Unsupervised Learning — A Novel Clustering Method for Rolling Bearing Faults Identification

Li Kai, Luo Bo, Ma Tao, Yang Xuefeng and Wang Guangming
1037 Luoyu Road, Hongshan District, Wuhan(430074) Hubei Province, P. R. China
15527965836@163.com

Abstract. To promptly process the massive fault data and automatically provide accurate diagnosis results, numerous studies have been conducted on intelligent fault diagnosis of rolling bearing. Among these studies, such as artificial neural networks, support vector machines, decision trees and other supervised learning methods are used commonly. These methods can detect the failure of rolling bearing effectively, but to achieve better detection results, it often requires a lot of training samples. Based on above, a novel clustering method is proposed in this paper. This novel method is able to find the correct number of clusters automatically the effectiveness of the proposed method is validated using datasets from rolling element bearings. The diagnosis results show that the proposed method can accurately detect the fault types of small samples. Meanwhile, the diagnosis results are also relative high accuracy even for massive samples.

1. Introduction
The rolling bearing is a key part which is widely used in rotating machinery, ranging from simple electric fans to complex machine tools [1]. The traditional technology of fault diagnosis usually fails in analyzing complex faults with no further fault information, and intelligent diagnosis theory based on machine learning has been central in study and application recently[2,3]. Correspondingly, supervised learning methods such as artificial neural networks(ANN), decision trees, support vector machines(SVM), etc have been used widely. Samanta et al. [4] employed ANNs and SVM to diagnose faults of bearings and utilized time domain features to characterize the health conditions of the bearing. Tamilselvan [5] used the deep belief network (DBN) in an aircraft engine fault diagnose is based on health state classification in a small sample with enough fault information. Widodo et al [6] proposed two methods of multi-class classification techniques for fault diagnosis through relevance vector machine(RVM) and support vector machines(SVM), aiming to detect the fault of low speed bearing. B.A. Paya and I.I. Esat[7] preprocessed the real time domain vibration signals of bearing and gear faults by wavelet transforms, then employed the neural network to perform fault detection. Hai dong shao et al[8] presented a novel optimization deep belief network(DBN) for rolling bearing fault diagnosis, this method has a good generalization performance in complex classifications. Feng Jia, Yaguo Lei et al[9] presented a novel intelligent method which overcome the deficiencies of the aforementioned intelligent diagnosis methods for detecting of rotating machinery with massive data, the DNNs are trained by using raw signals in the time domain. Oveall, supervised learning methods have been more mature in the application of fault diagnosis, but it generally requires a massive fault or health samples for training learning model. However, during the actual processing, the lifecycle of rolling bearing is relatively long. So it is difficult for acquiring a large number of fault samples.
Moreover, the unknown fault types may occur during production, and the learning model can not accurately detect the fault types at that time. In 2014, Alex Rodriguez and Alessandro Laio[10] proposed a novel approach, this algorithm assume that cluster centers are surrounded by neighbors with lower local density and that they are at a relatively large distance from any points with a higher local density. This novel clustering method can make up for the shortcomings of the above two methods. Meanwhile, it is suitable for nonspherical clusters and determine clustering centers automatically.

2. A brief introduction to the novel clustering method

The novel clustering algorithm seems to be able to meet the above requirements. The following is a brief introduction to the novel clustering method in two steps.

2.1. The local density $\rho_i$ and distance $\delta_i$

Cluster centers can be defined as two main features: (1) Cluster centers are surrounded by neighbors with lower local density. (2) They are at a relatively large distance from any points with a higher local density [9]. That is a cluster center has higher $\rho_i$ and large value of $\delta_i$ as compared with non-center data points. Suppose the dataset to be clustered is $S = \{x_i\}_{i=1}^N$. The cut-off local density $\rho_i$ of data point is defined as[8,9]:

$$\rho_i = \sum_j \chi(d_{ij} - d_c)
\tag{1}$$

Where $\chi(x) = 1$ if $x < 0$ and $\chi(x) = 0$ otherwise. $\rho_i$ is equal to the number of points that are closer than $d_c$ to point $i$. $d_c$ is a cut-off distance that to be selected. As a rule of thumb, one can choose $d_c$ so that the average number of neighbors is around 1 to 2% of the total number of points in the data set [9]. From equation (1), we find that cut off kernel are discrete numeric values. It is easy to cause the different data points tend to be the same local density, which will have an impact on our searching clustering centers. Thus, we choose Gaussian kernel to calculate the local density $\rho_i$. Gaussian kernel can obtain continuous value, which reduces the probability of the same local density of different data points. The Gaussian kernel local density $\rho_i$ of data point is defined as:

$$\rho_i = \sum_{jd_{ij}(i)} e^{-\left(\frac{d_{ij}}{d_c}\right)^2
\tag{2}$$

Above equation (1) and equation (2) where $d_{ij}$ is the distance between data points $i$ and $j$. In machine learning, the similarity measures include: Euclidean distance, Manhattan distance, Minkowski distance, cosine of included angle, etc. For clustering in rolling bearing fault, we find that the Euclidean distance is better than others. $\delta_i$ measured by computing the minimum distance between the point $i$ and the nearest point with higher density[9]:

$$\delta_i = \min_{j\rho_j > \rho_i} (d_{ij})
\tag{3}$$

3. Fault diagnosis of rolling bearing using the novel clustering method

3.1. Data description
The bearing fault data used here are provided by the Case Western Reserve University (CWRU) collected from the electrical engineering laboratory [10]. The experimental platform consists of two horsepower motors, two acceleration sensors, a torque sensor and an electronic control device, as shown in Fig.1. The single point failure of the bearing is arranged on the bearing using the electric spark machining technology. The fault diameter is 0.007, 0.014, 0.021, 0.028 inch respectively (1 inch=25.4 mm). Each bearing was tested at four different loads (0, 1, 2, and 3), 0,1,2,3 load corresponding to 1797rpm, 1772rpm, 1750rpm, 1730rpm running speed of motor respectively.

Experimental data of rolling bearings are acquired from the experimental system under four different operating conditions [11]: (1) the normal conditions; (2) with the inner race fault; and (3) with the ball fault. (4) With the outer race fault. In addition, the rolling bearing with outer race fault has three measurement directions which is 3 o'clock, 6 o'clock and 12 o’clock.

3.2. Data preprocessing

The number of fault sample in 12K drive end is 59, as shown in Table 1. Sampling time of each sample data is 10 seconds. Due to the different length of time domain vibration signal between different bearing fault types. However, we need to measure the similarity between different bearing fault types. However, we need to measure the similarity between different bearing fault types.

Therefore, the time domain signal data of each fault type need to be preprocessed, and the preprocessing steps are as follows: (1) the time domain signal data interception of each bearing fault. Each sample is taken from the first 12000 data points, as shown in Fig.2(1); (2) Calculating auto-power spectrum for the interception of the fault time domain vibration signal, as shown in Fig.2(2); (3) Smoothing the auto-power spectrum of each bearing fault sample, as shown in Fig.2(3). It is found that the smoothing process is very important to the clustering result. As the clustering process is very sensitive to the noise signal, so it is needed to eliminate the noise by smoothing. In addition, as there are many kinds of fault types in 12K drive end, in order to simplify the expression, we choose one kind of bearing fault types from three different fault positions (the ball fault, the inner race fault and the outer race fault) to represent the pretreatment process, as shown in Fig. 2.

3.3. Diagnosis results

After the raw bearing fault vibration signals are preprocessed in 3.2 section. The novel clustering method is used to cluster the 12K drive end bearing fault samples to find out the same type of faults. The similarity between each bearing fault sample is computed by Euclidean distance. The density is estimated by a Gaussian kernel with variance $d_c = 0.03$. The clustering decision graph is shown in Fig.3 (a) when the auto-power spectrum is not smoothed.
Table 1. Description of 12K drive end bearing fault datasets

| Fault position       | Fault diameter (inch) | Fault direction | Load condition |
|----------------------|-----------------------|-----------------|----------------|
| The ball fault       | 0.007                 | Non             | 0,1,2,3        |
|                      | 0.014                 |                 |                |
|                      | 0.021                 |                 |                |
|                      | 0.028                 |                 |                |
| The inner race fault | 0.007                 | Non             | 0,1,2,3        |
|                      | 0.014                 |                 |                |
|                      | 0.021                 |                 |                |
|                      | 0.028                 |                 |                |
| The outer race fault | 0.014                 | 6 o’clock       | 0,1,2,3        |
|                      | 0.021                 | 3 o’clock       | 0,1,2,3        |

Figure 2. The time domain vibration signal and spectra for three different positions of bearing fault: (a) The ball fault (The fault diameter is 0.007 inch under 0 load condition), (b) The inner race fault (The fault diameter is 0.007 inch under 0 load condition), (c) The outer race fault (The fault diameter is 0.007 inch and Fault direction is 3 o’clock under 0 load condition)
Figure 3. Cluster analysis of 12K drive end bearing fault. (a) The corresponding decision graphs, with the auto-power spectrum is not smoothed, (b) The corresponding decision graphs, with the auto-power spectrum has been smoothed and the centers colored by cluster, (c) The distribution graph of the sample after classification.

Fig. 3(b) shows the clustering decision graph after smoothing and the clustering results of each sample is depicted in Fig.3(c). Compared with Fig.3 (b), the exact number of cluster centers is not clear and a consequence of the sparsity of the data points in Fig.3 (a). Therefore, it is required to smooth the auto-power spectrum before clustering bearing fault samples. In addition, From Fig. 3 (b) and Fig.3 (c) shows that the 12K drive end bearing fault is divided into 15 clusters. In addition, according to the number of clustering centers we found using the novel clustering method, K-means was used to classify the rolling bearing fault samples again. Compared with the two clustering methods, both of the methods can identify all kinds of fault types, but the novel clustering method does not need to determine the number of clusters in advance and the accuracy rate of the classifying results is 100%.

4. Conclusions
Currently, supervised learning methods are commonly used for the fault diagnosis of bearing and other related parts and the accuracy is often very significant. However, supervised learning methods often require massive the known fault types as training samples, which is not so easy for the actual
processing conditions. And as the actual processing conditions are often more complex, so there will be some unknown fault types occasionally. It is a challenge for fault diagnosis using the supervised learning methods. On this basis, this paper proposes a novel method for classifying the fault types of rolling bearings, this clustering method that can find the cluster centers from the samples automatically is different from the traditional K-means and K-medoids, it does not need to determine the number of clusters in advance. The effectiveness of the proposed method is verified using 59 samples. 59 rolling bearing fault samples contain a lot of fault types, clustering results show that each of fault types can be separated; the accuracy can even reach 100%. But by looking for suspected clustering centers can distinguish between high similarities of fault types. On the other hand, the novel clustering method is sensitive to noise, thus the auto-power spectrum of the time domain signal of the bearing faults should be smoothed before classifying, so that the classifying effect will be more significant.

Acknowledgments
The research is supported by the National Natural Science Foundation of China under Grant nos. 51375193, the Postdoctoral Science Foundation of China under Grant nos. 2016M602282 and the State key laboratory smart manufacturing for special vehicles and transmission system nos.GZ2016KF008.the National Major Special Projects Foundation of China under Grant nos. 2014ZX04014101. the National Natural Science Foundation of China under Grant nos. 51775210

References
[1] Wang W and Lee H, An energy kurtosis demodulation technique for signal denoising and bearing fault detection Meas. Sci. Technol. 24 (2013)025-601
[2] Liu J, Wang W, Golnaraghi F and Liu K, Shannon wavelet spectrum analysis for bearing fault diagnostics Meas. Sci. Technol. 19(2008) 015-105
[3] Lei Y G, He Z J and Zi Y Y, Application of an intelligent classification method to mechanical fault diagnosis Expert. Syst. Appl. 36(2009) 9941–8
[4]B. Samanta, Artificial neural networks and genetic algorithms for gear fault detection, Mech. Syst. Signal Process. 18 (2004) 1273–1282.
[5] Tamilselvan P and Wang P F, Failure diagnosis using deep belief learning based health state classification Reliable. Eng. Syst. Saf. 115 (2013)124–35
[6] A. Widodo, E.Y. Kim, J.-D. Son, B.-S. Yang, A.C. Tan, D.-S. Gu, B.-K. Choi, J. Mathew, Fault diagnosis of low speed bearing based on relevance vector machine and support vector machine, Expert Syst. Appl. 36 (2009) 7252–7261.
[7]B.A. Paya, I.I. Esat, Artificial neural network based fault diagnostics of rotating machinery using wavelet transforms as a preprocessor, Mech. Syst. Signal Process. 11 (1997) 751–765
[8] H. Shao, H. Jiang, X. Zhang, and M. Niu,Rolling bearing fault diagnosis using an optimization deep belief network, Meas.Sci. Technol.26(2015)115-002
[9] Feng Jia, Yaguo Lei n, Jing Lin, Xin Zhou, Na Lu, Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data, Mech. Syst. Signal Process. 72-73 (2016) 303–315
[10] Rodriguez, A. and Laio,A.Clustering by fast search and find of density peaks. Science (New York, N.Y.), 344(2104)1492–1496
[11] X.S. Lou, K.A. Loparo, Bearing fault diagnosis based on wavelet transform and fuzzy inference, Mech. Syst. Signal Process. 18 (2004) 1077–109