

Research Article

Fault Diagnosis Method Based on Improved Evidence Reasoning

Jianbin Xiong ^{1,2}, Chunlin Li,¹ Jian Cen,¹ Qiong Liang,¹ and Yongda Cai³

¹School of Automation, Guangdong Polytechnic Normal University, Guangzhou 510000, China

²Guangdong Provincial Key Lab of Petrochemical Equipment Fault Diagnosis, Maoming 525000, China

³School of Information Engineering, Guangdong University of Technology, Guangzhou 510006, China

Correspondence should be addressed to Jianbin Xiong; 276158903@qq.com

Received 24 December 2018; Revised 31 January 2019; Accepted 13 February 2019; Published 4 March 2019

Academic Editor: Sajad Azizi

Copyright © 2019 Jianbin Xiong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Evidence reasoning (ER) combined with dimensionless index method can be used in rotating machinery fault diagnosis. In ER algorithm, reliability is mainly obtained in two ways: distance-based method and correlation measure by set theory. In practice, the distance-based method cannot generate high-discrimination reliability in high-coincidence data like dimensionless index data. Therefore, correlation measure by set theory method is used in fault diagnosis more frequently. Because correlation measure by set theory only considers upper bound and lower bound of fault data, we add a regularization term to calculate the relationship between the inner data. Experience result shows that fault diagnosis accuracy had improved, which illustrates that the new reliability can describe data relationship better.

1. Introduction

Rotating petrochemical machinery has become more and more complicated. For instance, the connection between parts is getting closer and closer. Its working and operating environment is more complex and demanding [1, 2]. Therefore, higher reliability and safety requirements are put forward for equipment design, structure, process, and operation state [3]. As a key component of petrochemical units, rotary units cover important engineering fields such as petrochemical, power, chemical, metallurgical, and mechanical manufacturing [3, 4]. Rotating unit equipment (such as generators, steam turbines, blowers, and large rolling mills) is often the plant's key equipment [5]. The operating condition not only affects the operation of the machine itself but also affects the subsequent production. Therefore, it is urgent to discuss and study the fault diagnosis technology of rotating units [6, 7].

At present, fault diagnosis methods can be divided into three types according to the diagnosis model: analytical based model, qualitative knowledge based model, and Dempster Shafer Theory based model. Fault diagnosis based on the analytic model is a method to find out the running rule of the object. By studying the intrinsic relation between dynamic parameters and response symptoms in fault state [8], the

information of normal operation and abnormal correlation is obtained. This kind of method is suitable for systems with an accurate quantitative mathematical model and a sufficient number of sensors. They can gain fault pattern recognition result by establishing physical model and mathematical model. Typical analytical model-based methods include the state estimation method, the parameter estimation method, the equivalent space method, and the analysis redundancy method.

Qualitative experience based fault diagnosis method is a kind of reasoning method based on the qualitative model. The core of this method is using incomplete prior knowledge to describe the function structure of the system and establish a qualitative model to realize reasoning. According to the model, the behavior of the system is predicted and compared with the actual system behavior to detect the failure of the system. This method usually includes expert system, graph search, and fault tree analysis. For complex fault diagnosis, because the number and combination of faults are unpredictable, the workload of constructing the qualitative model is relatively heavy. Especially for complex systems, unpredictable fault combinations will increase the scale of the model exponentially. Therefore, when this kind of method is applied, it is often used to analyze some specific complex faults [9, 10].

Evidence theory based fault diagnosis method is an inexact reasoning theory, which can deal with uncertainty information. The confidence interval is used to replace the probability, the event is represented by the set, and the Bayesian formula is replaced by the rule of evidence combination. The confidence function can be expressed directly by uncertainty and not knowing. In the application of composite fault diagnosis, the D-S evidence theory makes a decision result through the fusion reasoning of each evidence body on the same recognition frame [11–13]. At present, a large number of evidence theory based fault diagnosis methods are mainly aimed at the diagnosis of single fault, which requires that the elements in the identification framework have mutually exclusive relations [14]. But, for complex fault diagnosis, such settings have fundamental limitations. In order to extend its effectiveness in complex faults, the extended evidence theory (Dezert-Smarandache Theory, DS_mT), which uses the intersection of elements in the identification framework to represent concurrent and composite faults, is proposed. Taking [14] as an example, it gives an identification framework that can cover a single fault and a composite fault. It sets the correlation degree for different faults to the evidence. Each evidence in each group is decomposed into two kinds of evidence, which are independent and relevant. Then several independent pieces of evidence are fused by using the DS_mT combination rule, and the uncertainty of different independent source evidence is inferred; thus the identification of composite faults is realized. However, this method can not realize fault discrimination for conflicting evidence. Therefore, a fault diagnosis method based on evidence reasoning and the dimensionless index is proposed in [15]. This method can realize the fusion of multiple attributes and realize the diagnosis of conflicting evidence. However, there are many methods to calculate the reliability in the process of ER (e.g., [16] uses distance method and [17] uses set correlation measure), which has a great influence on reasoning results. Because of the coincidence of dimensionless index values between different faults, it is more effective to use the set correlation measure method. However, such a method will be affected by wild value in dimensionless indicators.

To solve the above problems, we propose a method to regularize the reliability value. We use the correlation coefficient as the regularization term to improve the reliability calculation formula. The Gini correlation coefficient is used in this paper because it can describe the relation of nonlinear data effectively [18]. Therefore, this paper uses the improved evidence reasoning algorithm and dimensionless index to carry out fault diagnosis. The main contributions of this paper are as follows:

(1) In traditional ER, the reliability only considers set interval in fault data, neglecting the impact of relationship between the inner data. In new proposed method, we used correlation coefficient as reliability regularization. In practice, the new reliability is more reliable.

(2) To achieve better fault diagnosis result, we combine improved ER and dimensionless indexes in rotating machinery fault diagnosis. The experimental result shows that the new method is better than the traditional one.

2. Related Work

2.1. Dimensionless Index. The dimensionless index is a value obtained by comparing two dimensions' values [19]. The value of the dimensionless index is determined by the nature or shape of the probability density function of vibration signal amplitude. The change of working condition has less effect on the dimensionless index, which is helpful to the time domain analysis of fault diagnosis. At the same time, the dimensionless index is a ratio, which has little to do with the sensitivity and magnification of the vibration detector. So the monitoring system does not need to be calibrated, which brings convenience to the fault diagnosis of the actual equipment. The accuracy of traditional dimensionless fault diagnosis is high for single fault diagnosis, but for complex fault diagnosis, it needs to be further improved [20, 21]. In order to overcome the shortcomings of traditional dimensionless index construction and improve the accuracy of fault diagnosis of rotating units, scholars have put forward new algorithms for fault diagnosis.

Consider a random signal with its amplitude and probability density function denoted by x and $p(x)$, respectively. Using these notations, various types of dimensional indexes can be defined as follows [22]:

Average amplitude:

$$\bar{X} = \int_{-\infty}^{+\infty} |x| p(x) dx \quad (1)$$

Root mean square value:

$$X_{rms} = \sqrt{\int_{-\infty}^{+\infty} x^2 p(x) dx} \quad (2)$$

Root mean square amplitude:

$$X_r = \left[\int_{-\infty}^{+\infty} \sqrt{|x|} p(x) dx \right]^2 \quad (3)$$

Kurtosis:

$$\beta = \int_{-\infty}^{+\infty} x^4 p(x) dx \quad (4)$$

Maximum value:

$$X_{max} = \max(X) \quad (5)$$

Dimensional indexes are sensitive to early fault data but they are affected in a nonlinear manner with the increase in the degree of failure, which results in diagnosis error. Therefore, we take a ratio of two-dimensional indexes to form a dimensionless index. Dimensionless indexes can eliminate the nonlinear change effect of dimensional index values. Various dimensionless indexes can be defined as follows:

Waveform index:

$$S_f = \frac{\left[\int_{-\infty}^{+\infty} |x|^2 p(x) dx \right]^{1/2}}{\left[\int_{-\infty}^{+\infty} |x| p(x) dx \right]} = \frac{X_{rms}}{|\bar{X}|} \quad (6)$$

Pulse index:

$$I_f = \frac{\lim_{l \rightarrow \infty} \left[\int_{-\infty}^{+\infty} |x|^l p(x) dx \right]^{1/l}}{\left[\int_{-\infty}^{+\infty} |x| p(x) dx \right]} = \frac{X_{max}}{|\bar{x}|} \quad (7)$$

Margin index:

$$M_f = \frac{\lim_{l \rightarrow \infty} \left[\int_{-\infty}^{+\infty} |x|^l p(x) dx \right]^{1/l}}{\left[\int_{-\infty}^{+\infty} |x|^{1/2} p(x) dx \right]^2} = \frac{X_{max}}{X_r} \quad (8)$$

Peak index:

$$C_f = \frac{\lim_{l \rightarrow \infty} \left[\int_{-\infty}^{+\infty} |x|^l p(x) dx \right]^{1/l}}{\left[\int_{-\infty}^{+\infty} |x|^2 p(x) dx \right]^{1/2}} = \frac{X_{max}}{X_{rms}} \quad (9)$$

Kurtosis index:

$$K_v = \frac{\beta}{X_{rms}^4} \quad (10)$$

We can express all the dimensionless indexes using the general equation (11). In (11), different dimensionless index equation can be generated by choosing different values for parameters l and m .

$$\xi_x = \frac{\left[\int_{-\infty}^{+\infty} |x|^l p(x) dx \right]^{1/l}}{\left[\int_{-\infty}^{+\infty} |x|^m p(x) dx \right]^{1/m}} \quad (11)$$

Equation (11) shows that dimensionless index calculation is based on the probability density function of the input signal. Hence, dimensionless index is a ratio that is not affected by the absolute level of the signal.

2.2. Fault Diagnosis Method Based on Evidence Theory and Dimensionless Index. Inspired by [16], we have the following fault diagnosis process. For a fault diagnosis problem y , assume that there are L basic attributes represented as $x_i (i = 1, \dots, L)$. Define the set of L basic attributes as a source of evidence $E = \{x_1, \dots, x_L\}$. According to the above description, assume that every attribute has its own weight $w = \{w_1, \dots, w_i, \dots, w_L\}$. Here, w_i represents the important degree of i_{th} attribute. The evaluation results of every attribute $x_i (i = 1, \dots, L)$ can simply represent the following reliability distribution form:

$$S(x_i) = \{(F_n, \beta_{i,n}(x_i)), n = 0, \dots, N-1\}, \quad i = 1, \dots, L \quad (12)$$

Note that $\beta_{i,n}(x_i) \geq 0$, $\sum_{n=0}^{N-1} \beta_{i,n}(x_i) \leq 1$, and $\beta_{i,n}(x_i)$ represent the reliability of attribute x_i and evaluate result point to fault F_n .

We use β_n as the reliability of problem y diagnosis to F_n . β_n is the final reliability which fuses all the attribute

evaluation results. The following is to use the ER algorithm proposed by Yang et al. to fuse the information [16].

Let $m_{n,i}$ denote the basic probability assignment value of basic attribute x_i . Support diagnosis problem is F_n . Otherwise, $m_{F,i}$ denotes the basic probability assignment value of it not assigned to any of the fault types. The value of $m_{F,i}$ describes the degree of uncertainty. The basic probability assignment value can be obtained as follows:

$$m_{n,i} = w_i \beta_{n,i}, \quad n = 0, 1, \dots, N-1 \quad (13)$$

$$m_{F,i} = 1 - \sum_{n=1}^N m_{n,i} = 1 - w_i \sum_{n=1}^N \beta_{n,i} \quad (14)$$

$$\bar{m}_{F,i} = 1 - w_i \quad (15)$$

$$\tilde{m}_{F,i} = w_i \left(1 - \sum_{i=1}^N \beta_{n,i} \right) \quad (16)$$

$$m_{F,i} = \bar{m}_{F,i} + \tilde{m}_{F,i}, \quad i = 1, 2, \dots, L \quad (17)$$

It is easy to find that $m_{F,i}$ is decomposed to two parts: $\bar{m}_{F,i}$ and $\tilde{m}_{F,i}$. $\bar{m}_{F,i}$ is affected by the weight of attribute and $\tilde{m}_{F,i}$ is affected by the attribute by incomplete evaluation information.

To obtain the final diagnosis result, we apply Dempster combination rules directly to get the final evaluation results:

$$\{F_n\} : m_n = K_L \prod_{i=1}^L (m_{n,i} + \bar{m}_{F,i} + \tilde{m}_{F,i}) \quad (18)$$

$$- \prod_{i=1}^L (\bar{m}_{F,i} + \tilde{m}_{F,i})$$

$$\{F\} : \bar{m}_F = K_L \prod_{i=1}^L \bar{m}_{F,i} \quad (19)$$

$$K_L = \left[\sum_{n=1}^N \prod_{i=1}^L (m_{n,i} + \bar{m}_{F,i} + \tilde{m}_{F,i}) - (N-1) \prod_{i=1}^L (\bar{m}_{F,i} + \tilde{m}_{F,i}) \right] \quad (20)$$

$$\{F_n\} : \beta_n = \frac{m_n}{(1 - \bar{m}_F)} \quad (21)$$

Note that β_n are final reliability of fault diagnosis problem to F_n . Therefore, we can obtain the result simply:

$$n = \arg \max(\beta_n), \quad n = 1, 2, \dots, N \quad (22)$$

According to the following description, reliability and attribute weight play an important role in ER algorithm. The reliability based on correlation measure by set theory will be introduced in Section 3.1. For attribute weight and detailed derivation of ER, see [16].

3. Method Description

3.1. Traditional Reliability Calculation Method. The value of reliability obtained method has a very important effect on the evidence reasoning result. This is because the reliability value can accurately reflect the fault feature information and directly determines the fusion weight of each dimensionless index in the process of data fusion. In addition, reliability also directly affects the calculation of Dempster combination rules, such as basic probability assignment. The traditional idea is to determine the reliability based on the distance between the input data and the average value of the data. This method can be used in low overlap data circumstances. But, for high-overlap dimensionless index data, we are willing to use correlation measure by set theory, because it is only affected by upper and lower bounds of index values.

The dimensionless index ξ_i can be denoted as interval form $[x_i, \bar{x}_i]$. When we want to obtain the correlation between $[x_i, \bar{x}_i]$ and $[c_{i,j}, d_{i,j}]$, It can be directly generated by

$$\alpha_{i,j}(x_i) = \frac{|[x_i, \bar{x}_i] \cap [c_{i,j}, d_{i,j}]|}{|[x_i, \bar{x}_i]| + |[c_{i,j}, d_{i,j}]| - |[x_i, \bar{x}_i] \cap [c_{i,j}, d_{i,j}]|} \quad (23)$$

In (23), $|[a, b]|$ denotes the length of set. Then, reliability can be obtained:

$$\beta_{i,j}(x_i) = \frac{\alpha_{i,j}(x_i)}{\sum_{j=0}^3 \alpha_{i,j}(x_i)} \quad (24)$$

Equation (23) shows that the reliability calculation method depends entirely on the interval value between the two groups of calculated data. This method will lose some information about the structure and correlation of the data, which leads to an inaccurate calculation of reliability. Therefore, a natural idea in the improved reliability design method is to add a regularization condition after the reliability calculation method. The new method should be able to represent the relationship between the two sets of data. In this paper, the calculation of the correlation coefficient is used to obtain the information inside the data.

3.2. Improved Reliability Method

Assumption 1. Given two sets of independent and identically distributed data sets of the same length X, Y , the two sets of data are matched one by one to form pairs. At first, we sort X_i and $Y_i (i = 1, 2, \dots, n)$. The notation $(X_{(i)}, Y_{[i]})$ denotes the pairs that are sorted based on values of X_i , such that $X_1 < X_2 < \dots < X_n (i = 1, 2, \dots, n)$. Additionally, $Y_{[1]}, Y_{[2]}, \dots, Y_{[i]}, (i = 1, 2, \dots, n)$ represent the values of Y paired with unsorted points X_i . Similarly, we can obtain $(X_{[i]}, Y_{(i)})$.

Definition 2. Based on the sorted data pairs, Gini correlation coefficient can be defined as [18]

$$r_G(Y, X) = \frac{(1/n(n-1)) \sum_{i=1}^n (2i-1-n) Y_{[i]}}{(1/n(n-1)) \sum_{i=1}^n (2i-1-n) Y_{(i)}} \quad (25)$$

$$r_G(X, Y) = \frac{(1/n(n-1)) \sum_{i=1}^n (2i-1-n) X_{[i]}}{(1/n(n-1)) \sum_{i=1}^n (2i-1-n) X_{(i)}} \quad (26)$$

In (25) and (26), n denotes the number of points in a data set. From (25) and (26), we can note that $r_G(Y, X) \neq r_G(X, Y)$. Hence, we defined a symmetric Gini correlation (SGC) coefficient as defined in the following equation:

$$r_G^*(X, Y) = \frac{1}{2} [r_G(Y, X) + r_G(X, Y)] \quad (27)$$

The correlation coefficient has the following properties: (1) correlation coefficient lies in the interval $[-1, +1]$; (2) the correlation between X and Y is a positive correlation or a negative correlation if the sign of the correlation coefficient is positive or negative, respectively; (3) if the correlation coefficient is 0, then X and Y are uncorrelated; (4) if magnitude of the correlation coefficient value is close to 1, it implies that the correlation between X and Y is stronger.

It can be seen from the above correlation equation that the Gini correlation coefficient calculation method is relatively simple, which provides the condition for real-time fault diagnosis. Gini correlation coefficient is more stable than other classical correlation coefficients in dealing with nonlinear data [18].

According to the correlation measure of the set, the same dimensionless index of different fault types is considered as the regularization based on the set correlation metric. The composition equation is as follows:

$$\hat{\alpha}_{i,j}(x_i) = \lambda_1 \alpha_{i,j}(x_i) + \lambda_2 r_G^*(x_i) \quad (28)$$

Then, the new reliability can be generated by

$$\hat{\beta}_{i,j}(x_i) = \frac{\hat{\alpha}_{i,j}(x_i)}{\sum_{j=0}^n \hat{\alpha}_{i,j}(x_i)} \quad (29)$$

According to the new reliability calculation equation, it can be found that the accuracy calculated in the same recognition frame is not lower than the accuracy of old reliability.

4. Experiment

4.1. Experiment Data. Experimental data is collected from large rotating machinery in petrochemical fault diagnosis experiment platform of multistage centrifugal fan fault diagnosis unit. The fault diagnosis unit consists of 11KW 5-stage centrifugal blower plus transmission, torque sensor, inverter motor, and several failure axes, tooth, and bearing members. The fault diagnosis unit can simulate common fault in multistage centrifugal blower unit. EMT390 data acquisition probe is placed in a position denoted with label "F" as shown in Figure 1. At the same time, the experimental data is read and stored using the Guangdong Provincial Key Laboratory system software. The originally collected data comprises the chassis vibration acceleration values. Since the different location of fault can have different effect on the operation of the entire axis, we can obtain the fault type by analyzing the chassis vibration acceleration information.

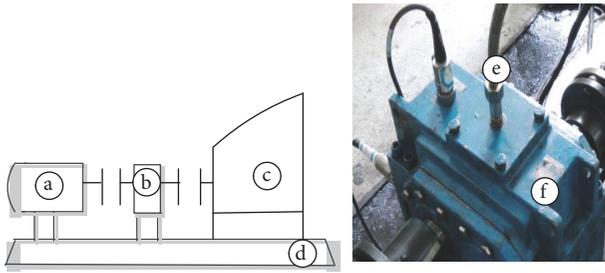


FIGURE 1: Petrochemical large-scale rotating equipment fault diagnosis experimental platform and data acquisition chassis location physical map. (a) Motor, (b) coupling or gearbox, (c) fans, (d) platform base, (e) oil tube, and (f) data acquisition probe placement.

4.2. Fault Diagnosis Model. Before data acquisition, the fault type, fault combination, motor speed and so on are determined. Then the lab staff change the normal parts of the unit and replace the corresponding fault parts according to the type of fault. Turn the machine on to a specified speed of 1000 rpm. Then the vibration acceleration of the housing is collected by the EMT390 data collector in the specified position. In order to facilitate the cross-use of data validation and diagnosis method, a fault type data acquisition process is completed by two people, each collecting two groups of data. The process is shown in Figure 2.

The vibration acceleration of all fault types is stored in the fault data folder, and 46 sets of data stored in one folder are read out by the data-reading program. There are 1024 vibration acceleration values in each group of data. In the process of dimensionless index calculation, five different dimensionless index values are calculated for 1024 vibration acceleration values. Therefore, each set of fault data contained 46×5 dimensionless values.

The fault diagnosis model is divided into five steps. First, the raw data are collected on the large petrochemical unit in Guangdong Petrochemical Equipment Fault Diagnosis Laboratory. Second, the dimensionless processing is used to extract the eigenvalues of the original data. Third, according to the composite degree of fault type, it can be divided into single fault and composite fault. Fourth, the input fault data is determined according to the fault type of fault data within the identification framework. Fifth, the fusion results are obtained to determine the diagnostic results. The specific steps can be described as follows.

Step 1. Determining the type of fault to be collected.

Step 2. Replacing the normal petrochemical unit parts to the designated fault parts.

Step 3. Electrifies the motor and debugs to 1000 rpm.

Step 4. Data acquisition personnel use EMT390 to collect vibration acceleration of the housing.

Step 5. Using data management software to read the sensor data and save it on the computer.

Step 6. Using MATLAB program to read the collected data and convert them into five dimensionless indexes.

Step 7. Calculating the initial reliability according to the correlation measure method of five dimensionless index sets.

Step 8. Calculating the correlation coefficient according to the five dimensionless index values to obtain the regularization term.

Step 9. Setting the parameter value to obtain the new reliability.

Step 10. Calculating the weight of each dimensionless index according to the result of new reliability calculation.

Step 11. Fusing reliability and weight according to Dempster combination rule.

Step 12. Finding out the fault type corresponding to the maximum reliability of the four fusion results.

In the process of experiment, we need to establish a fault identification framework and train the optimal parameters λ_1 and λ_2 corresponding to the framework by collecting multiple groups of data. So, in the experiment, the diagnosis effect of each recognition frame is optimized. When the optimization reaches a certain effect, we begin to consider the linkage diagnosis in each recognition frame. When unknown fault data are input, the optimal diagnosis results can be found in each identification framework. Therefore, the establishment of a relatively complete identification framework library is necessary for the practical application of the fault diagnosis method in the industry. Figure 3 is the basic structure of the recognition framework library.

The types of failures used in the experiment include two types, one fault, and more than two complex faults. In the experiment on this paper, we use three recognition frameworks. Identify faults included in frame 1: outer ring wear, inner ring wear, normal, and left bearing outer ring wear. The faults included in frame 2 include large gear missing teeth and left bearing outer ring wear composite failure, large gear missing teeth and left bearing missing ball composite failure, large and small gear missing teeth and outer ring wear composite failure, and large and small gear missing teeth and inner ring wear and composite failure. The faults included in frame 3 include large gear missing tooth, large and small gear missing tooth composite failure, bearing lacking ball, large gear missing tooth, and a left bearing inner ring wear composite failure.

4.3. Fault Diagnosis Result. According to Figures 4 and 5, the accuracy of fault diagnosis obtained by using the traditional method is 58.3% on identification frame 1 and 58.3% on identification frame 2. The accuracy of fault diagnosis obtained by using the proposed method is 75.0% on recognition frame 1 and 66.7% on recognition frame 2. Recognition frame 3 combines single fault and composite fault. Fault diagnosis result is shown in Figure 6. The accuracy of the method in [17] is 58.3% and the accuracy of the improved method is

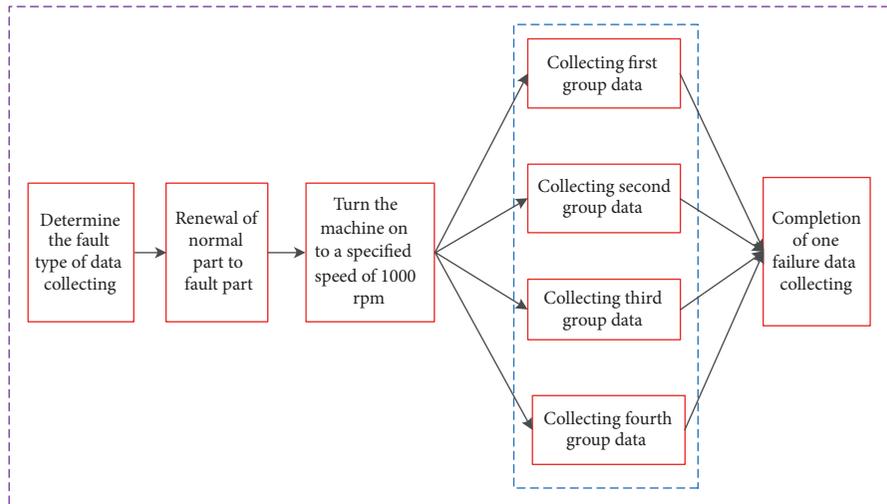


FIGURE 2: Experiment data collecting process.

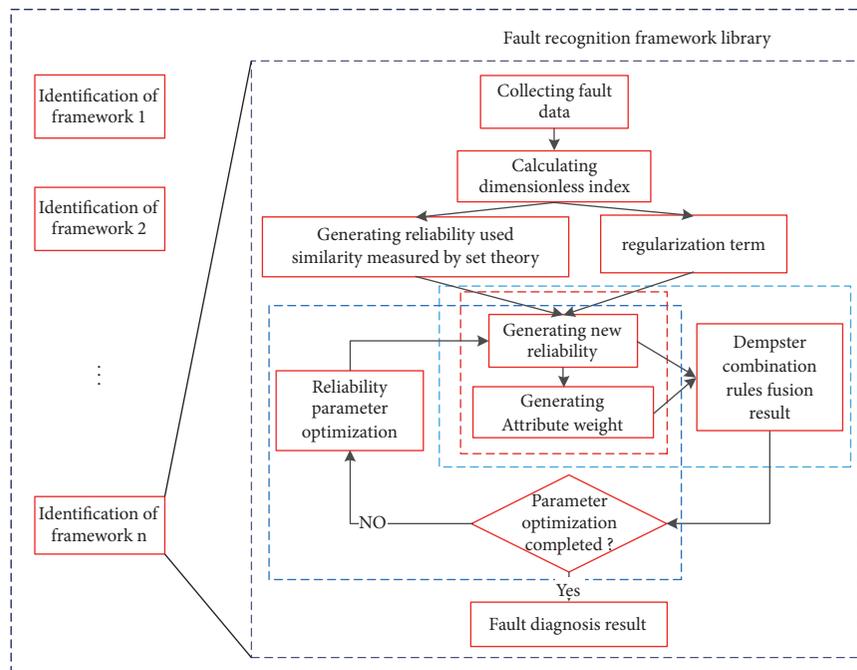


FIGURE 3: The basic structure of recognition frame library.

75.0%. The results show that the total diagnostic accuracy of traditional reliability calculation method is 58.3%, and that of the improved algorithm is 72.23%. The overall diagnostic accuracy has been greatly improved. From the fault type of error diagnosis, the main fault in identifying frame 1 is outer ring wear, while in frame 2, the main fault is the composite fault of large and small gear missing teeth and inner ring wear of left bearing. The fault in recognition frame 3 is the large gear tooth-missing fault.

It can be seen intuitively from the three diagrams of the experimental results that the actual fault identification effect of the improved evidence reasoning method of recognition frame 1, recognition frame 2, and recognition frame 3 is

better than [17]. The feasibility and accuracy of the proposed method in practical operation are verified. In terms of the overall diagnosis effect, the diagnosis effect of a single fault is better than that of a complex fault. This is because the information carried by the data collected by the single fault is easy to identify and distinguish. Because its fault data represent many features of fault, complex fault is prone to misdiagnosis.

5. Conclusion

In traditional evidence reasoning and dimensionless indexes combining fault diagnosis method, the diagnosis result is

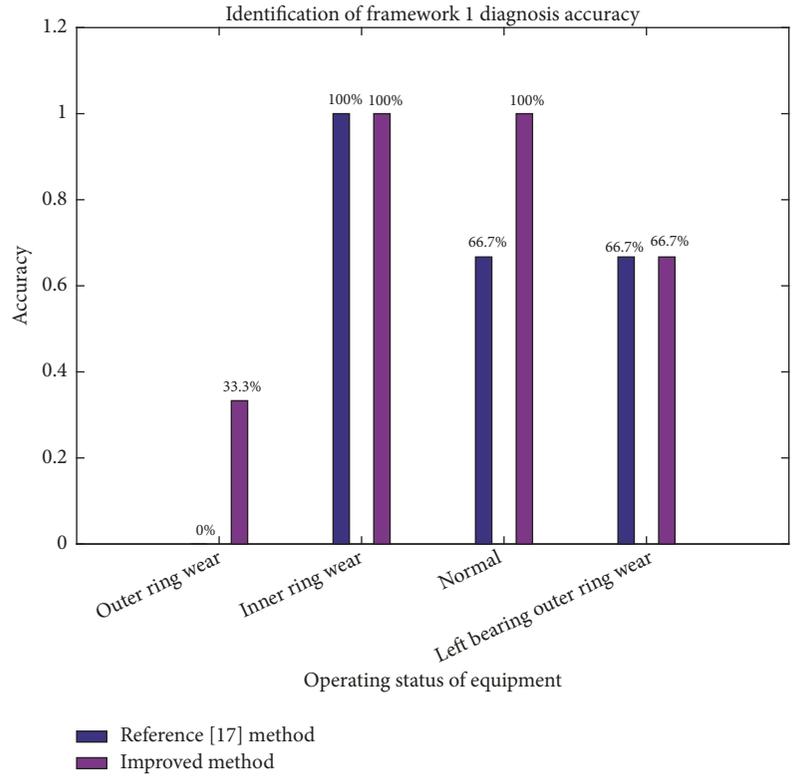


FIGURE 4: Diagnostic accuracy of each fault type in frame 1.

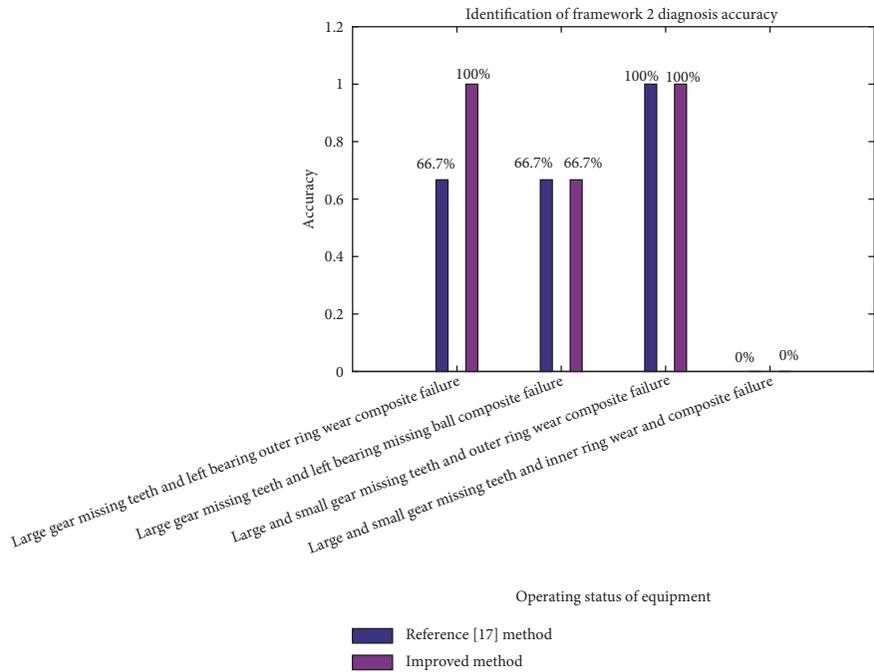


FIGURE 5: Diagnostic accuracy of each fault type in frame 2.

often wrong. This is largely due to the coincidence of the dimensionless indicators of different fault data. Because the reliability calculated based on dimensionless index is not accurate when the dimensionless index overlaps between

different faults, in this paper, an improved evidence reasoning method based on reliability regularization is proposed. The reliability regularization is mainly realized by calculating the Gini correlation coefficient between data. For the reason that

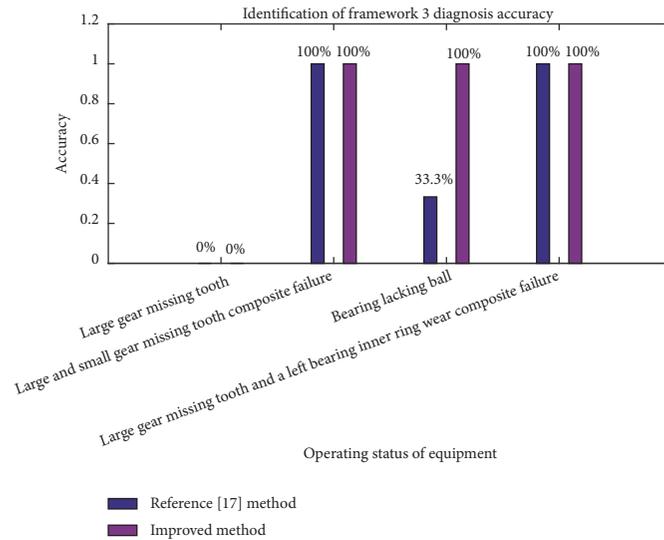


FIGURE 6: Diagnostic accuracy of each fault type in frame 3.

Gini correlation coefficient has a good ability to judge the nonlinear data, the regularized reliability value can better reflect the relationship between the two groups of data and obtain a more practical evaluation result. The experimental results show that the improved reliability method is closer to the actual fault situation and the diagnostic accuracy is improved.

Data Availability

All of the data used in this paper is collected from a large rotating machine in Guangdong Provincial Petrochemical Equipment Fault Diagnosis Key Laboratory. It could be obtained by contacting the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 61473331, in part by the Science and Technology Plan of Guangdong Province of China under Grant 2017A070712024, in part by the Sail Plan Training High-Level Talents of Guangdong Province of China, in part by the Introduction of Talents Project of Guangdong Polytechnic Normal University of China under Grant 991512203, in part by the 2016 Annual Scientific and Technological Innovation Special Fund to Foster Students Projects of Guangdong under Grant pdjh2016b0341, in part by the Guangdong University of Petrochemical Technology College Students Innovation Incubation Project under Grant 2015pyA006, in part by the Science and Technology Project of Guangzhou under Grant 201604010099, Grant 2016B030306002, and Grant 2016B030308001, in part by the Fundamental Research Funds for the Central Universities

under Grant x2jqD2170480, in part by Guangdong Province Science and Technology Major Special Projects under Grant 2017B030305004, by Guangdong Province Science and Technology Application of Major Special Projects under Grant 2016B020243011, and by Major Provincial Scientific Research Projects of Guangdong Normal Universities under Grant 2017KZDXM052.

References

- [1] Z. Gao, C. Cecati, and S. X. Ding, "A survey of fault diagnosis and fault tolerant techniques-Part II: fault diagnosis with knowledge-based and hybrid/active approaches," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 6, pp. 3768–3774, 2015.
- [2] J. Xiong, Q. Zhang, J. Wan, L. Liang, P. Cheng, and Q. Liang, "Data fusion method based on mutual dimensionless," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 2, pp. 506–517, 2018.
- [3] T. Bai, L. Zhang, L. Duan, and J. Wang, "NSCT-based infrared image enhancement method for rotating machinery fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 10, pp. 2293–2301, 2016.
- [4] S. Lu, J. Guo, Q. He, F. Liu, Y. Liu, and J. Zhao, "A novel contactless angular resampling method for motor bearing fault diagnosis under variable speed," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 11, pp. 2538–2550, 2016.
- [5] X. Zhang, L. Tang, and J. Decastro, "Robust fault diagnosis of aircraft engines: A nonlinear adaptive estimation-based approach," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 3, pp. 861–868, 2013.
- [6] A. E. Ashari, R. Nikoukhah, and S. L. Campbell, "Active robust fault detection in closed-loop systems: quadratic optimization approach," *IEEE Transactions on Automatic Control*, vol. 57, no. 10, pp. 2532–2544, 2012.
- [7] A. Grastien and A. Anbulagan, "Diagnosis of discrete event systems using satisfiability algorithms: a theoretical and empirical study," *Institute of Electrical and Electronics Engineers*

- Transactions on Automatic Control*, vol. 58, no. 12, pp. 3070–3083, 2013.
- [8] A. Bouchachia, “Adaptive computational intelligence for dynamical systems,” in *Intelligence for Nonlinear Dynamics and Synchronisation*, vol. 3, pp. 3–20, Atlantis: Atlantis Press, France, 2010.
- [9] A. H. Chen, L. L. Jiang, Y. L. Liu, and X. J. Li, “Fault diagnosis expert system model based on knowledge grid,” *Chinese Journal of Scientific Instrument*, vol. 30, no. 11, pp. 2450–2454, 2009.
- [10] D. L. Qin, *Research on Intelligent Diagnosis for Equipment Fault Based on Knowledge Management. [Ph.D. thesis]*, Hunan University, 2014.
- [11] S. Ding, S. Yin, K. Peng, H. Hao, and B. Shen, “A novel scheme for key performance indicator prediction and diagnosis with application to an industrial hot strip mill,” *IEEE Transactions on Industrial Informatics*, vol. 9, no. 4, pp. 2239–2247, 2013.
- [12] J. Chen and R. Patton, *Robust model-based fault diagnosis for dynamic systems*, Kluwer, Boston, Mass, USA, 1999.
- [13] H. Chen and S. Lu, “Fault diagnosis digital method for power transistors in power converters of switched reluctance motors,” *IEEE Transactions on Industrial Electronics*, vol. 60, no. 2, pp. 749–763, 2013.
- [14] J. Xiong, Q. Zhang, Z. Peng, G. Sun, W. Xu, and Q. Wang, “A diagnosis method for rotation machinery faults based on dimensionless indexes combined with k-nearest neighbor algorithm,” *Mathematical Problems in Engineering*, vol. 2015, Article ID 563954, 9 pages, 2015.
- [15] J. Xiong, Q. Zhang, Z. Peng, G. Sun, and Y. Cai, “Double Sample data fusion method based on combination rules,” *IEEE Access*, vol. 4, no. 99, pp. 7487–7499, 2016.
- [16] J. B. Yang and D. L. Xu, “Nonlinear information aggregation via evidential reasoning in multiattribute decision analysis under uncertainty,” *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 32, no. 3, pp. 376–393, 2002.
- [17] Q. H. Zhang, G. X. Sun, and X. S. Si, “A complex Fault diagnosis method based on the combination of artificial immune and evidence Theory,” Article ID 103678881, 2014.
- [18] W. Xu, Y. S. Hung, M. Niranjana, and M. Shen, “Asymptotic mean and variance of Gini correlation for bivariate normal samples,” *IEEE Transactions on Signal Processing*, vol. 58, no. 2, pp. 522–534, 2010.
- [19] Q. Zhang, A. Qin, L. Shu, G. Sun, and L. Shao, “Vibration sensor based intelligent fault diagnosis system for large machine unit in petrochemical industries,” *International Journal of Distributed Sensor Networks*, vol. 2015, no. 3, pp. 1376–1381, 2013.
- [20] Q. H. Zhang, Q. Hu, G. X. Sun, X. S. Si, and A. S. Qin, “Concurrent fault diagnosis for rotating machinery based on vibration sensors,” *International Journal of Distributed Sensor Networks*, vol. 75, no. 8, Article ID 472675, pp. 93–102, 2015.
- [21] G. H. B. Foo, X. Zhang, and D. M. Vilathgamuwa, “A sensor fault detection and isolation method in interior permanent-magnet synchronous motor drives based on an extended kalman filter,” *IEEE Transactions on Industrial Electronics*, vol. 60, no. 8, pp. 3485–3495, 2017.
- [22] Q. H. Zhang, *Machine unit fault diagnosis base on artificial immune system [Ph.D. thesis]*, South China University of Technology, 2004.

