Intelligent Fault Diagnosis Based on Granular Computing

Xiaoxu Yan, Zhousuo Zhang, Wei Cheng

Department of Mechanical Engineering, State Key Laboratory for Manufacturing Systems Engineering
Xi’an Jiaotong University, Xi’an 710049, P.R. China
E-mail: daokym@hotmail.com  zzs@mail.xjtu.edu.cn  xjtuwei@gmail.com

Abstract
This paper presents a new approach to intelligent fault diagnosis of the machinery based on granular computing. The tolerance granularity space mode is constructed by means of the inner-class distance defined in the attributes space. Different features of the vibration signals, including time domain statistical features and frequency domain statistical features, are extracted and selected using distance evaluation technique as the attributes to construct the granular structure. Finally, the proposed approach is applied to fault diagnosis of rolling element bearings, and testing results show that the proposed approach can reliably recognize different faulty categories and severities.

Keywords: Granular computing, tolerance relations, granularity structure, fault diagnosis

1. Introduction

Nowadays, the machineries are playing a very important role in many fields. And the normal operation of the machineries is very important. However, there are kinds of mechanical faults that occur frequently and cause great casualties and economical loss. In order to keep the machines performing at its best. Different methods of fault diagnosis have been developed and used effectively to detect the machine faults at an early stage, such as the neural network, fuzzy theory. They can have an effective diagnosis to the different faults, but to the diagnosis of the different stages of one fault, the existing methods can’t get good effect. And the diagnosis of the different stages of one fault is especially important and has great influence on the monitoring and diagnosis of machineries. In this paper, Granular Computing (GrC) is used. It can not only implement the diagnosis of the different faults, but effectively implement the diagnosis of the different stages of one fault. A complicate problem can be divided into several small ones which can be easily understood and solved according to the idea of GrC. Therefore, by constructing granularity structure, the faults information may be decomposed into different granularity levels, and each level can be clearly understand and analyzed. The relations of the granules and the different levels may provide us a good method to distinguish the different faults. Generally, we construct the granularity structure using the features exacted from the vibration signals of the machineries. However, some of the features contain too much unrelated information to the faults and there is a high degree of overlap among the values of these features of different faults. These features would confuse the granular structure and therefore, cause great performance degradation. The effect of the granularity structure may be greatly different with different features. So it’s very necessary to select the most useful features. Here, distance evaluation technique [10] is used to select the most superior features from the original features set.

Naturally, study of related models of granular computing is a meaningful direction. In the past ten years, there have been some studies about models of granular computing. Zadeh[4] proposed a general framework of granular computing based on fuzzy set theory. Lin [5, 6] and Yao[7] considered a model of granular computing using neighborhood systems. Pawlak[8] built a model of granular computing based on rough set theory. Recent years, shi zhongzhi[1] proposed a new mode called tolerance granularity space which has been proved to have good performance on information classification. In this paper, we propose a new approach to intelligent fault diagnosis of machinery based on tolerance granularity space.

The organization of this paper is as follows: we introduce the basic concept about tolerance granularity
space in section 2. Faults diagnosis mode is constructed in Section 3. In Section 4, we verify the good performance of this approach by means of the data of rolling element bearings.

2. Classification based on tolerance granularity space

We assume that information about objects in a finite universe is given by a decision table DT = (U, C ∪ D, Vα, f). The details can be found in Ref. [14]. And let B ⊆ C. Then the rule set F generated from DT and B consists of all rules with the form as follows:

\[ \land \{(a, v): a \in B \text{ and } v \in (Vα \cup \{\ast\}) \} \rightarrow d = v_d(1) \]

Where v_d ∈ Vα. The symbol \("\ast\"") means that the value of the corresponding attribute is irrelevant to the rule. The length of the rule is the number of attribute whose value of corresponding eigenvalues are less than or equal to \(r\), and remaining the \("\ast\"") formerly in \(w_j\), can be got by ergodicing every "0" with "1" in \(w_{j,1}\), and remaining the "1" formerly in \(w_{j,1}\). Until all the dimension of \(w\) are "1", the last level granules are got. And each of the higher level granules is a sub-granule of a certain granule in front level. A granule may have several sub-granules. After getting the \(EG\) of each \(TG\) in per level, and if the \(TG\) is \(CTG\), the classification rule can be formed as follows:

\[ IG_{i,j} = \eta_{i,j} \ast w_{i,j} \]

That means the values in \(IG\) corresponding with the "0" in \(w\) are "\(*\)\", and the others are the values in \(\eta\) corresponding with the "1" in \(w\). The length of the rule shows the level to which the \(TG\) belongs.

So, the more the attributes contains, the finer the objects set is granulated. We can select different levels for solving our problem according to different situation.

3. Fault diagnosis based on the structure of tolerance granularity space

And \(w_\alpha = 0\), because the last dimension denotes the decision label of the objects, so it isn’t involved in the calculation.

A tolerance granule (TG) is composed by two parts [13]: the intension of TG, named \(IG\), which is the decision rule defined in formula (1); the extension of TG, named \(EG\), which contains all the objects satisfying \(IG\) in the decision table. That is, \(TG = (IG, EG)\). Therefore, if all the classification labels of the objects in \(EG\) are the same, we call the \(TG\) consistent tolerance granules (CTG), the \(IG\) of this \(TG\) can be a classification rule, or else not.

A granular structure have \(m\) levels, \(m\) is equal to the dimension of \(w\), when only one dimension is "1" in \(w\), others are "0", and let "1" ergodics every dimension in \(w\), we can get all the 1th-level granules under different forms of \(w\): \(G_1 = \{TG_{1,1}, \cdots TG_{1,s}\}\). And each \(w\) corresponds with a granule. Then we get the \(j\)-th-level granules by granulating every granule in the \((j-1)\)-th level by means of the formula as follow:

\[ EG_j = \{x | (x, \eta_j) \in cp(\alpha, \beta | DIS, D) \wedge x \in EG_{j-1}\} \]

Where \(\eta_j\) can be got by ergodicing every "0" with "1" in \(w_{j,1}\), and remaining the "1" formerly in \(w_{j,1}\). Until all the dimension of \(w\) are "1", the last level granules are got. And each of the higher level granules is a sub-granule of a certain granule in front level. A granule may have several sub-granules. After getting the \(EG\) of each \(TG\) in per level, and if the \(TG\) is \(CTG\), the classification rule can be formed as follows:

\[ IG_{i,j} = \eta_{i,j} \ast w_{i,j} \]
The main task of intelligent fault diagnosis is faults classification. We need distinguish the different faults as well as the different stages of a fault with a high accuracy. Here, the different faults and the different severities of one fault can be divided into different information granules, and each granule may contain the different fault or the information of the different stages of one fault by constructing the tolerance granularity space mode. Fig-1 shows the mode of intelligent fault diagnosis. We can know that in a decision table, DT=<U, C ∪ D, V_a, f>, all the sample signals consist of the training samples set U, C is a feature set. And D contains the different faults expressed by different number. V_a contains all the eigenvalues of the features.

The results evaluation technique [10] can be used to select the most superior features from the original features. The features with the smallest distance evaluation criterion a_j will be the most superior features. As shown in formula (5), where, d_j(w) is the average distance of the C faults. d_j(b) is the average distance between the different faults samples.

\[
\begin{align*}
|d_{c,j}| = & \frac{1}{M_c \times (M_c - 1)} \sum_{l,m=1}^{M_c} |v_{m,c,j} - v_{l,c,j}|, \\
& l,m = 1,2, \cdots M_c, l \neq m \\
\end{align*}
\]

Where, \(v_{m,c,j}\) is the jth eigenvalue of the mth sample under the cth condition, \(M_c\) is the sample number of the cth condition. \(d_{c,j}\) is the jth inner-class distance under the cth condition. \(C\) is the number of the conditions.

So, if the samples x and \(x_j\) satisfy the tolerance proposition, the tolerance granules TG are generated as:

\[
\begin{align*}
EG = (x, x_j) \iff & dis(x, x_j | w_i) \leq 0 \\
x = (x_0, x_1, \cdots, x_n) \\
x_j = (x_0, x'_1, \cdots, x'_n) \\
\end{align*}
\]

\[
\begin{align*}
x \oplus x_j = & \begin{cases} 
0 & |x-x_j| \leq DIS \ast w_i \\
1 & \text{else}
\end{cases}
\end{align*}
\]

Where x is replaced by the next samples after the \(x_j\) ergodic all the samples in the training set. Here, in order to know whether the TG is CTG or not, we define the confidence of the granule TG [13] [15]:

\[
Conf(c_j) = \text{confidence} (class=c_j | a=v) = \text{P(class=c_j | a=v)}
\]

\[
P(class = c_i | a = v) = \frac{M_{class=c_i \land a=v}}{M_{a=v}}
\]

Where \(M_{class=c_i \land a=v}\) is the number of the samples with condition \(c_i \in D\) in the TG and \(M_{a=v}\) is the number of the samples with attribute set \(a \subset C, v \in V_a\). Because some features of the different faults may
overlap with the others, this will affect the judgment to the TG. So, for every TG, we define a threshold value $T_{\text{value}}$ as follows:

$$IG = \begin{cases} (r_1, r_2, \cdots, r_{n-1}, c) & \text{max}(\text{conf}) \geq T_{\text{value}} \\ 0 & \text{else} \end{cases}$$

(8)

$0.5 \leq T_{\text{value}} \leq 1$

Where $T_{\text{value}}$ denotes the reliability of the rule and we can give it different values in different situation. $r_i$ is the average of all the eigenvalues of each feature in the TG corresponding with the positions in $w_i$, and $r_j = \omega^*$ in $w_j = 0$. $c$ is the condition label with the maximum Conf in each TG.

As shown in Fig-1, we extract the eigenvalues of the superior features of the testing samples by distance evaluation technique. And then we calculate $R$ from the first level as follows:

$$R = \{IG_{i,j}, \cdots\} \Leftrightarrow \text{dis}(IG_{i,j}, y | w_{i,j}) \leq 0$$

(9)

Where, $i$ denotes the $i$th level, $j$ denotes the $j$th granule in $i$th level.

Step 1: If all the IGs in $R$ are the same $c$, the condition of $y$ is $c$.

Step 2: If all the IGs in $R$ are not the same $c$, and any granule has sub-granules, we calculate in the next level in the same way as step 1.

Step 3: If all the IGs in $R$ are not the same $c$ and none of the granules has sub-granules, we get the condition of $y$ as follows:

$$\text{Support}_{i,j} = \frac{M_{EG_{i,j}}}{M_{TG}}$$

(10)

$M_{EG_{i,j}}$ and $M_{TG}$ are the number of samples in the $EG_{i,j}$ and in the $TG$, respectively. So the IG with the maximal of $\text{Support}$ decides the condition of $y$. But if the TG with the maximal $\text{Support}$ is not one, we calculate $P$ as follows:

$$P = P(|(y - R_i|^2) = \sum_{j=0}^{n-1} (y_j - r_j)^2, y = (y_0, \cdots, y_{n-1}),$$

(11)

$$R_i = (r_0, \cdots, r_{n-1})$$

$R_i$ is the $IG_{i,j}$ with the maximal support. Then, the IG with the minimal $P$ decides the condition of $y$.

4. Experiments

The vibration data used in this paper have been obtained from the data set of the rolling element bearings [16]. The bearings are installed in a motor driven mechanical system, as shown in Fig. 2. A 2 hp, three-phase induction motor is connected to a dynamometer and a torque sensor by a self-aligning coupling. The dynamometer is controlled so that desired torque load levels can be achieved. An accelerometer is mounted on the motor housing at the drive end of the motor to acquire the vibration signals from the bearing. The data collection system consists of a high bandwidth amplifier particularly designed for the vibration signals and a data recorder with a sampling frequency of 12,800 Hz per channel. The data recorder is equipped with low-pass filters at the input stage for anti-aliasing. On the other hand, the frequency in the vibration signals of the system under study did not exceed 5000 Hz, and therefore the sampling rate is ample. Each bearing was tested under the four different loads (0, 1, 2 and 3 hp). We apply this proposed approach to the fault diagnosis of rolling elements bearings. In present paper, the original data is divided into some samples with 4096 data points. The detailed description of the data set is shown in Table 1.

![Fig-2 the experiment equipment](image)

**Table 1. Description of bearing data**

<table>
<thead>
<tr>
<th>Defect size (inches)</th>
<th>Speed (rpm)</th>
<th>Motor load (hp)</th>
<th>Operating condition</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>40/10</td>
<td>0</td>
<td>1725/3</td>
<td>Normal</td>
<td>1</td>
</tr>
<tr>
<td>40/10</td>
<td>0.014</td>
<td>1728/3</td>
<td>Outer race</td>
<td>2</td>
</tr>
<tr>
<td>40/10</td>
<td>0.017</td>
<td>1732/3</td>
<td>Ball</td>
<td>3</td>
</tr>
<tr>
<td>40/10</td>
<td>0.014</td>
<td>1733/3</td>
<td>Inner race</td>
<td>4</td>
</tr>
<tr>
<td>40/10</td>
<td>0.007</td>
<td>1732/3</td>
<td>Ball</td>
<td>5</td>
</tr>
<tr>
<td>40/10</td>
<td>0.007</td>
<td>1733/3</td>
<td>Inner race</td>
<td>6</td>
</tr>
<tr>
<td>40/10</td>
<td>0.007</td>
<td>1733/3</td>
<td>Outer race</td>
<td>7</td>
</tr>
</tbody>
</table>

The data set is composed by 350 samples with seven different operating conditions, including normal condition, outer race fault, inner race fault and ball fault under the 3hp load. And each condition has 50 samples among which 40 samples are training samples, the rest 10 are testing samples. The faulty defect size (diameter, depth) is 0.007 and 0.014 inches. Then we train the mode using the 280 training samples to get the granular structure, and test using the 70 testing sample. Fig-3 shows the frequency-domain waveform of the different faults. We extract 6 time-domain statistic features and 13 frequency-domain statistic features respectively, select three to six most superior features according to the distance evaluation criterion from these 19 features as shown in Fig-4 and form the
attribute set using these most superior features to construct the granular structure, respectively. Table 2 shows the classification accuracy with different number of superior features. Obviously, the accuracy changes from 91.43% to 98.57% as the granularity levels changing from two to six. This means that the finer we granulate, the more information about the faults we get.

Table 2. Performance comparison for different features

<table>
<thead>
<tr>
<th></th>
<th>Different superior features</th>
<th>Four poorest features</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Two</td>
<td>Three</td>
</tr>
<tr>
<td>Training</td>
<td>98.57</td>
<td>98.93</td>
</tr>
<tr>
<td>Testing</td>
<td>91.43</td>
<td>94.29</td>
</tr>
</tbody>
</table>

Here, we set the $T_{value} = 0.9$. Obviously, only 2 samples were wrongly classified when selecting four superior features. However, when we select four poorest features among these 19 features as shown in Fig-4, the rate of wrong classification is 38.57%, compared with 2.86% (four most superior features), it is much higher.

Table 3. Description of bearing data

<table>
<thead>
<tr>
<th>The number of training/testing samples</th>
<th>Defect size (inches)</th>
<th>Motor Speed (rpm)</th>
<th>Operating/Load (HP)</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>35/15</td>
<td>0.007</td>
<td>1732</td>
<td>3 Ball</td>
<td>1</td>
</tr>
<tr>
<td>35/15</td>
<td>0.014</td>
<td>1755</td>
<td>2 Ball</td>
<td>2</td>
</tr>
<tr>
<td>35/15</td>
<td>0.021</td>
<td>1756</td>
<td>3 Ball</td>
<td>3</td>
</tr>
</tbody>
</table>

In our study, we found that the $T_{value}$ can affect the classification results. When the $T_{value}$ is too big, we may miss some rules which should be right, and when the $T_{value}$ is too small, we may get some rules which should be wrong. Table 4 shows the different features under the seven fault conditions. These features can confuse thegranular structure. Therefore, this causes great performance degradation.

From the Fig-4, it can be seen that four poorest features are all in time-domain features, while four superior features are all in frequency-domain ones. This implies that time-domain features provide too little bearing fault-related information, and therefore, are unable to distinguish the seven classes.

Taking the six-level granular structure as example, we calculated the correct classification samples in each granularity level, shown in Fig-5. X-axis shows the seven conditions, and Y-axis indicates the ten testing samples of each condition, the total is 70.

This might suggest that some of the features contain too much fault-unrelated information and there is a high degree of overlap among the values of these features under the seven fault conditions. These features can confuse the granular structure. Therefore, this causes great performance degradation.

It can be easily found that the fourth and the fifth conditions are completely distinguished in the first level. Meanwhile, 80% of the sixth and seventh conditions are distinguished in the first level. While the first and second conditions are mainly distinguished in the third level. As to the third condition, most samples are distinguished in the second level; only one sample was wrongly classified in the forth level. This denotes that this granular structure has fine and precise decomposition to the seven mechanical conditions.
classification accuracy under different $T_{value}$ with four superior features shown in Fig-4. The faulty data is shown in Table 3.

The result shows that when the $T_{value}$ is 0.7 or 0.75, the rate of wrong classification (2.22%) is the smallest. It suggests that the confidence of each granule plays an important role to the granularity structure. And we can get a good hierarchical effect of the faults if we can get the most suitable $T_{value}$. In this paper, the $T_{value}$ was artificially set. So how to find the most superior $T_{value}$ needs further study.

Table 5. Performance comparison for different $T_{value}$

<table>
<thead>
<tr>
<th></th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.75</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Testing</strong></td>
<td>82</td>
<td>93.33</td>
<td>95.56</td>
<td>97.78</td>
<td>97.78</td>
<td>95.56</td>
<td>95.56</td>
</tr>
</tbody>
</table>

5. Conclusion

In this paper, we proposed a new approach to the intelligent fault diagnosis of mechanical equipments based on tolerance granular space. We selected two to six most superior features as attributes respectively, and get the high classification accuracy 98.57% under six features. The experimental results show that the construction of the attribute space has an important influence to the performance of the granular structure. And the approach enables the diagnosis of abnormalities in bearings and at the same time identification of the categories and severities of faults with a high accuracy. In the paper, the highest classification accuracy 97.78% is got under $T_{value}=0.5$ or 0.7. And it still needs further study on the method of finding the most superior $T_{value}$ to control the confidence of each granule. This approach not only extends the application field of granular computing (GrC), but also introduces a new and effective method for fault diagnosis.

Acknowledgement

This paper is supported by Nation High-tech R&D Program of China (863 Program) (2006AA04Z430).

Reference


