Received August 16, 2016, accepted August 26, 2016, date of publication September 7, 2016, date of current version November 18, 2016.

Digital Object Identifier 10.1109/ACCESS.2016.2604824

Double Sample Data Fusion Method Based on Combination Rules

JIANBIN XIONG[1,2], (Member, IEEE), QINGHUA ZHANG[1,2], ZHIPING PENG[1,2], GUOXI SUN[1,2], AND YONGDA CAI[1]

[1]School of Computer and Electronic Information, Guangdong University of Petrochemical Technology, Maoming 525000, China
[2]Guangdong Provincial Key Laboratory of Petrochemical Equipment Fault Diagnosis, Maoming 525000, China

Corresponding author: J. Xiong (276158903@qq.com)

This work is supported in part by the National Natural Science Foundation of China under Grant 61473331, Grant 61471133, Grant 61174113, and Grant 61272382, in part by the Natural Science Foundation of Guangdong Province of China under Grant 2014A030307049, in part by the Ordinary University Innovation of Guangdong Province of China under Grant 2015KTSCX094, in part by the Sail Plan Training High-level Talents of Guangdong Province of China, in part by the Science and Technology Plan of Guangdong Province of China under Grant 2015B020233019, in part by the Guangdong High-Level Personnel of Institutions of Higher Learning under Grant 2013 246.152, in part by the Scientific Research Foundation of Discipline and Specialty Construction in Higher Education of Guangdong, under Grant 2013KJCX0133, in part by the 2016 Annual Scientific and Technological Innovation Special Fund to Foster Students Projects of Guangdong, under Grant pdjh2016b0341, and in part by the 2015 Guangdong University of Petrochemical Technology College Students' Innovation Incubation Project under Grant 2015pyA006.

ABSTRACT This paper proposes a double sample data fusion method based on combination rules to improve the classification of dimensionless indices in petrochemical rotating machinery equipment. This method first collects the original data and counts the mutual dimensionless index as the body of evidence. The reliability of the body of evidence is then determined using a distance calculation method. Finally, the evidence reasoning method is used to fuse the mutual dimensionless index data based on reliability, and the type of fault is detected using the K-S test. A real-time data collection experiment shows that this method can identify the fault type for mutual dimensionless indices that have the appearance of coincidences or evidence conflicts. The experimental results also show that this method has a stronger ability to diagnose faults when compared with the K-nearest neighbor method, and exhibits an accuracy improvement of 9.45%.

INDEX TERMS Data fusion, fault diagnosis, mutual dimensionless index, evidence reasoning, K-S test.

I. INTRODUCTION

Rotating machinery (such as rotary bearings, steam turbines, compressors, and fans) is playing an increasingly important role in key equipment for important engineering sectors, including petroleum, the chemical industry, metallurgy, machinery manufacturing, and aerospace. This machinery has been developed to operate well under normal working conditions within the related industries. However, since rotating machinery equipment always works for an extended period of time, some mechanical wear and tear might occur and result in equipment faults. Therefore, it is very important to perform fault diagnosis on this type of equipment. Since the structure and processes of rotating machinery equipment are usually complex, multiple faults generally occur [1]. The characteristics of multiple different faults are difficult to distinguish because there is some uncertainty in the collected data when multiple faults occur. In other words, the diagnosis of multiple faults is a complex problem due to the correlation and uncertainty of faults, and this makes successful multiple fault diagnosis methods difficult to achieve [2]. When rotating machinery equipment breaks down, the most important characteristic is that the machine exhibits abnormal vibrations and noise. The fault information is reflected in the vibration signals in the amplitude domain, time domain and frequency domain [3]. The most basic and original methods operate on the vibration signal in the time domain, and it can be very beneficial to maintain the basic characteristics of a signal if the time domain signal can be used directly for the fault diagnosis. A probability density function can be used to derive the dimensional indices (such as the average or root mean square value) and dimensionless indices (such as the waveform index, margin index, or pulse index etc.) [4], [5]. The dimensional indices are general characteristics that are fault sensitive, but they are susceptible to outside interference. Conversely, the dimensionless indices are not sensitive to outside interference, but an overlap or coincidence exists between the fault coverage. In other words, the dimensionless indices have ranges with a common overlap for normal and abnormal equipment, which leads to failures in correctly diagnosing faults.

To deal with this problem, it is very important to use an effective data fusion method and fault diagnosis method. Traditionally, D-S theory is used [6], [7], which can accurately describe significant concepts including "uncertain" and "unknown" information. However, the sensor may exhibit mutual interference [8] due to natural or human interference in the actual information fusion system. Therefore, traditional D-S theory cannot deal with conflicting evidence effectively. Jianbin Xiong et al. presented four types of filtering methods [9] to solve the problem when dimensionless index volatility is large and the scope is difficult to determine. However, these methods are not able to resolve highly conflicting evidence sufficiently. Another method that combines a dimensionless index, evidence theory, and the K-Nearest Neighbor (KNN) method [10] to deal with conflicting evidence in fault diagnosis has been proposed. However, this method is not suitable for multiple fault diagnosis. Additionally, a method has been proposed that combines a static discounting factor with KNN [11] for fault diagnosis, and this improves the efficiency of the fault diagnosis by fusing the conflicting evidence in order to correct traditional methods. However, this method does not consider external factors that can have dynamic effects on the sensor. Therefore, the potential for interference in the data fusion becomes the key problem for accurate fault diagnosis.

A data fusion method based on evidence reasoning (ER) can solve the problem of interference in information fusion. Yang [12] performed research on quantitative and qualitative information distributed to a belief structure transformation problem, and proposed a conversion technology based on rule and utility information. Another study [13] improved the ER by defining general rules that information fusion methods need to satisfy, and Yang and Xu [14] analyzed the nonlinear characteristics of ER. Xiaosheng Si et al. [15] improved the reliability calculation method and built a new prediction model based on a reliability rule base. Changhua Hu et al. [16] proposed a reliability prediction model based on dynamic evidential reasoning that considered the time influence on reliability. This experiment showed that ER can accurately process the interference information. In 2014, Aisong Qin et al. used ER in rotating machinery equipment [17] and effectively improved the recognition rate for fault diagnosis. Until now, the ER algorithm has been the best nonlinear data fusion method for handling inaccurate, incomplete, or ambiguous data or random data. It has been widely used in fault forecasting, reliability prediction, multi-attribute decision analysis, environmental impact assessment, and pipeline detection. However, since the ER algorithm requires the use of reliability and weight and the formulas used in counting are flexible, the data fusion will be influenced when the counting formula is unsuitable.

For the method proposed in this paper, the mutual dimensional indices are first calculated. These include the mutual waveform index, the mutual peak index, the mutual pulse index, the mutual margin index, and the mutual kurtosis index. By obtaining a relatively accurate body of evidence, we can then identify the weight of different mutual dimensionless indices based on the body of evidence. Finally, the ER combination rule and the K-S test are used for fault diagnosis. In this paper, the creative combination of the K-S test with the ER method exploits the advantages of both methods to obtain a new data fusion method.

The remainder of this paper is organized as follows. Section II introduces the mutual dimensionless index, K-S test, ER, and builds an algorithm model. In section III, details of the experimental environment, steps, and results are given. In section IV, we summarize our main findings and provide conclusions.

II. RELATED DEFINITION AND CONSTRUCTION OF DIAGNOSIS MODEL

A. RELATED DEFINITIONS

1) NON-DIMENSIONAL INDEX DEFINITION

Since traditional dimensionless indices exhibit significant overlap between normal operation and abnormal operation, we propose a new method to calculate the mutual dimensionless index. This method can narrow the distance between the dimensionless internal structures of the index, and reduce the overlap of dimensionless indices of different fault types. Therefore, the fault diagnosis accuracy can be improved by using this mutual dimensionless index. The index calculation method can be described as follows:

Assumption 1: There is a set of observed signals $z(k)$, which consists more than one thousand signals. We can use formula (1) to separate the set into three discrete parts $s(k-k_0)$, $x(k)$, and $v(k)$

$$z(k) = c[s(k-k_0) + x(k) + v(k)] \tag{1}$$

Let

$$y(k) = x(k) + v(k) \tag{2}$$

The observed signal sample can be defined as:

$$z(k) = c[s(k-k_0) + y(k)] \tag{3}$$

Definition 1: Random variables s and y have the probability density functions $p(s)$ and $p(y)$, respectively, and the general equation of the mutual dimensionless index is:

$$MX_{SFR} = \frac{[\int_{\Re} |y|^l p(y)dz]^{1/l}}{[\int_{\Re} |s|^m p(s)dz]^{1/m}} = \frac{\sqrt[l]{E(|y|^l)}}{\sqrt[m]{E(|s|^m)}} \tag{4}$$

where SFR is the Signal Fault Ratio.

If $l = 2$, $m = 1$, the mutual waveform index MS_{SFR} can be defined as:

$$MS_{SFR} = \frac{[\int_{\Re} |y|^2 p(y)dz]^{1/2}}{[\int_{\Re} |s| p(s)dz]} = \frac{\sqrt{E(|y|^2)}}{E(|s|)} \tag{5}$$

If $l \to \infty, m = 1$, the mutual pulse index MI_{SFR} can be defined as:

$$MI_{SFR} = \lim_{l\to\infty} \frac{[\int_{\Re} |y|^l p(y)dy]^{1/l}}{[\int_{\Re} |s| p(s)ds]} = \frac{\lim_{l\to\infty} \sqrt[l]{E(|y|^l)}}{E|s|} \quad (6)$$

If $l \to \infty, m = 1/2$, the mutual margin index MCL_{SFR} can be defined as:

$$MCL_{SFR} = \lim_{l\to\infty} \frac{[\int_{\Re} |y|^l p(y)dy]^{1/l}}{[\int_{\Re} |s|^{1/2} p(s)ds]^2} = \frac{\lim_{l\to\infty} \sqrt[l]{E(|y|^l)}}{[E(\sqrt{|s|})]^2} \quad (7)$$

If $l \to \infty, m = 2$, the mutual peak index MC_{SFR} can be defined as:

$$MC_{SFR} = \lim_{l\to\infty} \frac{[\int_{\Re} |y|^l p(y)dy]^{1/l}}{[\int_{\Re} |s|^2 p(s)ds]^{1/2}} = \frac{\lim_{l\to\infty} \sqrt[l]{E(|y|^l)}}{\sqrt{(E|s|^2)}} \quad (8)$$

The mutual kurtosis index MK_{SFR} can be directly defined as:

$$MK_{SFR} = \frac{\int_{\Re} y^4 p(y)dy}{[\int_{\Re} |s|^2 p(s)ds]^2} = \frac{E(|y|^4)}{[E(|s|^2)]^2} \quad (9)$$

2) BASIC DEFINITION OF K-S TEST

Assumption 2: There are two independent signals: a test signal $\psi(x_{(i)})$ and a reference signal $\phi(x_{(j)})$. Using an i, j description, the two signals of the original time series are arranged in ascending order. The statistical distance d between the two signals can be calculated by [18]:

$$d = \max_{i,j} |\psi(x_{(i)}) - \phi(x_{(j)})| \quad (10)$$

Definition 2: Under assumption 2, the probability similarity of the two signals can be defined by $\mu(\nu)$:

$$\mu(\nu) = \phi(d\sqrt{\frac{\eta_1\eta_2}{\eta_1 + \eta_2}}) \quad (11)$$

Formulas (12), (13), and (14) are used to calculate the ϕ, λ, and η_e:

$$\phi(\lambda) = 2\sum\nolimits_{i=1}^{\infty} (-1)^{i-1} e^{2j^2\lambda^2} \quad (12)$$

$$\lambda = d(\sqrt{\eta_e} + 0.12 + \frac{0.11}{\sqrt{\eta_e}}) \quad (13)$$

$$\eta_e = \frac{\eta_1\eta_2}{\eta_1 + \eta_2} \quad (14)$$

Here η_1, η_2 are the representative data points of the test signal and the reference signal, respectively. η_e represents the effective points, and it has been proven that the result will be more accurate and reliable if there are enough effective points [18]. From formula (11), the probability similarity value $\mu(d)$ will tend towards one when the two signals are similar. In contrast, if the two signals are different, the probability similarity value $\mu(d)$ will tend towards zero.

3) ER BASIC DEFINITION

According to reference [14], for a fault diagnosis question q, it can be assumed that there are N basic properties which can be expressed as $\alpha_i(i = 1, \cdots, N)$. In this paper, the basic properties are all mutual and dimensionless. These N basic properties of the collection can be defined as a source of evidence: $E = \{\alpha_1, \cdots, \alpha_N\}$. Assuming that the weight of the properties is $\xi = \{\xi_1, \cdots \xi_i, \cdots, \xi_N\}$, and ξ_i shows the relative importance of the ith position property α_i, the weight can be normalized as follows:

$$0 \le \xi_i \le 1, \quad \sum\nolimits_{i=1}^{N} \xi_i = 1 \quad (15)$$

The fault diagnosis question q corresponding to the fault type can be delimited, and ε satisfies the requirements of the evidence reasoning recognition framework (a complete set of mutually exclusive properties). This can be shown by:

$$\varepsilon = \{\varepsilon_0, \cdots, \varepsilon_{M-1}\} \quad (16)$$

Where ε_m expresses the nth position of the fault.

For each property $\alpha_i(i = 1, \cdots, N)$, the evaluation result can be listed using the reliability distribution formula:

$$\Gamma(\alpha_i) = \{\varepsilon_m, (\theta_{i,m}(\alpha_i)), m = 0, \cdots, M-1\}, \quad i = 1, \cdots N. \quad (17)$$

Let $\theta_{i,m}(\alpha_i) \ge 0, \sum_{m=0}^{M-1} \theta_{i,m}(\alpha_i) \le 1$, where $\theta_{i,m}(\alpha_i)$ expresses the reliability that property α_i is evaluated as a fault ε_m. If $\sum_{m=1}^{M} \theta_{m,i} = 1$, the assessment of property α_i is complete, otherwise it is incomplete. The characteristics of properties $e_i(i = 1, \ldots, L)$ determine how $\theta_{m,i}$ is calculated. For example, numerical expressions are used for the quantitative property data and form the distribution that expresses this data. In order to handle quantitative and qualitative data within a unified reliable framework, YANG proposed an equivalent information conversion technology which converts numerical data, random data, and qualitative information into a reliable form.

Let θ_n express the reliability of question q diagnosing ε_m, and fuse all of the properties $\alpha_i(i = 1, \cdots, N)$ to obtain ε_m and describe the evidence reasoning fusion information [11].

Let $\beta_{m,i}$ donate the basic probability assignment value that the ith basis property, α_i, supports question q diagnosing fault ε_m. $\beta_{\varepsilon,i}$ expresses the basic probability assignment value which has not been assigned to any type of fault, and its size expresses the degree of uncertainty. The basic probability assignment value can be obtained in the following manner:

$$\begin{aligned}
&\beta_{m,i} = \xi_i\theta_{m,i}, \quad m = 0, 1 \cdots, M-1. \\
&\beta_{\varepsilon,i} = 1 - \sum_{m=1}^{M} m_{m,i} = 1 - \xi_i \sum_{m=1}^{M} \theta_{m,i} \\
&\bar{\beta}_{\varepsilon,i} = 1 - \xi_i, \quad i = 1, 2, \cdots, N. \\
&\tilde{\beta}_{\varepsilon,i} = \xi_i(1 - \sum_{i=1}^{M} \theta_{m,i}). \\
&\beta_{\varepsilon,i} = \bar{\beta}_{\varepsilon,i} + \tilde{\beta}_{\varepsilon,i}, \quad i = 1, 2, \cdots, N.
\end{aligned} \quad (18)$$

It can be found that the unassigned basic probability $\beta_{\varepsilon,i}$ can be divided into two parts: $\bar{\beta}_{\varepsilon,i}$ and $\bar{\beta}_{\varepsilon,i}$. $\bar{\beta}_{\varepsilon,i}$ is due to the relative weight properties $\alpha_i(i = 1, \cdots, N)$, and $\tilde{\beta}_{\varepsilon,i}$ is due to properties $e_i(i = 1, \ldots, L)$, which are the incomplete assessment information.

Since fault types of question q satisfy the requirements of the evidence theory recognition framework, the ER algorithm and the Dempster combination rule can be used to obtain the final evaluation result shown in formula (19):

$$\begin{aligned}\{\varepsilon_m\}: \beta_m &= K_N[\prod\nolimits_{i=1}^{N}(\beta_{m,i}+\bar{\beta}_{\varepsilon,i}+\tilde{\beta}_{\varepsilon,i}) \\ &\quad - \prod\nolimits_{i=1}^{N}(\bar{\beta}_{\varepsilon,i}+\tilde{\beta}_{\varepsilon,i})] \\ \{\varepsilon\}: \tilde{\beta}_{\varepsilon} &= K_N[\prod\nolimits_{i=1}^{N}(\bar{\beta}_{\varepsilon,i}+\tilde{\beta}_{\varepsilon,i}) - \prod\nolimits_{i=1}^{N}\bar{\beta}_{\varepsilon,i}] \\ \{\varepsilon\}: \bar{m}_{\varepsilon} &= K_N\prod\nolimits_{i=1}^{N}\bar{\beta}_{\varepsilon,i} \\ K_M &= [\sum\nolimits_{m=1}^{M}\prod\nolimits_{i=1}^{N}(\beta_{m,i}+\bar{\beta}_{\varepsilon,i}+\tilde{\beta}_{\varepsilon,i}) \\ &\quad -(M-1)\prod\nolimits_{i=1}^{N}(\bar{\beta}_{\varepsilon,i}+\tilde{\beta}_{\varepsilon,i})]^{-1} \quad (19) \\ \{\varepsilon_m\}: \theta_m &= \frac{\beta_m}{1-\bar{\beta}_{\varepsilon}} \\ \{\varepsilon\}: \theta_m &= \frac{\tilde{\beta}_{\varepsilon}}{1-\bar{\beta}_{\varepsilon}}\end{aligned}$$

Here, θ_m and θ_ε denote that question q is diagnosed with ε_m and ε fusion reliability, respectively. Therefore, the overall assessment result can be shown to be $q = \{(\varepsilon_m, \theta_m), (\varepsilon, \theta_\varepsilon), m = 1, 2, \cdots, M\}$.

4) ER RELIABILITY CALCULATION AND WEIGHT CALCULATION

During the ER process, a method of calculating the reliability and weight is necessary. In this paper, the following calculation method is used.

If the range of the jth fault's mutual dimensionless index P_i is $[a_{i,j}, b_{i,j}]$, the median is used to replace the range. The actual calculation of the index's mutual dimensionless distance from the center is then calculated to obtain the reliability. In other words, the nearer to the center it is, the greater the reliability of the distribution will be. This method avoids mutual dimensionless coincidences of different faults to a certain extent.

For a composite fault diagnosis of two faults, a recognition framework $\varepsilon = \{\varepsilon_0, \varepsilon_1, \varepsilon_2, \varepsilon_3\}$ is created, where ε_0, ε_1, ε_2, and ε_3 express the normal behavior, fault A, fault B, and the combination of fault A and B, respectively. An example of the detailed process involved in reliably transforming p_i will be used for illustration. It will be assumed that the mutual dimensionless index is p_i, and the ranges of the jth fault's mutual dimensionless index P_i is $[a_{i,j}, b_{i,j}]$ (here j includes $\varepsilon_0, \varepsilon_1, \varepsilon_2, \varepsilon_3$, i.e. $j = 1, 2, 3$).

From the above description:

$$\vartheta_{i,j} = \frac{a_{i,j}+b_{i,j}}{2} \text{ means } \varepsilon_{\text{j}}, \quad i = 1, \cdots, N;\ j = 0, 1, 2, 3 \tag{20}$$

Each p_i can then be transformed in the following manner:

$$\Gamma(\alpha_i) = \{(\varepsilon_j, \theta_{i,j}(\alpha_i)), i = 1, \cdots, N; j = 0, 1, 2, 3\} \tag{21}$$

Each $\theta_{i,j}(\alpha_i)$ can be calculated using the following formulas:

$$\theta_{i,j}(\alpha_i) = \frac{\vartheta_{i,j+1}-\alpha_i}{\vartheta_{i,j+1}-\vartheta_{i,j}} \quad \text{if } \vartheta_{i,j} \le \alpha_i \le \vartheta_{i,j+1}, \quad j = 0, 1, 2, 3 \tag{22}$$

$$\theta_{i,j+1}(\alpha_i) = 1 - \theta_{i,j}(\alpha_i) \quad \text{if } \vartheta_{i,j} \le \alpha_i \le \vartheta_{i,j+1}, \quad j = 0, 1, 2, 3 \tag{23}$$

$$\theta_{i,s}(\alpha_i) = 0 \quad \text{for } s = 0, 1, 2, 3,\ s \ne j, j+1 \tag{24}$$

After solving the equation $\Gamma(\alpha_i) = \{(\varepsilon_j, \theta_{i,j}(\alpha_i)), i = 1, \cdots, N; j = 0, 1, 2, 3\}$, ER fusion data and fault diagnosis can be used.

The method for confirming the property weight $\xi = \{\xi_1, \cdots \xi_i, \cdots, \xi_N\}$ will be described next. According to the body of evidence $\Gamma(\alpha_i) = \{\varepsilon_m, (\theta_{i,m}(\alpha_i)), m = 0, \cdots, M-1\}, i = 1, \cdots N.$, each mutual dimensionless index and minority subordinate can be transformed using the majority principle. We can then identify the property weight vector as $\xi = \{\xi_1, \cdots \xi_i, \cdots, \xi_N\}$. The details are described as follows:

Step 1: The body of evidence $\Gamma(\alpha_i) = \{(\varepsilon_m, \theta_{i,m}(\alpha_i)), n = 0, \cdots, M-1\}$ is converted into vector form, such as $\beta_i = (\theta_{i,0}(\alpha_i), \theta_{i,1}(\alpha_i), \cdots \theta_{i,M-1}(\alpha_i))$.

Step 2: The distance l_{ij} between bodies of evidence is calculated from $\Gamma(\alpha_i) = \{(\varepsilon_m, \theta_{i,m}(\alpha_i)), m = 0, \cdots, M-1\}$ and $\Gamma(\alpha_j) = \{(\varepsilon_m, \theta_{j,m}(\alpha_j)), m = 0, \cdots, M-1\}$, and then the evidence distances are built from the matrix $L_{N\times N} = [l_{ij}]$:

$$L_{N\times N} = \begin{bmatrix} l_{11} & l_{12} & \cdots & l_{1N} \\ l_{21} & l_{22} & \cdots & l_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ l_{N1} & l_{N2} & \cdots & l_{NN} \end{bmatrix} \tag{25}$$

Each l_{ij} can be calculated by using:

$$l_{ij} = \sqrt{\frac{1}{2}(\beta_i-\beta_j)(\beta_i-\beta_j)^T} \tag{26}$$

where l_{ij} should satisfy $0 \le l_{ij} \le 1$.

Step 3: The similarity r_{ij} between evidence $\Gamma(\alpha_j) = \{(\varepsilon_m, \theta_{i,m}(\alpha_j)), m = 0, \cdots, M-1\}$ and $\Gamma(\alpha_j) = \{(\varepsilon_m, \theta_{j,m}(\alpha_j)), m = 0, \cdots, M-1\}$ is calculated based on the evidence distance l_{ij}, and then a matrix is built using the evidence $\Gamma_{N\times N} = [r_{ij}]$ as:

$$L_{N\times N} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1N} \\ r_{21} & r_{22} & \cdots & r_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ r_{N1} & r_{N2} & \cdots & r_{NN} \end{bmatrix} \tag{27}$$

Each r_{ij} can be calculated from:

$$r_{ij} = 1 - l_{ij} = 1 - \sqrt{\frac{1}{2}(\beta_i-\beta_j)(\beta_i-\beta_j)^T} \tag{28}$$

where r_{ij} should satisfy $0 \le r_{ij} \le 1$.

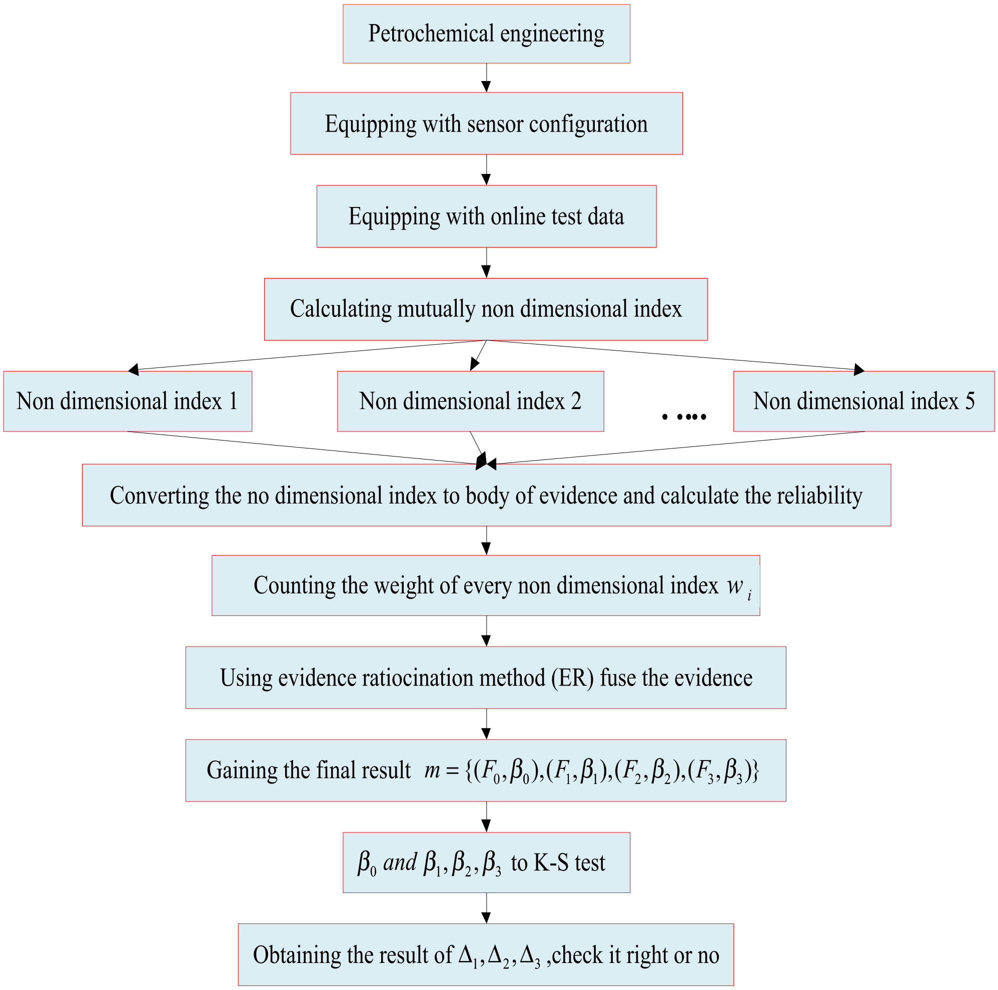


FIGURE 1. **The double sample data fusion model based on combination rules.**

Step 4: The total level of support, Γ_i of evidence $\Gamma(\alpha_i) = \{(\varepsilon_m, \theta_{i,m}(\alpha_i)), m = 0, \cdots, M-1\}$, is calculated using the following formula:

$$\Gamma_i = \sum\nolimits_{j=1}^{N} r_{ij} \tag{29}$$

Step 5: The base on total weight of support properties ξ_i is as follows:

$$\xi_i = \frac{\Gamma_i}{\sum_{j=1}^{N} \Gamma_i} \tag{30}$$

ER can be used to fuse data based on the calculation methods of reliability and weight.

5) SUMMARY RELATED DEFINITION

The K-S test can be used as a type of goodness of fit test to evaluate the similarity between two data samples. This method mainly uses ordered samples of a random variable to build a sample distribution function, and makes it possible to guarantee that a certain probability distribution function of another sample falls within a certain range. However, this method cannot resolve evidence conflicts effectively, and ER is needed to solve this problem. ER can fuse conflicting evidence and obtain the body of evidence, which is the easier test.

B. DOUBLE SAMPLE DATA FUSION METHOD BASED ON THE COMBINATION RULES OF MODEL BUILDING

The proposed method is divided into two parts: the first part collects fault data and the second part performs the experiments. Fault data collection is a basic step and, in this paper, equipment fault data of chassis vibration acceleration is obtained from real-time acquisition. 49 groups of data are used and each group contains 1024 vibration accelerations. The experimental process can be described as follows: first, a relatively accurate body of evidence is obtained by transforming the fault data into a mutual dimensionless index. Secondly, the mutual dimensionless index is used to calculate the reliability and weight of five mutual dimensionless indices. The reliability and weight are combined such that data fusion can be performed, and the different fault types are distinguished by using the K-S test. This process is shown in detail in Figure 1.

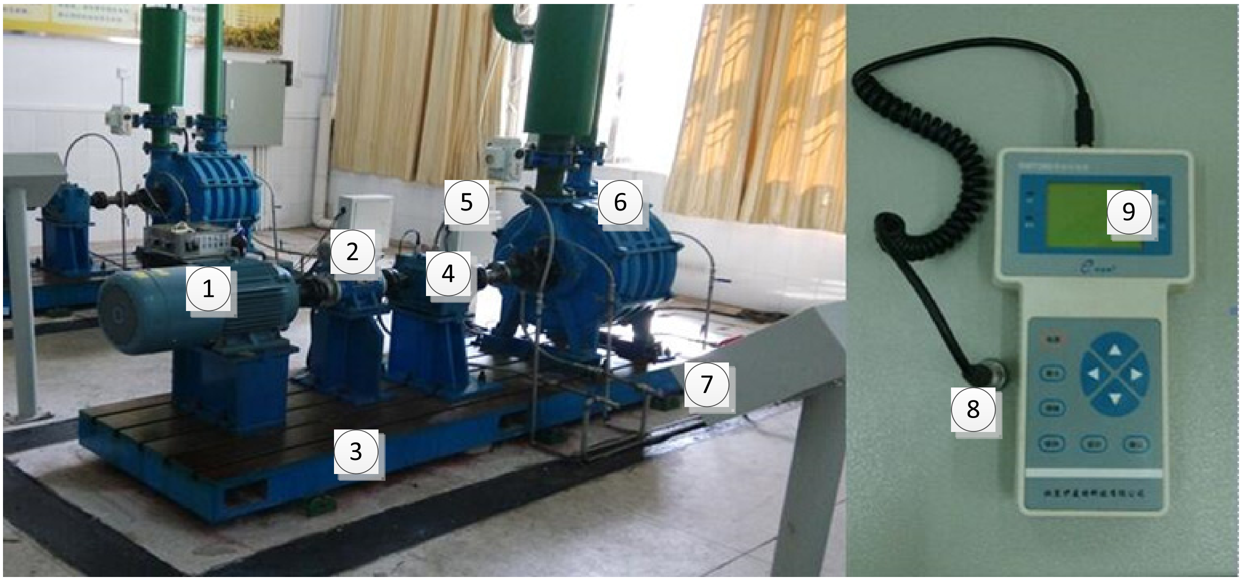


FIGURE 2. Petrochemical large rotary equipment fault diagnosis experiment platform and data Collector. It mainly consists of (1) electric motor, (2) gearbox, (3) base platform, (4) coupling, (5) oil pipe, (6) fan, (7) platform operation switch, (8) EMT390 sensor data acquisition probe, (9) EMT390 sensor data acquisition.

TABLE 1. Each number and each fault type for the bodies of evidence.

The data of the first data collector collecting		The data of the second data collector collecting	
Number	Fault type	Number	Fault type
a1	First collecting 49 groups of inner ring wear	f1	First collecting 49 groups of inner ring wear
a2	Second collecting 49 groups of inner ring wear	f2	Second collecting 49 groups of inner ring wear
b1	First collecting 49 groups of outer ring wear	g1	First collecting 49 groups of outer ring wear
b2	Second collecting 49 groups of outer ring wear	g2	Second collecting 49 groups of outer ring wear
c1	First collecting 49 groups of big gear teeth missing	h1	First collecting 49 groups of big gear teeth missing
c2	Second collecting 49 groups of big gear teeth missing	h2	Second collecting 49 groups of big gear teeth missing
d1	First collecting 49 groups of big gear teeth missing and inner ring wear combination	i1	First collecting 49 groups of big gear teeth missing and inner ring wear combination
d2	Second collecting 49 groups of big gear teeth missing and inner ring wear combination	i2	Second collecting 49 groups of big gear teeth missing and inner ring wear combination
e1	First collecting 49 groups of big gear teeth missing and outer ring wear combination	j1	First collecting 49 groups of big gear teeth missing and outer ring wear combination
e2	Second collecting 49 groups of big gear teeth missing and outer ring wear combination	j2	Second collecting 49 groups of big gear teeth missing and outer ring wear combination

III. EXPERIMENTS

A. EXPERIMENTAL ENVIRONMENT

This experiment was performed using a multistage centrifugal blower fault diagnosis unit at the Guangdong Provincial Key Laboratory, which diagnoses petrochemical equipment faults within a large petrochemical rotating machinery fault diagnosis test platform. The unit combines an 11kW 5-stage centrifugal blower with transmission, a torque sensor, an inverter motor, a standard plate, as well as various failure shafts, gears and bearings parts, as seen in Figure 2. The

TABLE 2. The results after evidence resoning form.

Zero fault type	1, 2, 3 fault group			Correct or incorrect	Zero fault type	1, 2, 3 fault group			Correct or incorrect
	a1	c1	d1			a2	c2	d2	
a2	0.8980	0.7959	0.8571	C	a1	0.8980	0.9184	0.8980	IC
c2	0.8776	0.7755	0.8776	IC	c1	0.8367	0.7959	0.8571	IC
d2	0.8980	0.8571	0.9184	C	d1	0.8776	0.8980	0.8980	IC
f1	0.8980	0.8163	0.8571	C	f1	0.9184	0.8980	0.8980	C
h1	0.8980	0.7755	0.8776	IC	h1	0.8980	0.8980	0.8571	IC
i1	0.7959	0.8163	0.8367	C	i1	0.8367	0.8367	0.8367	IC
f2	0.8163	0.8776	0.8571	IC	f2	0.7755	0.7959	0.8367	IC
h2	0.8980	0.7959	0.8776	IC	h2	0.8776	0.8776	0.8571	IC
i2	0.8776	0.8571	0.9184	C	i2	0.8776	0.8980	0.8571	IC
	f1	h1	i1			f2	h2	i2	
f2	0.8163	0.7959	0.7959	C	f1	0.7959	0.8776	0.8980	IC
h2	0.8776	0.9388	0.8163	C	h1	0.7959	0.9184	0.8571	C
i2	0.9184	0.8571	0.8571	IC	i1	0.7347	0.7959	0.8163	C
a1	0.8980	0.9184	0.7959	IC	a1	0.8163	0.9184	0.8776	IC
c1	0.8367	0.7755	0.8367	IC	c1	0.8571	0.7959	0.8571	IC
d1	0.8980	0.8776	0.8776	IC	d1	0.8163	0.8776	0.9184	C
a2	0.8571	0.8980	0.8367	IC	a2	0.7959	0.8980	0.8980	IC
c2	0.8776	0.9184	0.8367	C	c2	0.8163	0.8980	0.8776	C
d2	0.8980	0.8571	0.8367	IC	d2	0.8367	0.8776	0.8980	C

machine can simulate common faults that occur in multi-stage centrifugal blowers and their transmission. We use the EMT390 collector, developed by Beijing Yi Maite Technology Co., Ltd., to collect the various faults. The data is then saved in a data management system and an algorithm to calculate the dimensionless indices is implemented using MATLAB.

B. EXPERIMENTAL STEPS

Step 1: Some of the large petrochemical mechanical rotating equipment is replaced (e.g., the normal inner ring, outer ring, and big gear are replaced with a worn inner ring, a worn outer ring, and a big gear with teeth missing, respectively).

Step 2: To fix the data acquisition position, a label with a small circle that is the same size as the EMT390 sensor data acquisition probe is affixed to the equipment chassis.

Step 3: The EMT390 collector is used by two people to collect data twice, where each collection contains 49 groups of data and each group has 1024 vibration accelerations.

Step 4: MATLAB is used to read the collected data, and save it to the corresponding folder.

Step 5: MATLAB is used to calculate five dimensionless indices: the mutual waveform index, the mutual pulse index, the mutual margin index, the mutual peak index, and the kurtosis index. This data is then saved as the evidence source $E = \{\alpha_1, \cdots, \alpha_L\}$ for the ER.

Step 6: The center value of each mutual dimensionless index interval $[a_{i,j}, b_{i,j}]$ is found, and the reliability $\theta_{i,j}(\alpha_i)$ base is identified using the distance between each mutual dimensionless index and the center value.

Step 7: The weight $\xi = \{\xi_1, \cdots \xi_i, \cdots \xi_L\}$ of each evidence source is obtained using.

TABLE 3. The experimental results after evidence reasoning form.

Zero fault type	1, 2, 3 fault group			Correct or incorrect	Zero fault type	1, 2, 3 fault group			Correct or incorrect
	b1	c1	e1			b2	c2	e2	
b2	0.8980	0.8571	0.8571	C	b1	0.8776	0.8367	0.8571	C
c2	0.8367	0.7959	0.9388	IC	c1	0.8776	0.7755	0.8163	IC
e2	0.8776	0.8776	0.7755	IC	e1	0.8163	0.9184	0.7551	IC
g1	0.8163	0.8776	0.8571	IC	g1	0.8980	0.8571	0.8571	C
h1	0.8163	0.7551	0.9184	IC	h1	0.8367	0.9184	0.7551	C
j1	0.8980	0.8776	0.8367	IC	j1	0.9184	0.8571	0.8571	IC
g2	0.8571	0.8980	0.8571	IC	g2	0.8980	0.8571	0.8571	C
h2	0.8571	0.8163	0.9184	IC	h2	0.8571	0.8980	0.7755	C
j2	0.8571	0.8571	0.8980	C	j2	0.8776	0.8980	0.7959	IC
	g1	h1	j1			g2	h2	j2	
g2	0.8980	0.8367	0.8980	IC	g1	0.8980	0.8776	0.8980	IC
h2	0.8571	0.9184	0.8571	C	h1	0.8367	0.9184	0.8776	C
j2	0.8980	0.8776	0.8571	IC	j1	0.8980	0.8571	0.8776	IC
a1	0.8163	0.8367	0.8980	IC	a1	0.8367	0.8367	0.8980	IC
c1	0.8776	0.7755	0.8776	IC	c1	0.8980	0.7959	0.8367	IC
d1	0.8367	0.9184	0.8571	IC	d1	0.8571	0.9388	0.8776	IC
a2	0.8776	0.8367	0.8980	IC	a2	0.8980	0.8571	0.8776	C
c2	0.8367	0.9184	0.8571	C	c2	0.8367	0.8980	0.8776	C
d2	0.8776	0.7755	0.8776	IC	d2	0.8571	0.7755	0.8367	IC

Step 8: The composite formula of ER is used to calculate $\beta_{m,i}$, $\beta_{\varepsilon,i}$, $\bar{\beta}_{\varepsilon,i}$, $\tilde{\beta}_{\varepsilon,i}$, $\beta_{\varepsilon,i}$. The ER algorithm proposed by YANG is used, and the Dempster combination rule is applied to obtain the values of β_m, $\tilde{\beta}_\varepsilon$, $\bar{\beta}_\varepsilon$, K_i, θ_m, θ_ε. The fusion result is obtained from:

$$q = \{(\varepsilon_m, \theta_m), (\varepsilon, \theta_\varepsilon), m = 1, 2, \cdots, M\}$$

Step 9: The 0 fault and the 1, 2, and 3 fault for the K-S test are selected (here 0, 1, 2, 3 denote the fault type or the order of the group in the ER) to obtain the test result and judge whether or not the diagnosis result is true.

C. THE EXPERIMENTAL CONTRAST

In this experiment, two collectors are used to collect fault data at different time points. The data collections include the first and second inner ring wear, the outer ring wear, teeth missing on the big gear, a combination of teeth missing on the big gear teeth and inner ring wear, and a combination of teeth missing on the big gear teeth and outer ring wear in the same machine. Each collection has 49 groups and each group has 1024 numerical values. During the experiment, the collected data used to calculate the dimensionless indices is used as the evidence source for the ER. Therefore, 20 bodies of evidence are obtained. Alphanumeric codes are used to conveniently express the bodies of evidence, and these are detailed in Table 1.

Two data collectors are defined as the first step, and then the faults are collected including fault A, fault B, and fault A and fault B. These are combined to give twelve bodies of evidence that make up the associated body of evidence. For example, in Table 1, a1, a2, c1, c2, d1, d2, f1, f2, g1, g2, i1, and i2 comprise one associated body of evidence.

TABLE 4. The non-dimensional experimental results combined with the KNN algorithm.

Testing fault type	Practicing fault group					Correct or incorrect
	The first experiment					
	a1	b1	c1	d1	e1	
a2	Imaginary number					IC
b2	0.1820	0.2078	0.2333	0.1949	0.1821	IC
c2	0.1543	0.2178	0.2178	0.2178	0.1923	IC
d2	0.1013	0.2179	0.2038	0.2599	0.2171	C
e2	0.1235	0.2193	0.2057	0.2327	0.2189	IC
f1	0.2110	0.2253	0.2531	0.1557	0.1549	IC
g1	0.2277	0.2273	0.2124	0.1670	0.1657	IC
h1	Imaginary number					IC
i1	0.1320	0.2105	0.2108	0.2362	0.2105	C
j1	0.1128	0.2115	0.1975	0.2391	0.2391	IC
f2	0.2650	0.2026	0.2191	0.1563	0.1570	C
g2	0.1578	0.2233	0.2103	0.2107	0.1980	C
h2	0.1452	0.2235	0.2105	0.2234	0.1975	IC
i2	0.1320	0.2235	0.2235	0.2235	0.1975	IC
j2	Imaginary number					IC
	The second experiment					
	a2	b2	c2	d2	e2	
a1	Imaginary number					IC
b1	0.1582	0.2104	0.2234	0.1975	0.2105	IC
c1	0.1690	0.2329	0.2205	0.1824	0.1952	IC
d1	0.0949	0.2030	0.2296	0.2429	0.2295	C
e1	Imaginary number					IC
f1	0.2387	0.2534	0.1975	0.1554	0.1550	IC
g1	Imaginary number					IC
h1	0.1248	0.2181	0.2029	0.2352	0.2191	IC
i1	0.1320	0.2105	0.2235	0.2105	0.2235	IC
j1	0.1237	0.2057	0.2190	0.2191	0.2325	C
f2	0.2695	0.2386	0.1903	0.1589	0.1428	C
g2	0.1628	0.2247	0.2124	0.1877	0.2124	C
h2	0.1317	0.2230	0.2233	0.2244	0.1977	IC
i2	0.1304	0.1822	0.2335	0.2207	0.2333	IC
j2	Imaginary number					IC

Within this set, inner ring wear is fault A, teeth missing in the big gear is fault B, and the combination of teeth missing in the big gear and inner ring wear is the combination of faults A and B. As Table 2 shows, three fault data points are selected according to certain rules including fault A, fault B, and faults A and B combined in the associated body

TABLE 5. **The non-dimensional experimental results combined with the KNN algorithm.**

Testing fault type	Practicing fault group					Correct or incorrect
	The third fault experiment					
	f1	g1	h1	i1	j1	
a1	Imaginary number					IC
b1	0.2104	0.1973	0.1843	0.2105	0.1975	IC
c1	0.2394	0.1863	0.1727	0.2141	0.1875	IC
d1	0.1532	0.1532	0.2081	0.2496	0.2358	C
e1	0.1530	0.1526	0.2355	0.2226	0.2362	C
f1	Imaginary number					IC
g1	0.2303	0.1923	0.1796	0.2051	0.1927	IC
h1	0.1896	0.1892	0.1759	0.2294	0.2160	IC
i1	0.1599	0.1592	0.2168	0.2318	0.2323	IC
j1	0.1555	0.1692	0.1972	0.2390	0.2391	C
f2	Imaginary number					IC
g2	0.1856	0.1978	0.1729	0.2340	0.2097	IC
h2	0.1896	0.1624	0.1893	0.2294	0.2293	IC
i2	0.1737	0.1736	0.2000	0.2395	0.2132	C
j2	Imaginary number					IC
	The fourth experiment					
	f2	g2	h2	i2	j2	
a1	Imaginary number					IC
b1	0.1646	0.2150	0.2150	0.2149	0.1905	IC
c1	0.1844	0.2106	0.2105	0.2234	0.1711	IC
d1	0.1320	0.2105	0.2234	0.2235	0.2105	C
e1	0.1411	0.2121	0.2107	0.2113	0.2248	C
f1	0.2927	0.2252	0.1719	0.1711	0.1392	C
g1	0.1974	0.2363	0.2235	0.1844	0.1584	C
h1	0.1521	0.2151	0.2151	0.2275	0.1902	IC
i1	0.1356	0.2024	0.2308	0.2284	0.2027	IC
j1	0.1218	0.2295	0.2030	0.2428	0.2028	IC
f2	Imaginary number					IC
g2	Imaginary number					IC
h2	0.1517	0.2124	0.2120	0.2268	0.1971	IC
i2	0.1346	0.2313	0.2073	0.2194	0.2074	IC
j2	0.1321	0.2237	0.2235	0.2103	0.2103	IC

of evidence. These data points correspond to fault 1, 2, and 3 in the ER. One of the bodies of evidence is selected for the remainder of the associated body of evidence. The experimental steps described above are completed one by one. The experimental results in Table 2 use the inner ring wear, teeth missing in the big gear, and a combination of inner ring wear and teeth missing in the big gear as the three fault types. Results from an additional fault set are shown in Table 3, which uses outer ring wear, teeth missing in the big gear, and a combination of outer ring wear and teeth missing in the big gear as the three fault types.

Since the K-S test is based on the similarity probability between two bodies of evidence, the similarity probability will tend towards 1 if the two bodies of evidence are closer. In this paper, the test result is judged to be true or not by comparing the probability similarity value. For example, when fault 0 is a2 then a1, c1 or d1 can be selected as fault 1, 2, or 3. If the K-S test result shows that the fault 0 and fault 1 test results are bigger than fault 0 and fault 2 or fault 0 and fault 3, it is judged that the diagnosis result is correct, but if not, the result is incorrect. In the table, we use 'C' and 'IC' to represent whether the result is 'correct' or 'incorrect'.

Another experiment was performed to compare with the above experimental results. The dimensionless index and KNN combination were used for fault diagnosis [10] and the same experimental data from the experiment above was used. The experimental results are shown in Table 4 and Table 5. In these tables, the practicing fault group refers to the five fault types which are known (such as the first collection of 49 groups of faults A, B, C, D, and E). However, the fault type being tested is unknown. Additionally, the tested fault type is one of the five fault types, but the data has been collected by different data collectors at a different time (such as the second collection of 49 groups of fault A). A correct judgment is obtained when the type of test failure corresponds to the position of the fault type in the training results. For example, when the tested fault type is B, the second number is compared with the other four numbers. If the second number has the biggest value, the diagnosis result is judged to be correct.

1) EXPERIMENTAL DISCUSSION

Tables 2 and 3 show that there were 26 experiments with a correct diagnosis out of a total of 72 experiments using the double sample data fusion method based on combination rules. This translates to a fault diagnosis accuracy rate of 36.11%.

Tables 4 and 5 show that there were 16 experiments with a correct diagnosis out of a total of 60 experiments using the mutual dimensionless index and KNN combination. This translates to a fault diagnosis accuracy rate of 26.66%.

Analyzing the 72 experiments in Tables 2 and 3, it is found that the same number is shown twice or more for the 26 correct diagnoses in the experimental results. This shows that ER is good at fusing the five mutual dimensionless indices, but at the same time, it eliminates the characteristics of each index. This may be occurring due to a large flaw in how the reliability is calculated. In this paper, the arithmetic mean is obtained from a group of data and then the reliability is identified by calculating the distance between each data element in the group using the arithmetic mean. Although this method makes it convenient to obtain reliability, it ignores the distribution of each data element. The reliability may be incorrect when each group's distribution is not evenly distributed.

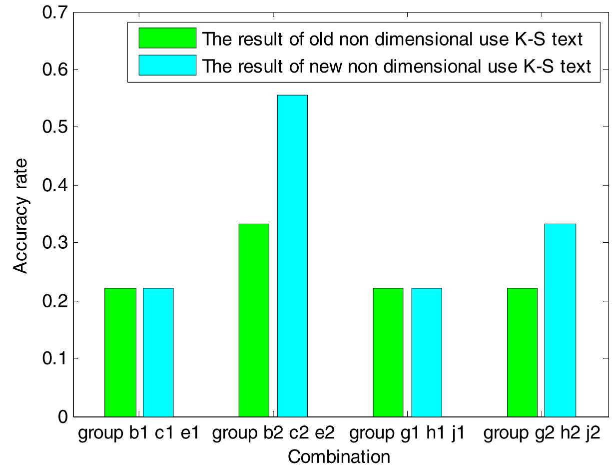


FIGURE 3. Old dimensionless and new dimensionless each data after using evidence reasoning fault diagnosis accuracy comparison chart.

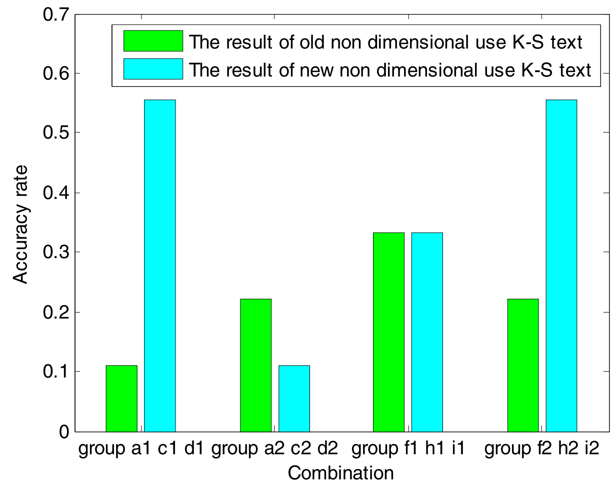


FIGURE 4. Old dimensionless and new dimensionless each data after using evidence reasoning fault diagnosis accuracy comparison chart.

Comparing the fault diagnosis accuracy rate between ER with the dimensionless index and the KNN combination method shows that ER has a greater ability to handle conflicting evidence. Currently, the accuracy of these two methods is not high enough to allow their use in practical situations. This suggests that there is a large external interference when the data is collected, which means that the original fault data does not represent the real characteristics for each fault type.

To further review the algorithm, the old dimensionless index and mutual dimensionless index are compared in this paper and the fault diagnosis accuracy rate is plotted, as shown in Figures 3 and 4. It is found that the mutual dimensionless index can correct weaknesses that are blurred or difficult to discern when using the old dimensionless index.

IV. CONCLUSION

Due to the fact that fault data is easily influenced by various external factors during data collection, some blurring or highly conflicting evidence may appear and lead to errors in the fault diagnosis results. In this paper, a double sample data fusion method is proposed for fault diagnosis, which is based on combination rules. This method uses a new proposed mutual dimensionless index as the body of evidence in the ER. The reliability is calculated and fused with the mutual dimensionless index of each fault type to reduce the effect of uncertainties. The experimental results show that this method can fuse the mutual dimensionless indices and accurately determine different fault types. However, some of the data fusion results are blurred. After analysis, it was found that the reliability calculation method is inadequate for this experiment because it does not consider the distributed nature of the body of evidence.

REFERENCES

[1] J. P. Xuan and T. L. Shi, "Classification feature extraction of multiple gear faults using genetic programming," *J. Eng. Mech.*, vol. 19, no. 1, pp. 70–74, 2006.

[2] Y. F. Xie, Q. H. Zhang, F. Duan, and G. X. Sun, "The application of a new dimensionless parameter in diagnosis of rotating machinery," *J. Guangdong Univ. Petrochem. Technol.*, vol. 22, no. 3, pp. 33–36, 2012.

[3] Y. B. Gu, *Power Generation Equipment State Maintenance Theory and Technology*. Beijing, China: China Electric Power Press, 2009.

[4] Q. H. Zhang, L. Q. Shao, H. F. Li, and Y. J. Zhu, "Research of concurrent faults diagnosis technology of rotating machinery based on the non-dimensional parameter," *J. Huazhong Univ. Sci. Tech. (Natural Sci. Ed.)*, vol. 37, pp. 156–159, 2009.

[5] L.-L. Jiang, H.-K. Yin, X.-J. Li, and S.-W. Tang, "Fault diagnosis of rotating machinery based on multisensor information fusion using SVM and time-domain features," *Shock Vibrat.*, vol. 2014, no. 1, pp. 153–154, 2014.

[6] A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping," *Ann. Math. Statist.*, vol. 38, no. 2, pp. 325–339, 1967.

[7] G. Shafer, *A Mathematical Theory of Evidence*. Princeton, NJ, USA: Princeton Univ. Press, 1976.

[8] X.-S. Si, C.-H. Hu, J.-B. Yang, and Q. Zhang, "On the dynamic evidential reasoning approach for fault prediction," *Expert Syst. Appl.*, vol. 38, no. 5, pp. 5061–5080, 2011.

[9] J. B. Xiong, Q. H. Zhang, G. X. Sun, Z. P. Peng, and Q. Liang, "Fusion of the dimensionless parameters and filtering methods in rotating machinery fault diagnosis," *J. Netw.*, vol. 9, no. 5, pp. 1201–1207, 2014.

[10] 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," *Math. Problems Eng.*, vol. 2015, no. 4, pp. 1–9, 2015.

[11] J. Xiong, Q. Zhang, G. Sun, X. Zhu, Z. Li, and M. Liu, "An information fusion fault diagnosis method based on dimensionless indicators with static discounting factor and KNN," *IEEE Sensors J.*, vol. 16, no. 7, pp. 2060–2069, Apr. 2016.

[12] J. B. Yang, "Rule and utility based evidential reasoning approach for multiattribute decision analysis under uncertainties," *Eur. J. Oper. Res.*, vol. 131, no. 1, pp. 31–61, 2001.

[13] J. B. Yang and D. L. Xu, "On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 32, no. 2, pp. 289–304, May 2002.

[14] J.-B. Yang and D.-L. Xu, "Nonlinear information aggregation via evidential reasoning in multiattribute decision analysis under uncertainty," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 32, no. 3, pp. 376–393, May 2002.

[15] X.-S. Si, C.-H. Hu, J.-B. Yang, and Z.-J. Zhou, "A new prediction model based on belief rule base for system's behavior prediction," *IEEE Trans. Fuzzy Syst.*, vol. 19, no. 4, pp. 436–451, Aug. 2011.

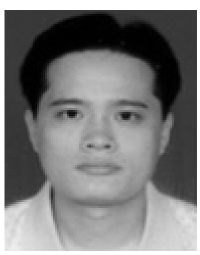
[16] C. H. Hu, X. S. Si, and J. B. Yang, "Dynamic evidential reasoning algorithm for system reliability prediction," *Int. J. Syst. Sci.*, vol. 41, no. 7, pp. 783–796, 2010.

[17] A. S. Qin, Q. H. Zhang, Q. Hu, and G. X. Sun, "Diagnosis model of rotating machinery fault based on evidence reasoning," *Mach. Tool Hydraulics*, vol. 42, no. 21, pp. 188–191, 2014.

[18] F. Y. Cong, J. Chen, and G. M. Dong, "Performance degradation assessment by Kolmogorov–Smirnov test and prognosis based on AR model," *J. Vibrat. Shock*, vol. 31, no. 10, pp. 79–82, 2012.

JIANBIN XIONG (M'12) was born in Hunan, China, in 1976. He received the B.E., M.A. and Ph.D. degrees from the Guangdong University of Technology, China. He is currently a Research Scientist with the Guangdong Provincial Key Laboratory of Petrochemical Equipment Fault Diagnosis, Guangdong University of Petrochemical Technology, China. He is currently an Associate Professor with the College of Computer and Electronics Information, Guangdong University of Petrochemical Technology, China. He has published over 40 papers in related international conferences and journal. His research interests include fault diagnosis, signal processing, image processing, information fusion, and computer applications.

QINGHUA ZHANG received the B.Sc. degree in electrical automation from Henan Polytechnic University, China, in 1985, the M.Sc. degree in industrial automation from the South China University of Technology, China, in 1995, and the Ph.D. degree in control theory and control engineering from the South China University of Technology, China, in 2004. He is currently a Professor and also the Dean of the Guangdong Provincial Key Laboratory of Petrochemical Equipment Fault Diagnosis, Guangdong University of Petrochemical Technology, China. His research interests include condition monitoring and fault diagnosis of rotating machinery, intelligence control, and applications of intelligent algorithms.



ZHIPING PENG was born in Fujian, China, in 1969. He received the B.E. and M.A. degrees from the Guangdong University of Petrochemical Technology, China. He is currently pursuing the Ph.D. degree with the South China University of Technology, China. He is currently with the School of Computer and Electronic Information, Guangdong University of Petrochemical Technology, China, as a Professor and Instructor of M.E. students. His current research interests include cloud compute, image processing, information fusion, and computer applications.

GUOXI SUN received the B.Sc. degree from the Department of Precision Instrument and Mechanology, Tianjin University, China, in 1992, and the M.Sc. and Ph.D. degrees from the South China University of Technology, in 1999 and 2006. respectively. He is currently an Associate Professor with the Guangdong University of Petrochemical Technology, China. His main research areas include algorithms, artificial intelligence, and fault diagnosis.

YONGDA CAI was born in Guangdong, China, in 1995. He is currently pursuing the bachelor's degree with the Guangdong University of Petrochemical Technology, China. He is a good Programmer in MATLAB and C language.

His current research interests include data fusion and fault diagnosis.