Fault Diagnosis of Motor Bearings based on BP Neural Network and PSO

Mei Yuan

School of Automation, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

yuanmei20@126.com

Abstract

A motor bearing fault diagnosis method based on BP neural network (BP) and particle swarm optimization (PSO) is proposed in this paper. The method uses PSO to optimize the weights and thresholds of neural network. Which improves the convergence rate, generalization performance and recognition accuracy of BP neural network. In terms of feature extraction, this paper combines the time-domain and wavelet packet energy characteristics of vibration signals so that the characteristics of the vibration signal has a good reliability and sensitivity. Experiments show that the fault diagnosis method have a good recognition effect on the outer ring fault, inner ring fault and ball fault of motor bearing, and has strong practicability.

Keywords

BP neural network; Particle swarm optimization; Wavelet packet analysis; Motor bearing; Fault diagnosis.

1. Introduction

As an important part of the motor, bearing in the event of an accident, it will endanger the vehicle traffic safety, and even cause serious traffic accidents. In order to reduce or eliminate accidents caused by bearing fault, it is very important to study the bearing fault diagnosis method.

The diagnosis of motor bearing mainly includes the vibration signal acquisition, fault feature extraction and feature recognition, among them, the fault feature extraction and feature recognition are the key. Fault signal feature extraction method is commonly used Local mean decomposition [1], Empirical mode decomposition [2] [3] and so on, but they all have the problem that the endpoint effect distorts the results. Fault diagnosis can be regarded as the process of fault pattern recognition, and designing a suitable classifier for fault pattern recognition is a critical step in fault diagnosis. The traditional pattern recognition methods often use Bayesian[4] criterion and logistic regression algorithm[5] to realize different classifiers, but these methods have their own limitations, For example, Bayesian theoretically solves the problem of optimal classifier, But it needs to solve the more difficult probability density estimation problem.

BP neural network[6] is an algorithm commonly used in pattern recognition, implements a mapping function from the input to the output, it has proved that it has the ability to realize any complex nonlinear mapping in mathematical theory, and has strong generalization ability, self-learning and adaptive ability. But there is a problem that the convergence rate is slow, easy to fall into the local minimum value, and the connection weight and the threshold value have a great influence on the network training .Therefore, many scholars combine the BP neural network with other algorithms to improve the recognition rate.

In the traditional motor rolling bearing diagnosis method, only the bearing fault of different fault types with the same severity is studied, and the extracted fault feature can't characterize the severity, resulting in a low recognition rate in a particular application. This paper studies the time-domain features and wavelet packet feature of fault signal, and presents a BP neural network motor bearing fault diagnosis method based on PSO algorithm. PSO algorithm is used to optimize the weights and thresholds of BP neural network, which speeds up the convergence and improves the fault diagnosis rate.

2. The Feature Extraction of Vibration Signal Section Headings

When the bearing fault occurs, some characteristic parameters of the vibration signal will change with the nature and size of the fault, these features can be used as the basis for fault diagnosis .There are many parameters that can be used to characterize the bearing state, Common parameters include peak factor, pulse factor, margin factor, waveform factor, peak and so on. Different types of parameters have corresponding sensitive fault types. For example, For example, the peak is more sensitive to surface failure, the root mean square is susceptible to wear fault, and skewness is more sensitive to impact fault. In this paper, considering the advantages of vibration signal in time-domain and wavelet packet energy, The time-domain feature and the energy feature based on wavelet packet decomposition ,so that the eigenvector characterizing the vibration signal has good reliability and sensitivity.

2.1 Time-Domain Feature Extraction

The parameters that can be used to characterize the fault characteristics are limited, in the selection should follow the high sensitivity, high reliability and can achieve the principle. Because the bearing vibration signal is aperiodic random signal, can't express into a specific mathematical model, but in a large number of repeated tests will appear a statistical law, get its important features. There will be a certain statistical regularity, and its important characteristics. In view of the fact that the time-domain statistical characteristic parameters are more sensitive to the load size and speed and depend on historical data, In this paper, five dimensionless parameters with sensitive sensor bearing defects are selected as the time-domain feature elements, respectively, kurtosis factor, margin factor, peak factor, pulse factor, waveform factor. They are not affected by the absolute level of the signal without considering the relative standard values.

$$\text{Kurtosis factor: } K_r = \sum_{i=1}^{n} x_i^4 \Big/ n x_{rms}^4 \tag{1}$$

Where, $x_{rms} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}$, the kurtosis factor is independent of the speed, load and size of the rolling bearing. When the rolling bearing fails, the kurtosis factor and the peak factor have similar trends.

$$\text{Peak factor: } C = x_{peak} / x_{rms} \tag{2}$$

The peak factor is not affected by the bearing size, speed and load. When the fault occurs, the peak factor can effectively predict the fault of the rolling bearing.

$$\text{Margin factor: } L = x_{peak} / x_d \tag{3}$$

Where, $x_d = \left(\frac{1}{n}\sum_{i=1}^{n}\sqrt{|x_i|}\right)^2$, the margin factor is more sensitive to impulse fault, especially an early failure occurs, the value has changed significantly.

$$\text{Waveform factor: } Q = x_{peak} / |x_m| \tag{4}$$

Where, $x_{peak} = \frac{1}{m}\sum_{j=1}^{m} x_{pj}$, x_{pj} is the peak present in the signal ,where $j = 1, 2 \cdots m$. The waveform factor has a high sensitivity to pitting fault and wear fault of rolling bearings.

$$\text{Pulse factor: } S = x_{rms} / |x_m| \tag{5}$$

Among the five characteristic parameters, the kurtosis factor, the crest factor and the margin factor have strong sensitivity, and the waveform factor has good stability.

2.2 Wavelet Packet Energy Feature Extraction

Since the scale function of the multiresolution analysis varies in binary, so its frequency resolution is poor in the high frequency band. The wavelet packet decomposition[7][8] can divide the frequency band into multiple layers, so as to further decompose the high frequency part which is not subdivided by the multi-resolution analysis, resolution can be improved and the frequency band matching the signal spectrum can be selected adaptively according to the characteristics of the analyzed signal. Based on the above advantages, the wavelet packet analysis is used to extract the wavelet packet energy characteristics of the vibration signal.

Wavelet packet decomposition and reconstruction algorithm is as follows:

From $\left\{d_t^{j+1,n}\right\}$ to solve $\left\{d_t^{j,2n}\right\}$ and $\left\{d_t^{j,2n+1}\right\}$, the wavelet packet decomposition algorithm is shown in formula (6):

$$\begin{cases} d_t^{j,2n} = \sum_k h_{k-2l} d_k^{j+1,n} \\ d_t^{j,2n+1} = \sum_k g_{k-2l} d_k^{j+1,n} \end{cases} \tag{6}$$

From $\left\{d_t^{j,2n}\right\}$ and $\left\{d_t^{j,2n+1}\right\}$ to solve $\left\{d_t^{j+1,n}\right\}$, the wavelet packet reconstruction algorithm as shown in formula (7):

$$d_t^{j+1,n} = \sum_k (h_{l-2k} d_k^{j,2n} + g_{l-2k} d_k^{j,2n+2}) \tag{7}$$

Where, $g_k = (-1)^k h_{1-k}$, belong to the orthogonal relationship.

In this paper, Daubechies4 wavelet packet is used to decompose the vibration signal sampling sequence to three layers of orthogonal wavelet, and eight bands are obtained. let wavelet packet decomposition sequence of energy is $e_j\,(j = 0, 1 \cdots 7)$, $e_j = \sum_{k=1}^{n} \left|x_{jk}\right|^2$, $x_{jk}\,(j = 0, 1 \cdots 7, k = 1, 2 \cdots n)$ is discrete point amplitude of reconstructing signal for each band. Take the energy of the wavelet packet decomposition sequence of each layer as the constituent element of the wavelet packet energy eigenvector, get the wavelet packet energy feature vector $T = (e_0, e_1, \cdots e_7)$.

As shown in Fig 1, when the bearing is normal operation, its vibration signal energy is mainly distributed nodes 0 and 1, while fault occurs, the energy is focused on the nodes 2 and 6, and different fault, the energy proportion in the two frequencies is not the same.

Integrated time-domain and band energy extracted features, the feature vector for bearing fault diagnosis in this paper is set to:

$$T' = (Kr, L, C, S, Q, e_0, e_1, e_2, e_3, e_4, e_5, e_6, e_7) \tag{8}$$

In order to facilitate the subsequent data processing convenience and to speed up the convergence, the extracted features are also normalized.

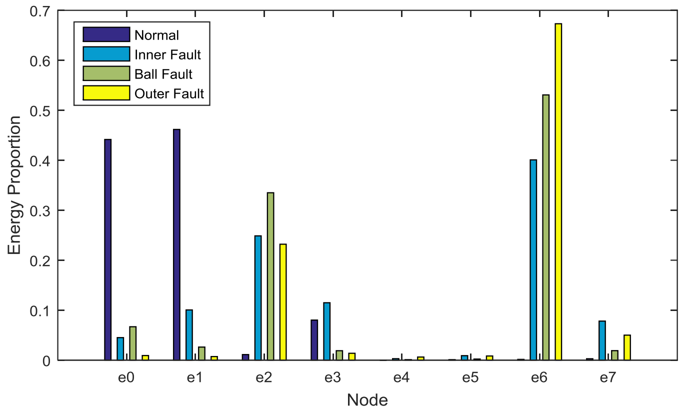

Fig.1 The same severity of the energy distribution of different faults

3. Design of Bearing Fault Detection Model

3.1 BP Neural Network

BP (Back Propagation) neural network[9] is a kind of multi-layer feedforward network with error back propagation, which is widely used, such as classification, function approximation, data compression and so on. Fig.2 is a typical BP neural network model topology diagram, in the BP neural network, each sample contains m input and n output. BP neural network model topology is divided into three layers, The first layer is the input layer, the last layer is the output layer, between the input layer and the output layer is hidden layer. 1989 Robert Hecht-Nielsen proved that a BP network with an implicit layer could approach a continuous function of any closed interval, so a three-layer BP network could achieve dimension-to-dimension mapping.

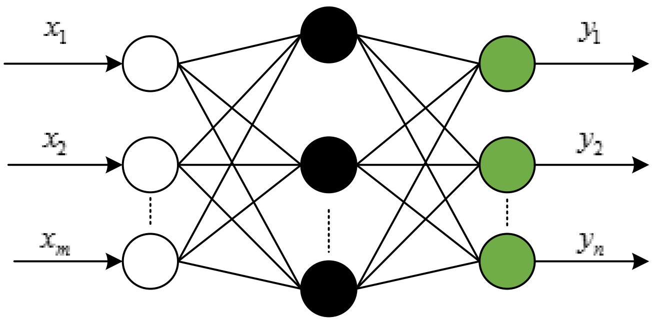

Fig.2 The model of BP neural network

In the BP neural network, the number of input nodes is equal to the number of input samples, and the number of output nodes is equal to the number of output results. Therefore, only the hidden layer node is to be determined. The number of hidden nodes can be determined by the following formula:

$$h=\sqrt{m+n}+a \tag{9}$$

Where m is the number of input nodes, n is the number of output nodes, a usually between 1 and 10, and is an adjustment constant, and h represents the number of hidden layer nodes that need to be determined.

BP neural network includes two processes, one is the forward transfer process, calculate the input and output of the hidden layer neurons and the output layer neurons, the second is the reverse transfer process, modify the weights and thresholds. In the forward transfer process, the output value of each node is determined by the output value of all nodes in the previous layer, the weights of the current node and the previous node, the threshold of the current node, and the activation function. Calculated as follows:

$$x_j = f(\sum_{i=0}^{m-1} w_{ij} x_i + b_j) \tag{10}$$

Where, x_j represent the output value of nodes j, w_{ij} represent the weight between the node i and the node j, x_i represent the output value of the nod i, b_j represent the threshold of the node j, f represent the activation function, the activation function generally selects a s-type function or a linear function. The main purpose of the error signal reverse transfer process is to repeatedly modify the weights and thresholds, its learning rule is to use the gradient descent algorithm, through the reverse propagation to constantly adjust the parameters of each node and mapping weights, so that the network error square sum and minimum.

3.2 Particle Swarm Optimization

Particle Swarm Optimization (PSO) [10] algorithm originated from the study of foraging behavior of birds, often used to solve the optimal solution problem. The realization process is: initialize a group of random particles, find the optimal solution by iteration. In each iteration, the particle updates itself by tracking two "extremes", one for the particle itself to find the optimal solution, called the individual extremum. The other is the optimal solution currently found for the whole population, called the global extremum.

There are m particles in the N-dimensional target search space to form a community. The position of particle i in the N-dimensional search space is represented by the vector $\vec{x}_i = (x_{i1}, x_{i2}, x_{i3} \cdots x_{iN})$ .the "flying" velocity is represented by the vector $\vec{v}_i = (v_{i1}, v_{i2}, v_{i3} \cdots v_{iN})$, $i = 1, 2 \cdots,\ m$ .The particle velocity and position update formula is as follows:

$$v_{in}^{k+1} = \omega v_{in}^{k} + c_1 r_1 (p_{in}^{k} - x_{in}^{k}) + c_2 r_2 (p_{gn}^{k} - x_{in}^{k}) \tag{11}$$

$$x_{in}^{k+1} = x_{in}^{k} + v_{in}^{k} \tag{12}$$

Where $i = 1, 2 \cdots,\ m$, $n = 1, 2 \cdots,\ N$, k is the number of iterations. c_1 , c_2 is the learning factor, is a nonnegative constant. r_1 , r_2 is the random number between [0,1]. w is the inertia weight. $v_{in} \in [-v_{\max}, v_{\max}]$, $v_{\max}$ is a constant, set by the user. p_{in} is the optimal position searched by particle i . p_{gn} is the optimal position searched by the entire particle swarm. The iteration termination condition is generally chosen as the maximum number of iterations or the optimized position of the particle swarm search so far satisfies the adaptation threshold.

3.3 PSO-BP Neural Network Model

BP algorithm is a common neural network training algorithm, but the BP algorithm based on gradient descent depends on the choice of initial weights, the training speed is slow and easy to fall into the local optimal problem, resulting the classification model of training is inconsistent and unpredictable. Particle swarm optimization (PSO), as an evolutionary algorithm, has the advantages of strong ability of global search, high robustness and fast convergence speed. Therefore, the combination of PSO and BP neural network to optimize the threshold and connection weight of neural network can overcome the problem of BP neural network, not only has the advantage of BP neural network, but also can improve the convergence of neural network Speed and learning ability.

The thresholds and network weights needed to be adjusted in the neural network are mapped into the particles of PSO. The dimension of the particle can be expressed as $D = IH + HO + H + O$, where I is the number of output neurons, H is the number of neurons in the hidden layer, O is the number of neuron in the output layer.

Through the competition between the particles and cooperation, constantly updated to optimize these parameters, in order to achieve network training.

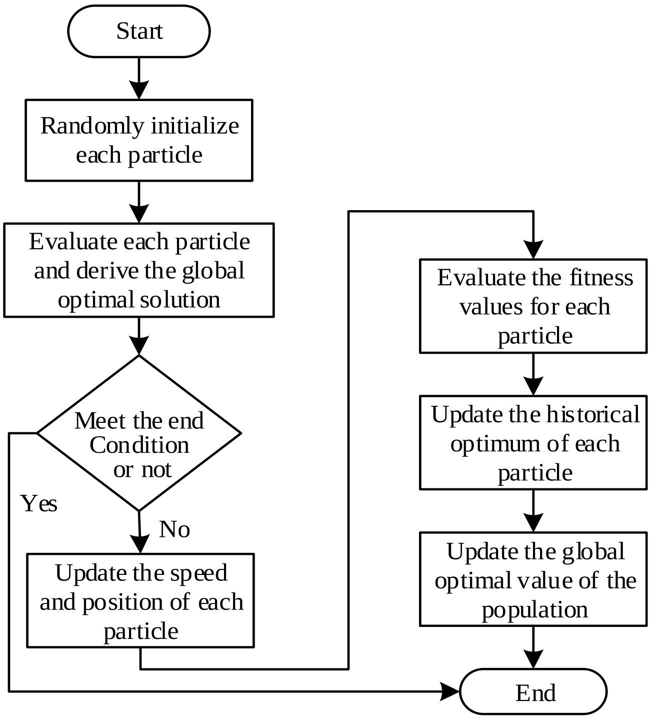

Fig.3 Flow chart of PSO optimization BP parameters

4. experimental simulation analysis

4.1 Training Samples

The data set for fault diagnosis of motor bearings in this paper is derived from the bearing data center of Case Western Reserve University in the United States. Bearing model is 6205-2RS JEM SKF, and the motor is unloaded at a speed of 1797 rpm/s, and use the acceleration sensor to capture the drive side vibration signal at a frequency of 12 kHz. Fault types include inner fault, outer fault and ball fault. Each type of fault consists of three different fault diameters, 0.007 inches, 0.014 inches, and 0.021 inches respectively. In addition, it also includes data under normal operating conditions, so a total of 10 types of data, and each type of data has 240 samples.

4.2 Test Scenarios

In this paper, the fault diagnosis method of motor bearing based on BP neural network and the improved BP neural network fault diagnosis method are compared and analyzed. In order to verify the effectiveness of the proposed method, three scenarios were set up for testing:

1) Without considering the severity of the failure, identify the motor bearing inner fault, outer fault, ball fault and normal conditions, a total of four kinds of situations; the time-domain characteristics of the bearing are extracted and the features are identified by BP neural network.

2) Considering the influence of bearing failure of different severity on the recognition rate, each type of fault is divided into slight faults, medium faults, and serious faults; Identify the motor bearing inner fault, outer fault, ball fault and normal conditions, a total of ten kinds of situations; the time-domain characteristics of the bearing are extracted and the features are identified by BP neural network.

3) Considering the influence of bearing failure of different severity on the recognition rate, each type of fault is divided into slight faults, medium faults, and serious faults; identify the motor bearing inner fault, outer fault, ball fault and normal conditions, a total of ten kinds of situations; take the time-domain features and energy characteristics based on the wavelet, and the BP neural network algorithm based on particle swarm optimization is used to identify the features.

4.3 Results Analysis

Test scenarios one: Take time-domain features, and the hidden layer of the neural network consists of two layers; Only use a 0.07 inch diameter bearing failure; fault types include inner fault, outer fault, ball fault and the normal, and include a total of four types of sample data, each type includes 180 set of training samples and 60 set of testing samples.

The experimental results show that under the condition of without considering fault severity, only using time-domain features as the BP neural network input, can reach 100% (240/240) of the fault diagnosis, the total time spent on training and testing is about 102 seconds.

Table 1. Fault diagnosis based on BP Neural Network without considering fault severity

Fault Category	Fault Diameter (inches)	Fault severity	Label（Y）	Training Samples	Test Samples
Normal	0	Normal	0	180	60
Inner Fault	0.007	Slight	1	180	60
Outer Fault	0.007	Slight	2	180	60
Ball Fault	0.007	Slight	3	180	60
The experimental results				100%	

Testing scenarios two: Take time-domain features, and the hidden layer of the neural network consists of two layers; faults contain 0.007, 0.014 and 0.021 inches in diameter; fault types include inner fault, outer fault, ball fault and the normal, and include a total of ten types of sample data. The training samples and testing samples as shown in Table 2.

Table 2. Fault diagnosis based on BP Neural Network with Time-Domain feature

Fault Category	Fault Diameter (inches)	Fault severity	Label（Y）	Training Samples	Test Samples
Normal	0	Normal	0	180	60
Inner Fault	0.007	Slight	1	60	20
	0.014	Medium	1	60	20
	0.021	Serious	1	60	20
Outer Fault	0.007	Slight	2	60	20
	0.014	Medium	2	60	20
	0.021	Serious	2	60	20
Ball Fault	0.007	Slight	3	60	20
	0.014	Medium	3	60	20
	0.021	Serious	3	60	20
The experimental results				85.0%	

The experimental results show that under the condition of considering the fault severity, only using time-domain features as the BP neural network input, the fault recognition rate is 85.0% (204/240) of the fault diagnosis, the total time spent on training and testing is about 132 seconds.

Testing scenarios three: Take the time-domain characteristics and based on wavelet packet energy feature; Using BP algorithm of particle swarm optimization; Using data fault is 0.007, 0.014 and 0.021 inches in diameter, Fault types including inner fault and outer fault, ball fault and the normal, including a total of ten types of sample data. The training samples and test samples as shown in Table 3.

Table 3. Bearing fault diagnosis based on PSO-BP algorithm

Fault Category	Fault Diameter (inches)	Fault severity	Label（Y）	Training Samples	Test Samples
Normal	0	Normal	0	180	60
Inner Fault	0.007	Slight	1	180	60
	0.014	Medium	2	180	60

	0.021	Serious	3	180	60
Outer Fault	0.007	Slight	4	180	60
	0.014	Medium	5	180	60
	0.021	Serious	6	180	60
Ball Failure	0.007	Slight	7	180	60
	0.014	Medium	8	180	60
	0.021	Serious	9	180	60
The experimental results				99.1667%	

In Fig.4, the experimental results show that the accuracy of PSO-BP neural network is 99.1667% (595/600), and the recognition rate has been improved remarkably by using the time-domain feature and the energy characteristics based on wavelet packet.

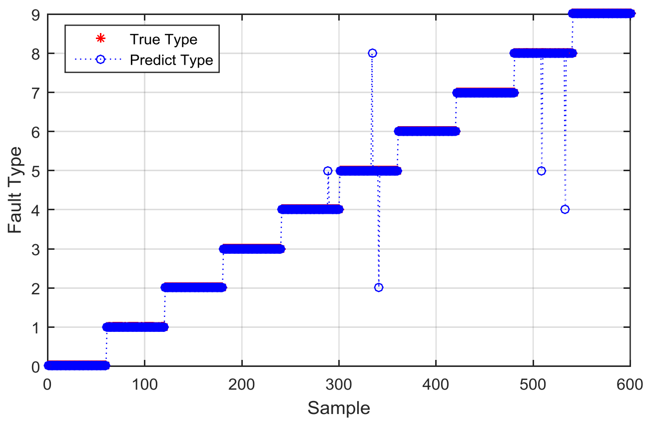

Fig.4 Fault diagnosis results based on PSO-BP algorithm

5. Conclusion

In the fault diagnosis of bearings, the common diagnostic method does not consider the influence of bearing failure of different severity on the fault diagnosis. Therefore, although the selected feature vector is low, the classification accuracy is higher. But in practical applications, if still using the same characteristics and algorithms, classification effect is not good.

Based on the practical application of fault diagnosis, this paper validates the influence of bearing failure of different severity on classification accuracy. On this basis, the time-domain characteristics of the vibration signal and the energy characteristics based on the wavelet packet are analyzed, and the extracted features have good reliability and sensitivity in the actual diagnosis.

Not only the feature selection has a greater impact on the classification results, but also the appropriate network connection weights and threshold parameters can improve the classification effect. In this paper, particle swarm optimization algorithm is used to optimize the parameters of BP neural network. The experimental results show that the PSO- BP neural network has the advantages of both PSO and BP neural network, which can accelerate the convergence speed and improve the generalization ability of the model.

To sum up, the bearing fault diagnosis method proposed in this paper not only has high diagnostic rate, but also has strong practicability.

References

[1] Bu Y, Wu J, Ma J, et al. The rolling bearing fault diagnosis based on LMD and LS-SVM[C]// Chinese Control and Decision Conference. 2014:3797-3801.

[2] Zhou W, Habetler T G, Harley R G. Bearing Fault Detection Via Stator Current Noise Cancellation and Statistical Control[J]. IEEE Transactions on Industrial Electronics, 2008, 55(12):4260-4269.

[3] Zhang X, Liang Y, Zhou J, et al. A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM[J]. Measurement, 2015, 69:164-179.
[4] Wong P K, Zhong J, Yang Z, et al. Sparse Bayesian extreme learning committee machine for engine simultaneous fault diagnosis[J]. Neurocomputing, 2016, 174:331-343.
[5] Pandya D H, Upadhyay S H, Harsha S P. Fault diagnosis of rolling element bearing by using multinomial logistic regression and wavelet packet transform[J]. Soft Computing, 2014, 18(2):255-266.
[6] Zhang L, Zhou Z. The BP Neural Network Fault Diagnosis of Bearings Based on Wavelet and Information Granulation[J]. Mechanical Science & Technology for Aerospace Engineering, 2012.
[7] Zhang Z, Wang Y, Wang K. Fault diagnosis and prognosis using wavelet packet decomposition, Fourier transform and artificial neural network[J]. Journal of Intelligent Manufacturing, 2013, 24(6):1213-1227.
[8] Zhao L Y, Wang L, Yan R Q. Rolling Bearing Fault Diagnosis Based on Wavelet Packet Decomposition and Multi-Scale Permutation Entropy[J]. Entropy, 2015, 17(9):6447-6461.
[9] Qin T, Yang Y, Cheng H, et al. Rolling bearing fault diagnosis based on intrinsic mode function energy moment and BP neural network[J]. Journal of Vibration Measurement & Diagnosis, 2008.
[10] Kennedy J, Eberhart R. Particle swarm optimization[C]// IEEE International Conference on Neural Networks, 1995. Proceedings. IEEE, 1995:1942-1948 vol.4.