The Fault Diagnosis of Rolling Bearings Based on Wavelet Packet Entropy and Affinity Propagation Clustering

Huizhong Huang¹, Fei Xu¹ and Zhansi Jiang¹

¹School of Mechanical Engineering, Guilin University of Electronic Technology, Guangxi, 541004, PR China
*Corresponding author’s e-mail: 425824473@qq.com

Abstract. In this paper, a fault diagnosis method based on wavelet packet entropy and affinity propagation clustering method is proposed for fault diagnosis of rolling bearings. The features of fault bearings are extracted by wavelet packet entropy method, and affinity propagation algorithm is used to cluster the features to realize the fault diagnosis of rolling bearings. The bearing fault data of case Western Reserve University are used to verify the proposed method, and the results show that the diagnosis of rolling bearings can be achieved by this method effectively.

1. Introduction

As an important part of rotating machinery, the failure of rolling bearings may cause serious consequences. There are two processes included by the fault diagnosis of rolling bearings: the feature extraction and pattern recognition. Wavelet packet entropy has been widely used in research and production because of its advantages in feature extraction. In reference [1], the wavelet packet entropy characteristics of gas concentration change are extracted by using the wavelet packet decomposition method based on reconstructed signal energy, so as to make an early warning of coal and gas outburst. In reference [2], the signal is decomposed by wavelet packet to obtain the node signal with fault characteristic components, and the last layer node signal is screened and the sensitive node signal is obtained according to the criterion of correlation number-wavelet packet entropy, and the bearing fault diagnosis is realized by reconstruction of signals. In reference [3], the collected current signal is transformed by wavelet packet, and its feature entropy vector is extracted as the sample data set, and the Mahalanobis distance classification method is used for fault diagnosis of frequency converter. In reference [4], the wavelet packet entropy and manifold learning method are combined to successfully realize the identification of caving coal and rock. In reference [5], an intelligent fault diagnosis method for train bearings based on wavelet packet entropy and multi-core learning is proposed. Affine propagation clustering determines the clustering centre by iterative method, and realizes the clustering process quickly and effectively. Reference [6] shows that the sliding time window method based on affinity propagation clustering has higher transaction extraction efficiency and solves the irrationality of the fixed time segment partition method. Reference [7] makes use of the advantage of affinity propagation clustering to automatically obtain the number of categories, and effectively subdivides speech into non-semantic speech and mute segments, far-field noise segments and so on. Reference [8] verifies that affinity propagation clustering not only excavates the local feature information of training samples, but also effectively reduces the computational complexity of local algorithms and avoids the occurrence of overlearning problems. Compared with the traditional load curve clustering method,
affinity propagation algorithm has better clustering effect in reference [9]. In reference [10], the vibration signal of rolling bearings is decomposed and calculated by EEMD method, and the entropy value is used as the input of affinity propagation clustering algorithm, and the fault diagnosis of rolling bearings is successfully realized. This paper attempts to use wavelet packet entropy to extract the fault features of rolling bearings, and to use affinity propagation algorithm to cluster the features, so as to realize the fault diagnosis of rolling bearings.

2. Algorithm

2.1. Wavelet packet entropy

Wavelet transform inherits and develops the idea of localization of short-time Fourier transform, and overcomes the shortcomings of window size changing without frequency. It can provide a time-frequency window with frequency change, and it is a tool for signal time-frequency analysis and processing. Only the low frequency part of the signal can be further decomposed by wavelet transform, while the high frequency part of the signal cannot be decomposed by it. Hence, a large class of signals with low frequency information as the main component can be better represented by wavelet transform, while the signals with a large number of details cannot be processed well. However, wavelet packet decomposition can solve the above problems. It can provide a good decomposition of the high frequency part, and the time frequency localization analysis of the signal containing a large number of medium and high frequency information can be carried out. Figure 1 is a schematic diagram of wavelet packet signal decomposition band division.

![Figure 1. Diagram of wavelet packet decomposition](image)

Entropy is the probability of the appearance of a particular information. Entropy $H$ can be defined in this way:

$$H = -\sum_{i=1}^{L} P_i \cdot \log P_i$$

Where $L$ is the total number of signal source state; The probability of $P_i$ signal values. $\sum_{i=1}^{L} P_i = 1, P_i \in [0,1]$. The more orderly the system performance, the smaller the corresponding entropy value, and the more chaotic the system performance, the greater the corresponding entropy value.

Wavelet packet entropy is based on wavelet packet decomposition and the wavelet packet coefficient is used to calculate the entropy value.

2.2. Affine propagation clustering algorithm

Affine propagation clustering (AP) is a fast and effective clustering method. It does not need to determine the number of the clusters in advance, and the clustering center is determined by the iterative method to get better results. The affine propagation clustering can be classified according to the industrial process data, so that the clustering result is more matched with the characteristics of the process objects.

The number of samples set up for clustering is $N$, and the affine propagation clustering is initially
used as a potential clustering center for the \( N \) sampling points, and the cluster center is selected for iteration according to the similarity criteria before the adjacent sample points. \( S(i, j) \) represents the similarity of any two sample points \( x_i \) and \( x_j \) the similarity of the similarity that \( N \times N \) can be represented by the similarity matrix \( S \). The AP algorithm redefines the elements on the main diagonal \( S(k, k) \) of \( S \), called "bias parameters", represented by \( p \). When initializing \( S \), the bias parameter generally takes the same value, the value is larger, the more classes the cluster gets, and the less the other way around, the result of the cluster can be adjusted by changing its value. When prior knowledge is unknown, \( P \) generally takes the median or minimum value of \( S \), or adjusts according to the actual situation.

\[
P = \beta \cdot \text{median}(S)
\]  

(2)

Where \( \beta \) can be adjusted for different processes. \( R(i, k) \) is set as the appropriateness of the clustering center, and \( A(i, k) \) is the appropriateness to choose as clustering center. The AP algorithm gets the appropriate cluster center by circularly updating \( R(i, k) \) and \( A(i, k) \). The update formula is as follows:

\[
\begin{align*}
R(i, k) &= S(i, k) - \max_{j \neq k} \{ A(i, j) + S(i, j) \} \\
A(i, k) &= \min \left\{ 0, R(k, k) + \sum_{j \neq i} \max \{ 0, R(j, k) \} \right\}
\end{align*}
\]  

(3)

When \( x_k \) satisfies the sum of \( A(i, k) + R(i, k) \) as the maximum of which, \( x_k \) is the cluster center of \( x_j \). Through the above steps, you can get the optimal clustering results.

### 3. Experiment

#### 3.1. Test bench configuration

The experimental data for this example come from the Electrical Engineering Laboratory of Case Western Reserve University. The bearings are mounted on the drive end of the drive motor, and the entire experimental device is shown in Figure 2. A 1.5kw 3PH induction motor is connected to a torque sensor via a self-calibrated coupling and the load motor that drives the analog fan. The load on the drive motor is mainly regulated by the analog fan. The vibration acceleration sensor is placed outside the enclosure and above the support bearing at the driving end of the induction motor, and the data acquisition system includes a data recording device with a sampling frequency of 48kHz and a bandwidth amplifier, with a useful signal component of no more than 5000Hz, so the selected sampling frequency can meet the standard. At the same time, the data recording device is equipped with a low-pass filter to resist mixed filtering.

![Figure 2. Rolling bearing fault simulation device](image)

In this example, the rotation speed of the motor is about 1797r/ min, the load is set to 0, the experimental bearing is a deep groove ball bearing produced by the SKF company of Sweden, the fault category is the fault of the inner ring of the outer ring of the Outer, and the fault of the bearing
damage is obtained by electric machining. The same degree of damage was simulated for each different fault, with a depth of 0.007 inches. The experimental parameters related to the rolling bearing are shown in Table 1.

| Signal category | $f_r$ (r/min) | $f_s$ (kHz/s) | inch |
|-----------------|--------------|--------------|------|
| Outer           | 1797         | 12           | 0.007|
| Inner           | 1797         | 12           | 0.007|
| Ball            | 1797         | 12           | 0.007|
| Normal          | 1797         | 12           | 0 |

### 3.2. Test process

2048 points are taken from four sets of experimental data for Fourier transformation to obtain its time domain map and frequency domain map, which is shown in Figure 3.

Using db5 wavelet, 3 layers of wavelet packet decomposition is used for each set of data, and 8 layers of wavelet packet decomposition coefficient is reconstructed to obtain energy $E$ for each node. The total energy of the wavelet packet is found, and the energy distribution of each band of the wavelet packet is found after the classification. Finally, the wavelet packet entropy $p$ is extracted. The entropy of four groups of wavelet packets is formed into a new matrix, and the results are obtained from the affine propagation cluster as follows Figure 4:

![Wavelet packet entropy and AP cluster diagnostic results](image)

As can be seen from the figure above, four different types of bearing signals can be completely distinguished by the wavelet packet entropy and affine propagation clustering method. In order to reflect the effectiveness of the method, the wavelet packet entropy and ADAP clustering are combined
to analyze the same data. Get the results as shown below:

![Figure 5. Wavelet packet entropy and ADAP cluster diagnostic results](image)

From Figure 5, it can be seen that the results obtained by using wavelet packet entropy and ADAP clustering cannot distinguish four different types of bearing signals well.

4. Conclusion
In this paper, a method based on the clustering of wavelet packet entropy and affine propagation is proposed for rolling bearing fault diagnosis. Through the above examples, we can get that this method can be used to differentiate fault well. Compared with the experimental results of wavelet packet entropy and ADAP clustering, it possesses a better recognition rate. thus proving the validity of this method. Hence, this method can be proved effective.

Acknowledgments
The work is supported by the National Natural Science Foundation of China (No. 51565008), Natural Science Foundation of Guangxi (No. 2017JJA160071z) and the Guangxi Key Laboratory of Manufacturing System & Advanced Manufacturing Technology (No. 1514030006K).

References
[1] Tu N W, Yan X. (2019) Study on Early-warning of Coal and Gas Outburst Based on Wavelet Packet Entropy and Data Fusion. Control Engineering of China, 26: 759-764.
[2] Yang S, Zhu A H, Yao D C, Yang J W, Bai Y L. (2018) Weak Fault Signal Extraction of Gear Box Bearing Based on MCKD and Wavelet Packet Entropy. Modular Machine Tool & Automatic Manufacturing Technique, 05: 131-134+138.
[3] Wang R, Wang X. (2017) Fault Diagnosis Based on Entropy of Wavelet Packet and Mahalanobis Distance for Cascade Frequency Converter. Process Automation Instrumentation, 38: 27-30+36.
[4] Li Y M, Fu S C, Zhou J Y, Zong K, Li R, Wu M. (2017) Collapsing coal-rock identification based on wavelet packet entropy and manifold learning. Journal of China Coal Society, 42: 585-593.
[5] Zhou P T, San Q, Ye Y G. (2017) Application of wavelet packet entropy and multi kernel learning in fault diagnosis of train bogie bearing. Journal of Yanshan University, 41: 401-406.
[6] Lin B, Zeng Y, Li X N, Chen W W, Luo Y. (2016) An analysis method of power communication alarm information based on affinity propagation clustering. International Electronic Elements, 24: 142-145.
[7] Lin Q, Tu Z Z, Wang Q W, Guo Y T. (2019) Voice activity detection based on affinity propagation clustering. Journal Of Anhui University(Natural Science Edition), 43: 27-32.
[8] Zhang W, Xu A Q, Ping D F, Xia F. (2018) Localized multi-kernel diagnosis model for avionics based on affinity propagation clustering. Journal of Beijing University of Aeronautics and Astronautics, 44: 1693-1704.

[9] Peng B, Li Z H, Li M, Yang Y, Xu W, Ma W H. (2019) Power Load Profiles Clustering Method based on Affinity Propagation Algorithm. Mechanical & Electrical Engineering Technology, 48: 183-186+224.

[10] Xu F, Fang Y J, Kong Z M. (2017) Roller Bearings AP Clustering Fault Diagnosis Based on EEMD Sample Entropy. Instrument Technique and Sensor, 06: 129-135.