots of fault diagnosis methods, but basically divided into three steps: acquisition of the diagnostic information; fault feature extraction; state recognition and fault diagnosis.

The vibration signal of the fan rotation process has closely related to the type, extent and location of the fan mechanical failure, contains a wealth of equipment state information. From the perspective of information theory, the vibration signal has the characteristic of maximum information entropy. From the perspective of practical application, the vibration signal has a series of advantages, including detection convenient, wide applicability, non-contact measurement and multidimensional measurement. Therefore, collecting the vibration signal is the most widely used method for fault monitoring and diagnosis currently. It is necessary to analyze the signals after sufficient vibration signals are collected. The traditional signal feature extraction methods are based on the stationary of the signal and can not analyze the non-stationary signal. Due to the driving force, the damping force, the nonlinearity of the elastic force and the nonlinearity of the mechanical system in the fan operation process, the detected vibration signal is a non-stationary signal, so the traditional signal feature extraction methods have a large limitation for the fan vibration signal. Wavelet packet analysis is the improvement of wavelet analysis [1-2], which takes the advantages of the short time Fourier transform and wavelet transform into account. The signal is decomposed into different frequency bands based on the multiresolution analysis theory of wavelet packet decomposition, which has many advantages for handling the discontinuous, saltatorial and non-stationary vibration signals [3-4]. Therefore, wavelet packet transform is used as a common method to extract the fault feature for

Received January 26, 2013; Revised April 22, 2013; Accepted April 30, 2013

its excellent time-frequency resolution property [5-7]. The category, severity and site of the fan mechanical failures have a great impact on the frequency band energy feature of wavelet packet reconstruction, thus the wavelet packet reconstruction frequency band energy feature can be the fault signature of the fan.

The irregularity and complexity of the vibration signal can reflect the occurrance and development of fault. There are lots of methods to measure signal complexity, including the Lempel-Zi complex degree, approximate entropy and sample entropy and permutation entropy. Yan [8] used approximate entropy for bearing condition monitoring and achieved good results. The sample entropy is the improved algorithm of the approximate entropy, and its superiority is less dependent on the length of the time series, therefore the ability to distinguish of sample entropy is stronger than the approximate entropy. Zhao Zhihong and Yang Shaopu [9] proposed a bearing fault diagnosis method based on wavelet packet transform and sample entropy. By the high fault identification rate, the sample entropy analysis of the vibration signal is an effective fault feature extraction method.

The neural network is an effective method of fault diagnosis based on its input and output nonlinear mapping, parallel processing and a high degree of self-organization and self-learning ability [10]. BP neural network is one of the most widely used network in the various fields. But there are many inadequacies, for example, network training requires a sufficient sample size, Network is easy to get the local minima rather than the global optimum value, the learning efficiency is low for too many training times, slow convergence. Thus, the fault diagnosis method using improved BP neural network has better timeliness and higher accuracy rate. In addition, SVM is also an effective fault diagnostic method because of its efficiency of the calculation, strong adaptability and strong classification ability [11-14][24]. SVM can achieve satisfactory results without data preprocessing. In the application of support vector machine, the selection of the appropriate parameters, especially the punishment factor c and the kernel function parameter σ, are important issues that need to be addressed.

The remaining sections of this paper are organized as follows. Section 2 gives a brief review of the principle of the wavelet packet denoising, decomposition and reconstruction, and the method to extract the energy feature of frequency band after wavelet packet decomposition and reconstruction. Section 3 gives a brief review of the sample entropy algorithm. Section 4 depicts the support vector machine multi-fault classifier, and details the process of using the particle swarm algorithm to optimize the support vector machine multi-fault classifier. Section 5 gives a brief review of the principle of the BP neural network, and details the improved method of the standard BP neural network. Section 6 is the focus of this article, including two diagnostic methods of fan machinery failures. Concluding remarks based on the results constitute the contents in Section 7.

2. Wavelet Packet Principle and Energy Feature Extraction

2.1. Wavelet Packet Decomposition and Reconstruction

In engineering applications, the wavelet packet decomposition can be attributed to the following recursive [25]:

$$W_{2n}(t) = \sum_k h_k W_n(2t-k)$$

$$W_{2n+1}(t) = \sum_k g_k W_n(2t-k)$$

Thus let the signal $W_1(t)$ pass through the high-pass and low-pass combined orthogonal mirror filters h_k, g_k. The essence of wavelet packet decomposition is to decompose a signal into different frequency band step by step. If the original signal data is sufficient, the frequency band can be divided to fine enough, the frequency bandwidth of the wavelet packet decomposition Δf and decomposition level j and sampling frequency f_s satisfies the relationship as follows:

$$\Delta f = f_s / 2^{j+1}$$

The appropriate selection of decomposition layer can get the desired frequency band by the above formula, thereby the characteristic signals of the original signal can be separated from the interference and noise. The frequency of the component retained in the characteristic signals needn't be precisely positioned since the frequency bands have a certain width, so wavelet packet decomposition has a certain ability to adapt the frequency drift.

The recursive formulas of the wavelet packet reconstruction are as follows:

$$W_2(2t) = \sum_k h_{2k+1} W_{2n}(t-k) + \sum_k g_{2k+1} W_{2n+1}(t-k)$$

$$W_n(2t-1) = \sum_k h_{2k} W_{2n}(t-k) + \sum_k g_{2k} W_{2n+1}(t-k)$$

The advantage of using wavelet packet to reconstruct the signal is that the information on the frequency band can be retained according to the demand and the remaining frequency band (interference, noise) will be cleared. Thus the noise signal and the characteristic signal can be decomposed into different frequency bands in order to reconstruct the characteristic signal without the noise and interference.

2.2. Wavelet Packet Energy Feature Extraction
2.2.1. Wavelet Packet Denoising

The original vibration signal of the fan includes a lot of noise. Therefore, the energy of the noise will drown the energy feature of the fan running state, if the frequency band energy feature is directly extracted by using the decomposition and reconstruction wavelet packet. So the wavelet packet denoising of the original vibration signal is necessary. Wavelet packet denoising steps are as follows:

(1) Use wavelet packet to decompose the fault signals. In this paper, the layer of the wavelet packet decomposition is five.

(2) Calculate the best tree. In this paper, the best tree is calculated based on the Shannon entropy standards.

(3) Calculate the threshold of wavelet packet decomposition coefficients. Choose an appropriate threshold value for each wavelet packet decomposition coefficients and implement the coefficient threshold quantization. Based on the characteristics of the noise and signal, this paper integrated used the adaptive threshold selection algorithm of the Stein unbiased risk estimation theory, selected $\sqrt{2\ln(length(x))}$ as the threshold algorithm and manually adjusted the threshold value based on the effect of denoising.

(4) The wavelet packet reconstruction: Implement wavelet packet reconstruction after quantize the wavelet packet decomposition coefficients of the 5th floor.

The selection of the threshold and the threshold quantization are the most critical issues among the four steps above, to some extent, they are directly related to the quality of denoising.

2.2.2. Wavelet Packet Reconstruction Frequency Band Energy Feature Extraction

The wavelet packet energy feature extraction steps of fault signal are as follows:

(1) The vibration signal S is decomposed by l layer wavelet packet, and wavelet packet decomposition coefficients of the l layer are extracted.

(2) Reconstruct decomposition coefficients in order to extract the signal of each frequency band. $S_j\,(j = 0,1,\cdots,2^l-1)$ represents the reconstructed signal of the each node of the l layer, then the total signal S can be expressed as follows:

$$S = \sum_{j=0}^{2^l-1} S_j$$

(3) Seeking the total energy of each frequency band signal. The corresponding energy $E_j\,(j = 0,1,\cdots,2^l-1)$ of each band reconstructed signal $S_j\,(j = 0,1,\cdots,2^l-1)$ can be expressed as follows:

$$E_j = \int \left|S_j(t)\right|^2 dt = \sum_{k=1}^{n} \left|x_{jk}\right|^2$$

Where, $x_{jk}\,(j = 0,1,\cdots,2^l-1;\, k = 1,2,\cdots,n)$ represents the amplitude of the discrete points of the reconstructed signal S_j, n represents the number of the discrete points of the vibration signal S.

(4) Wavelet packet energy feature vector T is established. $T=[E_0,E_1,E_2,\cdots,E_{2^n-1}]$ Through the process of normalization, we can get vector T'.

$$E'=\sum_{j=0}^{2^n-1}E_j$$

$$T'=T/E'$$

3. The Sample Entropy Algorithm

Assume that the original data is $x(1),x(2),\cdots x(N)$, including N points. The sample entropy of the original data is calculated as follows [15]:

(1) The vectors $X_m(1),\cdots X_m(N-m+1)$ are composed by the ordinal number of the original data, and the dimension of each vector is m. $X_m(i)=\{x(i),x(i+1),\cdots,x(i+m-1)\},1\le i\le N-m+1$

(2) $d[X_m(i),X_m(j)]$ represents the distance between the vector $X_m(i)$ and the vector $X_m(j)$. $d[X_m(i),X_m(j)]=\max_{k=0,\cdots,m-1}(x(i+k)-x(j+k))$

(3) B_i is the number of the $j\,(1\le j\le N-m,\ j\ne i)$ when $d[X_m(i),X_m(j)]\le r$. Thus $B_i^m(r)$ is defined as follows: $B_i^m(r)=\dfrac{1}{N-m-1}B_i,1\le i\le N-m$

(4) $B^m(r)$ is defined as follows: $B^m(r)=\dfrac{1}{N-m}\sum_{i=1}^{N-m}B_i^m(r)$

(5) A_i is the number of the $j\,(1\le j\le N-m,\ j\ne i)$ when $d[X_{m+1}(i),X_{m+1}(j)]\le r$. Thus $A_i^m(r)$ is defined as follows: $A_i^m(r)=\dfrac{1}{N-m-1}A_i$

(6) $A^m(r)$ is defined as follows: $A^m(r)=\dfrac{1}{N-m}\sum_{i=1}^{N-m}A_i^m(r)$

Finally, the sample entropy of the original data $x(1),x(2),\cdots x(N)$ is defined as follows:

$$SampEn(m,r)=\lim_{N\to\infty}\{-\ln[\frac{A^m(r)}{B^m(r)}]\}$$

The value of sample entropy is associated with the value of m and r. According to the findings of the literature [16], the calculated sample entropy has reasonable statistical property when $m=1$ or $m=2$, $r=0.1\mathrm{Std}\sim0.25\mathrm{Std}$ (Std is the standard deviation of the original data $x(i)$, $i=1,2,\cdots,N$). In the present study, $m=2$, $r=0.2\mathrm{Std}$.

4. SVM Multiple Faults Classifier and Optimization

SVM is a supervised learning method. On the one hand, it has its superiority in solving the small sample size problem, nonlinearity and high dimensional pattern recognition. On the other hand, it can be used in function fitting and other machine learning problems. In comparison to neural network, SVM has a lot of advantages. For instance, first, it has good generation ability. Second, it can solve the gender problem of computational complexity input vector. Moreover, SVM can finally translate into convex optimization problems which guarantee global optimality of the algorithm, avoiding the smallest problems which cannot be solved by the neural network. Thus, scholars pay more and more attention to the SVM method in the application of fault diagnosis.

SVM is a new learning machine designed for binary classification problems based on statistical learning theory and structural risk minimization principle. Need to adopt a multi-class SVM approach for multi-classification problems. The existing multi-class SVM methods include one-to-many, one-to-one, directed acyclic graph, error correction coding, decision tree and so on. In this paper, the one-to-one multi-taxonomy is used, which is suitable for online fault diagnosis. The performance characteristics of one-to-one compared with other classification methods are as follows [18, 19]: The training time of the one-to-one method is almost the same as that of the directed acyclic graph, error correction coding and decision tree, but shorter than that of one-to-many method. The one-to-one method does not require certainty topology design, thus the recognition rate of the one-to-one method is the highest among methods described above.

Studies [20-22] have found that selecting a different kernel function has little effect on the promotion performance of support vector machine. Instead, the choice of error penalty factor c and kernel function parameter σ is the key factor. Only selecting a suitable kernel function parameter can the data of the sample be projected onto the appropriate feature subspace to get the better performance SVM model. Therefore, the most important thing for SVM optimization is to optimize its parameters which include error penalty factor, insensitive parameters and kernel function parameters, and so on.

The particle swarm optimization algorithm works well on solving highly nonlinear problems with relatively short optimization time and good optimization effects. This article uses a particle swarm optimization algorithm to optimize the value of the error penalty factor and the kernel function parameter of SVM. And the basic steps are as follows:

(1) Divide the training dataset of the samples into k free subsets according to k cross-validation requirements.

(2) Contract a value range of optimization parameter c and σ, and encode the position vector of each particle with two-dimensional coding to create an initial particle swarm.

(3) Select training set for each particle corresponds to the parameters for cross-validation, and choose the accuracy rate of the forecast model as the objective function value corresponds to the particle.

4) Iterate the particles in the particle swarm as follows.

$$v_i^{k+1} = \omega v_i^k + c_1 r_1 \left(p_i^k - x_i^k\right) + c_2 r_2 \left(p_g^k - x_i^k\right), i = 1,2,\cdots$$

$$x_i^{k+1} = x_i^k + v_i^{k+1}, i = 1,2,\cdots$$

Where, v_i^{k+1} is the flight speed of the i particle in the $k+1$ generation, x_i^{k+1} is the flight speed of the i particle in the $k+1$ generation, p_i^k is the position of the i particle from the first generation to the k generation, p_g^k is the best position of particles from the first generation to the k generation, $p_i^k - x_i^k$ is the individual cognition, $p_g^k - x_i^k$ is the social cognition, ω is the inertia factor expressed the degree that one trust oneself, c_1、c_2 is normal number known as the acceleration factor, r_1、r_2 is a random number of (0,1).

(5) Evaluate all particles with the objective function values. And when the currently assessing value is superior to its historical evaluation, treat it as the best historical evaluation to inform the current optimal position vector of the particle.

(6) Seek the global optimal solutions, and update it if its value is better than the current history. The search is stopped and the optimal value of c and σ is outputted when the termination criteria is reached, otherwise go back and search again.

5. BP Neural Network and Optimization

The BP neural network is a kind of multilayer feedforward neural network based on the error back pass algorithm, constituted by an input layer, one or more hidden layers and an output layer. The neurons between layers are connected by the connection weights. The input-output relationship of the BP network is represented as follows using the s-type activation function:

Input: $net = x_1 w_1 + x_2 w_2 + \cdots + x_n w_n$

Output: $y = f(net) = \dfrac{1}{1+e^{-net}}$

Error function: $e = \dfrac{1}{2}\sum_{o=1}^{q}\left(d_o(k) - yo_o(k)\right)^2$

Where, f is the activation function, k is the total number of samples.

The learning process of BP neural network is that network connection weights are changed constantly with the stimulation of external input samples in order to make the output of the network constantly close to the desired output. BP algorithm is the most widely used algorithms in artificial neural network, but BP neural network has some deficiencies. For instance, its convergence rate is slow, it cannot guarantee convergence to a global minimum point; its network structure is difficult to determine. Therefore, this paper puts forward a way to improve the standard BP algorithm by increasing the momentum term and the adaptively adjust learning rate [23].

(1) Increase the momentum term. When the network amends its weights with the momentum term, the network considers the gradient descending direction adjustment of the moment t and the gradient descending direction adjustment of the moments before the moment t, which can effectively suppress the network into local minima by reducing the oscillation trend and the error surface sensitivity of the local details of the network.

Therefore, the weight adjustment amount of the last time is partly superimposed to the weight adjustment amount of this time. α is momentum coefficient, $\alpha \in (0,1)$. If W is a matrix of a layer weight, X is a layer input vector, then the adjustment amount of weight appended a momentum term is expressed as:

$$\Delta W(t) = \eta\delta X + \alpha \Delta W(t-1)$$

Where, δ is the error signal of the neural elements in the current layer, η is the adaptive learning rate.

(2) Learning rate controls the changing size of each step along the gradient direction corresponding to the weights in the weight space. The higher the learning rate is, the greater the amount of weight adjustment is. That is, oscillation may occur when the learning rate is too high. Therefore, the learning rate should be adaptively adjusted during the training of the network, which can improve training speed and shorten the learning time on the basis of the network stability.

We can get the network weights adjustment amount of the moment $t+1$ combined with the two methods described above:

$$\Delta W(t+1) = \begin{cases} \beta\eta\delta X + \alpha \Delta W(t) & E_{RMS} \uparrow \\ \eta\delta X + \alpha \Delta W(t) & E_{RMS} \text{ unchanged} \\ \mu\eta\delta X + \alpha \Delta W(t) & E_{RMS} \downarrow \end{cases}$$

Where, E_{RMS} is the total error of the network at the moment t.

If E_{RMS} increases after weight adjustment, the adjustment is invalid, $\eta(t+1) = \beta\eta(t), \beta < 1$. Otherwise, $\eta(t+1) = \mu\eta(t), \mu > 1$. If E_{RMS} remains unchanged, the learning rate will not change.

6. Fan Machinery Fault Diagnosis

6.1. Experimental Simulation of Fan Machinery Faults

Different mechanical vibration experiments were done on the 4-73 No.8D centrifugal fan test bench of North China Electric Power University. This type of fan is one of 4-73 series which is widely used in domestic power plant, and this series fan is usually very easy to meet the similarity law in the operation, thus the fan mechanical fault characteristic simulation has certain representativeness and instructional significance. The experiment selected the Schenck Germany IN-81 eddy current displacement probe and the PCB's Model 481 signal filter. Five eddy current sensors was mounted on both sides of the fan bearing, which non-contact measured the horizontal direction, vertical direction and axial vibration displacement signal of the bearing. The distribution of the mechanical vibration measuring points is shown in Figure 1. Using the non-contact measurement, the eddy current sensors were fixed on a specially machined vibration sensor holder through the threaded connection. The sensor mounting bracket was fixed on the basic rail with the bolt. This paper simulates 13 different running states of the fan, which are shown in Table 1. And the fan fault test system was used to collect the vibration signals in different running states, thus a sample set of the 4-73No.8D vibration signal was formed. Each measuring point measured 20 vibration signal samples in each state, that is, a total of $13 \times 5 \times 20$ vibration signal samples were collected. The sampling frequency is 1600Hz and the speed of the fan is 1200rpm.

Table 1. The Fan Running State

The category label of of each fan running state	The fan running state of the simulation	commentary
1	Normal	
2	Rotor unbalanced (1)	The rotor unbalance fault was simulated by installing the unbalanced mass on the six uniform angle positions close to the outer edge of the front disc. Different severity and different positions of the rotor unbalanced fault were simulated by adjusting the quality and location of the unbalanced mass.
3	Rotor unbalanced (2)	
4	Rotor unbalanced (3)	
5	Rotor unbalanced (4)	
6	Angular misalignment	Coupling of the bench is the rigid coupling. The angular misalignment, parallel misalignment (light) and parallel misalignment (serious) of the coupling were simulated.
7	Parallel misalignment (slight)	
8	Parallel misalignment (serious)	
9	Bearing loose (all loose slight)	Different severity and different positions of the bearing loose were simulated by adjusting the tightness of the bolts on the different parts of the bearing.
10	Bearing loose (all loose serious)	
11	Bearing loose (the left loose)	
12	Static and kinetic friction (1)	The varying degrees of the static and kinetic friction of the single point, multi-point and local face were simulated.
13	Static and kinetic friction (2)	

6.2. Extraction of Fan Vibration Signal Wavelet Packet Energy Eigenvector

Getting the signal sample set of five test points under 13 different running states, all the vibration signals were denoised by using wavelet packet. And the decomposition and reconstruction of 5-layer wavelet packet was applied to collect the normalized frequency band energy feature vector of each vibration signal. Taking a vibration signal sample under the angular misalignment condition for example, the effect of denoising is shown in Figure 2-a and Figure 2-b, and the normalized energy distribution of 32 frequency bands is shown in Figure 2-c. Obviously, the method of wavelet de-noising applied in this paper not only has good effect but also can keep the features of the vibration signal. Moreover, the frequency band energy ratio of higher frequency is nearing zero. For the purpose of avoiding vibration signal feature vector being high dimension, considering the normalized energy distributions of every band after the wavelet packet denoising, decomposition and reconstruction, $F_i = [E_{i0}, E_{i1}, E_{i2}, E_{i3}, E_{i4}]$ is used to express the vibration signal eigenvector of the i test point. E_{i0}, E_{i1}, E_{i2} and E_{i4} each means the normalized energy corresponding to the frequency band (f) for 1-4 times of the vibration signal of the i test point that being done with 5-layer decompositions and reconstruction. But E_{i5} means the sum normalized energy of all the frequency band that $\geq 5f$. During the experiment, 5 eddy current sensors were used to measure the vibration signal of the bearing on different positions and directions under different running state. Therefore, a comprehensive analysis of the vibration signal for 5 test points with the same time is needed to build the wavelet packet energy feature vector F for each state.

$$F = [F_1, F_2, F_3, F_4, F_5] = [E_{10}, \cdots, E_{14}, E_{20}, \cdots, E_{24}, E_{30}, \cdots, E_{34}, E_{40}, \cdots, E_{44}, E_{50}, \cdots E_{54}]$$

Based on the actual data of vibration, we make use of 5-layer wavelet packet denoising and wavelet packet decomposition and reconstruction for $13 \times 5 \times 20$ samples to get the wavelet packet energy feature vector F which is the syncretic result of five test points under various conditions. Then, wavelet packet energy feature vectors of 260 samples in 13 different conditions were got. But here only display one wavelet packet energy feature vector for each sample in each condition, which are shown in Table 2. Due to the high dimension of the energy feature vector F , this paper uses support vector machine to reduce it. Detailed comparison of fan fault diagnosis results with fault energy eigenvector F_i or F are shown in the table above.

Table 2. The Normalized Wavelet Packet Energy Eigenvector of a Single Sample of Each Fan Running State

Test poin-ts	Frequ-ency band	The category label of of each fan running state												
		1	2	3	4	5	6	7	8	9	10	11	12	13
1	1f	0.3056	0.0952	0.0196	0.0268	0.0481	0.0695	0.1638	0.1448	0.0285	0.0514	0.0898	0.0472	0.1993
	2f	0.5843	0.6079	0.8137	0.8349	0.7954	0.6246	0.4683	0.6568	0.471	0.5366	0.3736	0.7679	0.6797
	3f	0.0573	0.0975	0.0252	0.005	0.0276	0.1721	0.1742	0.0855	0.2347	0.207	0.2896	0.0388	0.0459
	4f	0.0491	0.1966	0.1404	0.1327	0.128	0.1333	0.1888	0.1109	0.2335	0.1883	0.2228	0.1361	0.0725
	>=5f	0.0037	0.0028	0.0012	0.0007	0.0009	0.0005	0.0049	0.0019	0.0323	0.0167	0.0242	0.01	0.0027
2	1f	0.0515	0.0675	0.0598	0.0348	0.0364	0.1386	0.0486	0.071	0.0877	0.0653	0.0343	0.0356	0.0483
	2f	0.8491	0.7641	0.79	0.786	0.8066	0.7597	0.7954	0.8218	0.7489	0.8096	0.8034	0.8221	0.8209
	3f	0.0088	0.068	0.0135	0.0078	0.0218	0.029	0.0621	0.0169	0.03	0.0162	0.0247	0.0084	0.0164
	4f	0.0892	0.0985	0.1362	0.1705	0.134	0.0693	0.0916	0.0898	0.1291	0.1071	0.1284	0.1198	0.1094
	>=5f	0.0014	0.0019	0.0005	0.0008	0.0012	0.0034	0.0023	0.0004	0.0043	0.0017	0.0091	0.0141	0.0051
3	1f	0.2127	0.2051	0.9583	0.6305	0.1666	0.0214	0.0175	0.0141	0.0744	0.1368	0.0819	0.3794	0.1087
	2f	0.6622	0.7089	0.0106	0.1342	0.6774	0.8294	0.8335	0.8283	0.7475	0.7006	0.7691	0.3213	0.658
	3f	0.0348	0.0391	0.0027	0.1195	0.0712	0.0105	0.0199	0.0086	0.0515	0.0571	0.0495	0.0782	0.1249
	4f	0.0869	0.0357	0.0043	0.109	0.0841	0.1381	0.1278	0.1474	0.1256	0.1036	0.097	0.1274	0.106
	>=5f	0.0034	0.0112	0.0241	0.0067	0.0007	0.0007	0.0012	0.0016	0.001	0.0018	0.0025	0.0937	0.0023
4	1f	0.0317	0.0501	0.0121	0.0172	0.1972	0.0612	0.0538	0.072	0.0273	0.0166	0.0148	0.041	0.0655
	2f	0.7661	0.7688	0.7955	0.6826	0.2226	0.7836	0.7759	0.774	0.8248	0.8151	0.8301	0.6104	0.7599
	3f	0.0221	0.0462	0.041	0.1205	0.2715	0.0077	0.0111	0.0085	0.0066	0.0093	0.0067	0.0333	0.0278
	4f	0.1758	0.1317	0.1449	0.1632	0.1519	0.1474	0.1584	0.1454	0.1396	0.157	0.1441	0.1051	0.1457
	>=5f	0.0044	0.0032	0.0064	0.0165	0.1568	0.0001	0.0007	0.0001	0.0017	0.002	0.0043	0.2102	0.0011
5	1f	0.0384	0.0149	0.0812	0.0685	0.3134	0.0194	0.0776	0.0319	0.0522	0.0404	0.0213	0.3131	0.0551
	2f	0.8186	0.8251	0.7392	0.7552	0.2137	0.758	0.7753	0.8315	0.763	0.777	0.7595	0.0704	0.7833
	3f	0.0216	0.0194	0.042	0.0424	0.0225	0.0582	0.016	0.007	0.0255	0.0442	0.0556	0.1075	0.0217
	4f	0.1128	0.1365	0.0534	0.076	0.0156	0.1586	0.1309	0.129	0.1289	0.1349	0.1344	0.0839	0.1212
	>=5f	0.0085	0.0041	0.0842	0.0578	0.4348	0.0057	0.0003	0.0006	0.0304	0.0036	0.0292	0.425	0.0186

6.3. Extraction of the Sample Entropy Feature Vector

During the experiment $13\times 5\times 20$ vibration samples were collected. The number of sampling points is 512. Each vibration signal can be expressed as $x(1), x(2), \cdots, x(512)$. The sample entropy of each signal was extracted to build the sample entropy feature vector of each fan running state. And the sample entropy feature vector is shown as follow:

$$SE = [SamEn_1, SamEn_2, \cdots SamEn_5]$$

Each of $SamEn_1, SamEn_2, \cdots SamEn_5$ means the sample entropy of the corresponding vibration singles from five test points under each state in the same time. Finally, the sample entropy feature vectors of 260 fault samples under 13 different states were extracted. But here only one sample entropy feature vector for each single sample in each state are shown as Table 3. Where ‘1’ represents the state in which the fan is running.

Table 3. The Sample Entropy Eigenvector of a Single Sample of Each Fan Running State

The category label of each fan running state	The input value of BP neural network input layer nodes Sample entropy of each measuring point 1	2	3	4	5	The target output of BP neural network output layer nodes
1	1.798	1.325	1.695	0.407	0.997	1000000000000
2	1.529	1.436	1.835	0.784	0.725	0100000000000
3	0.895	0.886	2.113	0.614	1.484	0010000000000
4	0.704	0.624	1.948	0.854	1.512	0001000000000
5	0.703	0.783	1.762	1.634	1.803	0000100000000
6	1.479	1.356	0.615	0.286	0.421	0000010000000
7	1.377	1.412	0.797	0.320	0.649	0000001000000
8	1.561	1.357	0.657	0.302	0.743	0000000100000
9	1.191	0.795	1.295	0.370	0.785	0000000010000
10	1.067	0.975	1.473	0.427	0.749	0000000001000
11	1.268	0.843	1.665	0.533	0.888	0000000000100
12	0.777	0.939	1.488	0.800	1.225	0000000000010
13	1.135	1.017	1.176	0.554	0.952	0000000000001

6.4. Fault Diagnosis based on Wavelet Packet Energy Eigenvector

In the condition of fault feature vector F_i or F , a SVM multi-class classifier was used for fan fault diagnosis after getting the wavelet packet energy feature vector F_i of the vibration single in 13 different fan running states. And diagnostic steps are as follows:

(1) 130 groups of the wavelet packet energy feature vectors (F_i or F) of 13 different states composed of 10 groups of each state were taken as the training samples. The remaining 130 were taken as test samples.

(2) The training samples were chosen as input vectors for the improved SVM multi-class classifier to train, and then the fan machinery fault diagnosis model was established. If the test sample is imported to the model, the diagnostic results will be obtained.

6.4.1. Fault Diagnosis based on Wavelet Packet Energy Feature Vectors of Single Measurement Point

The accuracy rate of the fan mechanical failure diagnosis is 65.38%,43.85%,53.85%, 54.62% and 69.23% taken $F_1, F_2, \cdots F_5$ as the feature state of the running states. The diagnostic results are shown in Figure 1. The numbers on the category label of the above figure correspond to 13 running states (1-normal, 2-unbalanced (1), 3-unbalanced (2), 4-unbalanced (3), 5-unbalanced (4), 6-angular misalignment, 7-parallel misalignment (light), 8-parallel misalignment (serious), 9-bearing loose (all loose light), 10- bearing loose (all loose serious), 11-bearing loose (the left loose), 12-static and kinetic (1) and 13-static and kinetic (1)).

Obviously, if a single measurement point is adopted, there will be a high probability of false alarm, false negatives and confusion. In other words, this method will lead to low fault recognition rate.

6.4.2. Fault Diagnosis based on Wavelet Packet Energy Feature Vectors of Multiple Measurement Points

The accuracy rate of the fan mechanical failure diagnosis is 94.6% taken F as the feature state of the running states. The diagnostic results are shown in Figure 2. It can be seen that there is no false alarm, false negatives and confusion with fault diagnosis based on wavelet packet energy feature vectors of multiple points. Moreover, the recognition and accuracy of the category, severity and location of failure are high. And there are seven samples misdiagnosed, including, two samples of parallel misalignment (light) state are misdiagnosed as parallel misalignment (serious) state; two of bearings looseness (all loose light) state are misdiagnosed as the bearing looseness (all loose serious) state; one of bearings looseness (all loose light) state are misdiagnosed as the bearing looseness (the left loose) and two of bearings looseness (the left loose) state are misdiagnosed as the bearing looseness (all loose light) state. The reasons result in a small probability of misclassification are that the feature vector F of different severity vibration signal in parallel misalignment fault are similar, and the feature vector F of different fault location and severity vibration signal in bearing looseness fault are similar.

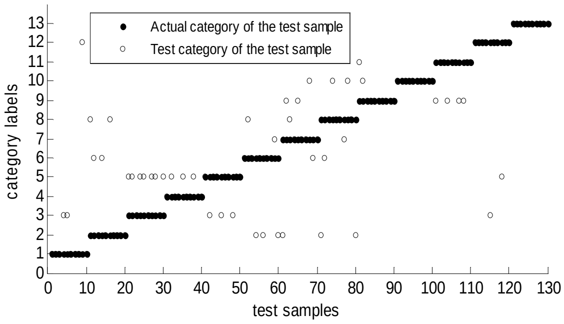

Figure 1. Fault Diagnosis Chart based on the Analysis of Single Measuring Point Vibration Signal

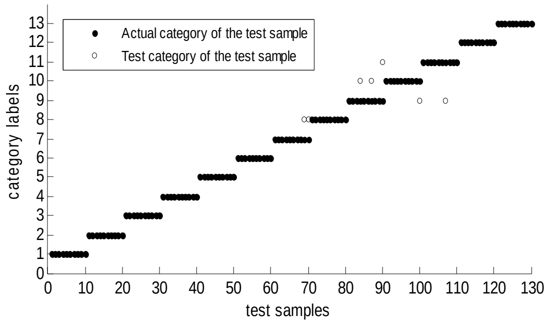

Figure 2. Fault Diagnosis Results Chart based on the Comprehensive Analysis of Five Measuring Point Vibration Signal

Table 4. Output Value of Improved BP Neural Network

Sample	The output value of each node of the network output layer												
No	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.9968	0.0066	0.0010	0.0000	0.0000	0.0000	0.0001	0.0002	0.0000	0.0006	0.0003	0.0000	0.0037
2	0.0101	0.9947	0.0013	0.0000	0.0015	0.0002	0.0025	0.0005	0.0000	0.0000	0.0005	0.0006	0.0019
3	0.0013	0.0000	0.9980	0.0006	0.0000	0.0000	0.0007	0.0000	0.0000	0.0017	0.0054	0.0005	0.0001
4	0.0020	0.0000	0.0012	0.9985	0.0061	0.0000	0.0001	0.0000	0.0014	0.0001	0.0003	0.0001	0.0009
5	0.0000	0.0023	0.0000	0.0035	0.9961	0.0000	0.0040	0.0000	0.0000	0.0000	0.0005	0.0039	0.0027
6	0.0000	0.0020	0.0000	0.0000	0.0001	0.9976	0.0008	0.0013	0.0006	0.0037	0.0000	0.0001	0.0004
7	0.0000	0.0007	0.0000	0.0000	0.0005	0.0003	0.9993	0.0015	0.0000	0.0001	0.0000	0.0004	0.0006
8	0.0001	0.0008	0.0001	0.0000	0.0004	0.0005	0.0013	0.9447	0.0000	0.0008	0.0000	0.0002	0.0065
9	0.0004	0.0000	0.0005	0.0016	0.0000	0.0005	0.0020	0.0020	0.9984	0.0014	0.0000	0.0000	0.0069
10	0.0002	0.0000	0.0013	0.0001	0.0000	0.0023	0.0003	0.0005	0.0467	0.9959	0.0021	0.0002	0.0000
11	0.0002	0.0000	0.0020	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	1.0000	0.0016	0.0000
12	0.0000	0.0004	0.0024	0.0001	0.0009	0.0000	0.0003	0.0000	0.0000	0.0000	0.0002	0.9994	0.0003
13	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0001	0.0000	0.0007	0.9970

6.5. Fan Machinery Fault Diagnosis based on the Sample Entropy Feature Vector

The improved BP neural network was utilize, whose neuron number of the input layer, hidden layer and output layer were 5, 15, and 13. Through training and testing, the results show that the diagnostic accuracy of the network is the highest when the transfer function of the input layer neurons and the hidden layer neurons, the hidden layer neurons and the output layer neurons are tansig(), logsig(). The neural network training method with the momentum term and the adaptive learning rate was utilized based on the supervised learning. The expected error minimum value of the network was le-5. The maximum number of iteration was 2000. The learning rate η was 0.01. The Momentum factor was 0.90.

The sample entropy feature vectors of 260 fault samples under 13 different states were extracted. 130 sample entropy feature vectors of 13 different states composed of 10 groups of each state were taken as the training samples. If the test samples of a network are the input values of BP neural network input layer nodes shown in Table 3, the output values of the improved BP neural network are shown in Table 4. Comparing with the target output of BP neural network output layer nodes shown in Table 3, the accuracy rate of fault diagnosis is 100%. If taking the remaining 130 as test samples of the improved BP neural network, the accuracy rate of fault diagnosis is 99.23%. On the other hand, if taking the remaining 130 as test samples of the standard BP neural network, the accuracy rate of fault diagnosis is 92.31%.

The results show that the improved BP neural network can reduce training times, improve the efficiency of learning and effectively suppress the network caught in local minima, thus, the improved BP neural network significantly improve the diagnostic accuracy of fan mechanical failures.

7. Conclusions

(1) In this paper, thirteen different fan running states were simulated and plenty of vibration signals were obtained through tests. Then utilizing the wavelet packet analysis method, the vibration signals could be denoised, decomposed and reconstructed, and the various band normalized energy of wavelet packet of the each vibration signal was calculated to make the feature vector F of 260 groups of fan running states as the sample set of fault diagnosis. After that, the sample entropy of each vibration signal was extracted so that the feature vector SE of 260 groups of fan running states was gotten as the sample set of fault diagnosis.

(2) Through the multi-fault classifier of the improved SVM, the sample sets of single point wavelet packet energy feature vector F_i and multi-point feature vector F were classified. The accuracy rates of the former are 65.38%、43.85%、53.85%、54.62% and 69.23%, which is apparently lower than the latter's 94.6%. In other words, the wavelet packet energy feature vector F can reach a good characterization for the various fan running states. The fault diagnosis method based on the wavelet packet analysis and the improved SVM has the high recognition ability for the category, severity and location of the fan mechanical failures.

(3) The fault diagnosis method based on the feature vector SE of the sample entropy and improved BP neural network is put forward, and its accuracy rate can reach up to 99.23%. Thus, the sample entropy of the vibration signal is made as the characteristics of the fan fault diagnosis, which has higher recognition ability for the category, severity and location of the fan mechanical failures than the wavelet packet energy feature. Furthermore, the recognition precision of this method applied in the fault diagnosis of the fan exceeds the method based on the wavelet packet analysis and the improved SVM.

Acknowledgements

This work is supported by the Fundamental Research Funds for the Central Universities (12MS116).

References

[1] Niu Peifeng, Zhang Jun. Application of the Wavelet Transform Technique in Turbo-generator Fault Diagnosis. *Journal of Chinese Society of Power Engineering.* 2007; 27(1): 76-80.
[2] Li K, Chen P, Wang H Q. Intelligent Diagnosis Method for Rotating Machinery Using Wavelet Transform and Ant Colony Optimization. *Sensors Journal.* 2012; 12(7): 2474-2484.
[3] Wang Wenyi. Early detection of gear tooth cracking using the resonance demodulation technique. *Mechanical Systems and Signal Processing.* 2001; 15(5): 887-903.
[4] Nikolaou NG, Antoniadis IA. Rolling element bearing fault diagnosis using wavelet packets. *NDT & E International.* 2002; 35(3): 197-205.
[5] Liang Ping, Bai Lei, LONG Xinfeng, et, al. Turbine rotor vibration faults diagnosis based on wavelet packet analysis and neural network. *Control Theory & Applications.* 2007; 24(6): 981-985.
[6] Han Lei, Hong Jie, Wang Dong. Fault diagnosis of aero-engine bearings based on wavelet package analysis. *Journal of Propulsion Technology*. 2009; 30(3): 328-332.
[7] Peng Wenji, Luo Xingqi. *Research on Vibrant Fault Diagnosis of Hydro-turbine Generating Unit Based on Wavelet Packet Analysis and Support Vector Machine*. Proceedings of the CSEE. 2006; 26(24): 164-168.
[8] Yan Ruqiang, Gao RX. Approximate entropy as a diagnostic tool for machine health monitoring. *Mechanical Systems and Signal Processing.* 2007; 21(2): 824-839.
[9] Zhao Zhihong, Yang Shaopu. Roller Bearing Fault Diagnosis Based on Wavelet Packet Transform and Sample Entropy. *Journal of Vibration, Measurement & Diagnosis*. 2012; 32(4): 640-644.
[10] Wang Jinwei. Research on the Improved Algorithm of the BP Artificial Neural Network. *China Water Transport.* 2008; 8(01): 157-158.
[11] Zheng Ruirui, Zhao Jiyin, Zhao Tingting, etal. *Power Transformer Fault Diagnosis Based on Genetic*

Support Vector Machine and Gray Artificial Immune Algorithm. Proceedings of the CSEE. 2011; 31(7): 56-63.
[12] Bai Weidong, Yan Jianhua, Ma Zengyi, etal. Method of Flame Identification Based on Support Vector Machine. *Journal of Chinese Society of Power Engineering*. 2004; 24(4): 548-551.
[13] Liu Yongbin, He Qingbo, Kong Fanrang, etal. Fault Diagnosis of Internal Combustion Engine Using PCA and SVM. *Journal of Vibration, Measurement & Diagnosis*. 2012; 32(2): 250-255.
[14] Yang Honghui, Chen Zhaoji, Dai Jian. Fault Diagnosis for Fan Based on Self-Adaptive Boosting SVM Ensemble. *Measurement & Control Technology.* 2010; 29(7): 72-74.
[15] Alcaraz R, Rieta JJ. A review on sample entropy applications for the non-invasive analysis of atrial fibrillation electrocardiograms. *Biomedical Signal Processing and Control*. 2010; 5(1): 1-14.
[16] Pincus SM. Assessing serial irregularity and its implications for health. *Annals of the New York Academy of Sciences.* 2002; 954: 245-67.
[17] JAK Suykens, J Vandewalle. Least squares support vector machine classifiers. *Neural Processing Letters*. 1999; 9(3): 293-300.
[18] Hao Zhifeng, LIu Bo, YANG Xiaowei. *A comparision of multiclass support vector machine algorithms*. Proceedings of the International Conference on Machine Learning and Cybernetics. 2006:4221-4226.
[19] Yan Zhigang, DU Peijun. Generalization performance analysis of M-SVMs. *Journal of Data Acquisition & Processing*. 2009; 24(4): 469-475.
[20] Shi Zhibiao, SONG Quangang, MA Mingzhao, etal. Fault Diagnosis of Steam Turbine Based on MPSO-SVM Algorithm. *Journal of Chinese Society of Power Engineering.* 2012; 32(6): 454-457.
[21] Wu Ming-sheng. Parameter Optimization Method Research and Application of RBF Neural Network and SVM. Central south university. 2007 (in Chinese).
[22] Zhang Nan. Study on the Parameter Optimization in Support Vector Machine. Northwest university, 2008.
[23] Huang Yong, GUO Xiaoping. Fault Diagnosis in Diesel Engine Based on Improved BP Artificial Neural Network. *Auto Mobile Science & Technology*. 2009; (2): 55-58.
[24] Yongxin Feng, Tao Yang, Wenguang Yang, Dongxiang Jiang. Study of Fault Diagnosis Method for Wind Turbine With Decision Classification Algorithms and Expert System. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2012; 10(5): 905-910.
[25] Jianzhao Huang, Jian Xie, Hongcai Li, Gui Tian, Xiaobo Chen. Self-adaptive Decomposition Level De-noising Method Based on Wavelet Transform. *TELKOMNIKA Indonesian journal of Electrical Engineering*. 2012; 10(5): 1015-1020.