Rolling bearing fault diagnosis based on information fusion using Dempster-Shafer evidence theory

Di Pei, Jianhai Yue and Jing Jiao
School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, No.3 Shangyuancun, Haidian District, Beijing, China
E-mail: 15121291@bjtu.edu.cn

Abstract. This paper presents a fault diagnosis method for rolling bearing based on information fusion. Acceleration sensors are arranged at different position to get bearing vibration data as diagnostic evidence. The Dempster-Shafer (D-S) evidence theory is used to fuse multi-sensor data to improve diagnostic accuracy. The efficiency of the proposed method is demonstrated by the high speed train transmission test bench. The results of experiment show that the proposed method in this paper improves the rolling bearing fault diagnosis accuracy compared with traditional signal analysis methods.

1. Introduction
In recent years, many methods have been used for bearing fault detection. The methods are broadly classified as acoustic measurement [1], current and temperature monitoring [2], wear debris detection [3] and vibration analysis. Compare all diagnosis methods, vibration analysis is widely used. However, vibration analysis also has its disadvantage. For example, using a single sensor has the limitation in the amount and type of data. Once the sensor failed, bearing vibration signal cannot be obtained. So, using multiple sensors to detect the working state of bearing can get more diagnostic information and improve diagnostic reliability.

Multi-sensor information fusion is a combination of data from multiple sensors and related information to obtain more detailed and more accurate reasoning results than a single sensor. Evidence theory is initially based on Dempster’s work and extended later by Shafer. As an extension to Bayes theory, D-S evidence theory uses belief and plausibility functions to quantify evidence and uncertainty. D-S evidence theory models how the uncertainty of a given hypothesis or discourse diminishes as pieces of evidence accumulate during the reasoning process [4].

The applications of D-S evidence theory in mechanical fault detection and diagnosis are reported in references [4-7]. In reference [4], D-S evidence theory is used for engine valve fault diagnosis. Reference [5] proposed a method that combine SVM with D-S evidence theory to resolve conflicting results generated from each SVM model. In reference [6], neural network and D-S evidence theory are combined to fuse the data from vibration signal and stator current signal for motor fault diagnosis. Reference [7] describes how evidence theory and fuzzy set theory are combined to improve the detection quality of weld defects.

2. Fault diagnosis based on D-S evidence theory

2.1. Basic of D-S evidence theory

[continued]
Let $\Theta$ be a finite non-empty set of mutually $n$ exhaustive and exclusive hypotheses about some problem domain. Evidence theory first assumes the definition of a set of hypotheses $\Theta = \{A_1, A_2, \ldots, A_n\}$, called the frame of discernment. The set composed with the proposition $A$ of $\Theta$ is denoted as $2^\Theta$.

$$2^\Theta = \{\emptyset, \{A_1\}, \{A_2\}, \ldots, \{A_i\}, \{A_i, A_j\}, \{A_i, A_j, A_k\}, \ldots, \Theta\}$$

(1)

Assuming $\Theta$ is the frame discernment, if the function $m$: $2^\Theta \rightarrow [0,1]$ satisfies

$$m(\emptyset) = 0, \quad \sum_{A \subseteq \Theta} m(A) = 1, \quad \forall A \subseteq \Theta,$$

(2)

$m$ is called the basic probability assignment (BPA), also called the mass function. $m(A)$ expresses the $A$’s value of basic probability assignment obtained by analysis.

The belief function ($Bel$) is defined as:

$$\forall A \subseteq \Theta, \quad Bel(A) = \sum_{B \subseteq A} m(B),$$

(3)

$Bel$ is used to express the belief degree of $A$’s corresponding proposition.

Assuming that the function $Pl$: $2^\Theta \rightarrow [0,1]$ satisfies

$$\forall A \subseteq \Theta, \quad Pl(A) = 1 - Bel(\bar{A}) = \sum_{A \neq B} m(B),$$

(4)

$Pl$ is called the plausibility function, reflecting the confidence degree of $A$’s corresponding proposition.

Assuming that $m_1, m_2, \ldots, m_n$ are $n$ BPAs in the same $\Theta$. According to the Dempster’s orthogonal rule, we have

$$m_{\Theta \cdots \Theta}(A) = \begin{cases} \frac{1}{K} \sum_{E_i = A} m_1(E_1) \cdot m_2(E_2) \cdots m_n(E_n), & A \neq \emptyset, \\ 0, & A = \emptyset \end{cases}$$

(5)

$$K = \sum_{E_i = \emptyset} m_1(E_1) \cdot m_2(E_2) \cdots m_n(E_n) > 0,$$

(6)

the value of $K$ reflects the conflict degree of evidence. The higher value of $K$, the higher degree of evidence confliction, the less the information obtained after evidence combination.

2.2. Calculation of mass functions

In this paper, the mass functions are calculated by feature vector method introduced in Ref. [4]. Assuming that the number of bearing fault is $N$, and these faults are independent of each other.

Let

$$X = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$$

(7)

$X$ is a feature vector that is composed of the feature parameters of vibration signal of known fault bearings, $x_i$ represents the $i$th feature, $n$ is the number of feature parameters. Thus, for $N$ types of bearing faults (including normal state), the working state can be described by a matrix $H$:

$$H = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{Nm} \end{bmatrix}$$

(8)

$X_j = [x_{j1}, x_{j2} \ldots x_{jn}]$, and $x_{ji}$ is the $i$th feature parameter of the $j$th fault, $j=1,2,\ldots,N$, $i=1,2,\ldots,n$. The matrix $H$ can be regarded as the fault prototype matrix. Similarly, the measurement feature vector $S_k$ can be defined as

$$S_k = [s_{k1}, s_{k2} \ldots s_{km}], \quad k = 1, 2, \ldots, M$$

(9)

$s_{ki}$ is the $i$th element of $S_k$, $m_k$ is the number of feature parameters obtained by the $k$th sensor, $\sum m_k = n$. The distance between $S_k$ and $X_j$ is denoted as


\[ d_{ij} = \begin{cases} \left[ \sum_{i=1}^{m_k} (s_{ki} - x_{kj})^a \right]^{1/a} & k = 1 \\ \left[ \sum_{i=1}^{m_k} (s_{ki} - x_{j(i+m_k)})^a \right]^{1/a} & k > 1 \end{cases} \]

\[ k = 1, 1, 2, \ldots, M \quad j = 1, 2, \ldots, N \]  

(10)

Then, defining \( p_{kj} = 1/d_{kj} \), the value of \( p_{kj} \) can be regarded as the BPA of the \( j \)th fault, as far as the \( k \)th sensor is concerned, that is,

\[ m_k(A_j) = p_{kj} \]  

(11)

3. Experiment and analysis

3.1. Experiment Setup

The experiment of this paper is based on the transmission test bench of high speed train. The test bench is mainly composed of motor, gear box, rolling bearing, bearing seat, wheelset, transmission shaft and generator. The model of bearing in this experiment is 351306, and the bearing can be changed. Acceleration sensors are set on bearing seats, as shown in figure 1.

Figure 1. The transmission test bench of high speed train

In experiment, we set bearing 1 as normal (N), inner ring fault (I) and outer ring fault (O) respectively. Bearing 2, bearing 3 and bearing 4 are set as normal. In this paper, in order to calculate the fault prototype matrix \( H \), three groups of experiments (denote as experiment A) were set up. Each group of experiment A has 60 times measurement. In order to calculate the measurement feature vector \( S_k \) \((k=1,2,3,4)\), 60 groups of experiments (denote as experiment B) were set up. Each group of experiment B has 3 times measurement in normal, inner ring fault and outer ring fault condition of bearing 1 individually. The settings of experiment A and B are shown in Table 1.

Table 1. The settings of experiment A and B

| Experiment settings | Group | Bearing 1 | Bearing 2 | Bearing 3 | Bearing 4 |
|---------------------|-------|-----------|-----------|-----------|-----------|
| A                   | 1     | N         | N         | N         | N         |
|                     | 2     | I         | N         | N         | N         |
|                     | 3     | O         | N         | N         | N         |
|                     | 1     | N         | I         | O         | N         |
|                     | 2     | N         | I         | O         | N         |
| B                   | 60    | N         | I         | O         | N         |

3.2. Fault diagnosis and analysis

Before using D-S evidence theory for fault diagnosis, experiment A was done to collect 180 groups of data for diagnosis based on traditional vibration signal analysis methods: time domain analysis and
frequency domain analysis. Then, we carry on the fault diagnosis by using D-S evidence theory. The peak-to-peak value and the peak frequency of bearing 1 vibration signal from experiment A are selected as the feature parameters.  

Then, we carry on the fault diagnosis by using D-S evidence theory. The fault prototype matrix $H$ can be calculated by feature parameters. In order to calculate the measurement feature vector $S_i$ and fuse sensors data, experiment B was designed. The mass function can be calculated by equation (9) to (11). Data of sensor 1 and sensor 2, sensor 1 and sensor 3, sensor 1, sensor 2, sensor 3 and sensor 4 were fused respectively according to equation (5) and (6). Meanwhile, we set up a contrast test that only judge the bearing 1 condition by sensor 1. The fault diagnosis results of bearing 1 from different diagnostic methods are shown in table 2.

**Table 2.** Comparison of different diagnostic methods in diagnostic accuracy

| Correct number of diagnosis | Fault diagnosis methods |
|-----------------------------|-------------------------|
|                             | Time domain analysis    | frequency domain analysis | $S_1$ | $S_1+$ | $S_2$ | $S_2+$ | $S_3$ | $S_3+$ | $S_4$ | $S_4+$ |
| N                           | 45                      | 48                      | 42     | 47     | 48     | 57     |
| I                           | 42                      | 45                      | 39     | 49     | 51     | 54     |
| O                           | 45                      | 45                      | 45     | 51     | 54     | 57     |
| Accuracy (%)                | 73.33%                  | 76.67%                  | 70.00% | 81.67% | 85.00% | 93.33% |

The diagnosis accuracy based on time domain analysis and frequency domain analysis are 73.33% and 76.67% respectively. In the case of non-fusion, the basic probability assignment of the fault obtained by the sensor 1 is used only, and the corresponding condition of maximum basic reliability distribution value is taken as the diagnosis result, the diagnosis accuracy is 70%, lower than time domain analysis and frequency domain analysis. Fuse the data from sensor 1 and sensor 2, the diagnosis accuracy reaches to 81.67%; Fuse the data from sensor 1 and sensor 3, the diagnosis accuracy reaches to 85.0%; Fuse the data from the sensor 1, sensor 2, sensor 3 and sensor 4, the diagnosis accuracy reaches to 93.33%.

4. Conclusion

The experiment result based on transmission test bench of high speed train allow to draw the following conclusions:

1. Compared to traditional vibration signal analysis methods for rolling bearing fault diagnosis, using D-S evidence theory to fuse the sensors data can improve the diagnosis accuracy remarkably.

2. The diagnosis accuracy cannot be increased only by taking mass function of D-S evidence theory as judgement without evidence fusion.

3. When fusing the data from sensors, the location and number of sensors have an impact on the accuracy of the diagnosis to some extent, in general, the more the number of evidence, the higher the diagnostic accuracy.

Acknowledgments

The study is supported by Technology Division of China Railway (M15D00230).

References

[1] Hecke B V, J Yoon and D He 2016 *Applied Acoustics* **105** 35-44

[2] Seo J J, Yoon H, Ha H, Hong D P and Kim W 2011 *Advanced Materials Research* **31** 295-297 1544-47

[3] Zmitrowicz A 2005 *Journal of Theoretical & Applied Mechanics* **43** 3-35

[4] Basir O and Yuan X 2007 *Information Fusion* **8** 379-386

[5] Hui K H, Meng H L, Leong M S and Al-Obaidi S M 2017 *Engineering Applications of Artificial Intelligence* **57** 160-170

[6] Yang B S, Kim K J 2006 *Mechanical Systems & Signal Processing* **20** 403-420

[7] Kaftandjian V, Dupuis O, Babot D and Zhu Y M 2003 *Pattern Recognition Letters* **24** 547-564