Research Article

Improvement of Adaptive GAs and Back Propagation ANNs Performance in Condition Diagnosis of Multiple Bearing System Using Grey Relational Analysis

Lili A. Wulandhari,[1] Antoni Wibowo,[2] and Mohammad I. Desa[3]

[1]School of Computer Science, Bina Nusantara University, Jakarta 11480, Indonesia
[2]School of Quantitative Sciences, UUM College of Arts and Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia
[3]Advanced Informatics School (AIS), Universiti Teknologi Malaysia, 54100 Kuala Lumpur, Malaysia

Correspondence should be addressed to Lili A. Wulandhari; lili.wulandhari@gmail.com

Condition diagnosis of multiple bearings system is one of the requirements in industry field, because bearings are used in many equipment and their failure can result in total breakdown. Conditions of bearings commonly are reflected by vibration signals data. In multiple bearing condition diagnosis, it will involve many types of vibration signals data; thus, consequently, it will involve many features extraction to obtain precise condition diagnosis. However, large number of features extraction will increase the complexity of the diagnosis system. Therefore, in this paper, we presented a diagnosis method which is hybridization of adaptive genetic algorithms (AGAs), back propagation neural networks (BPNNs), and grey relational analysis (GRA) to diagnose the condition of multiple bearings system. AGAs are used in the diagnosis algorithm to determine the best initial weights of BPNNs in order to improve the diagnosis accuracy. In addition, GRA is applied to determine and select the dominant features from the vibration signal data which will provide good diagnosis of multiple bearings system in less features extraction. The experiments results show that AGAs-BPNNs with GRA approaches can increase the accuracy of diagnosis in shorter processing time, compared with the AGAs-BPNNs without the GRA.

1. Introduction

A bearing is a device widely used in industries to minimize friction on rotating part of machine by giving smooth metal balls or roller and a smooth inner and outer metal surface for the balls to roll against. Unfortunately, bearing is one of the machine parts which has high percentage of failures compared to other components [1]. Based on previous study, bearings contribute 40–50% causes of machine failure [2]. Therefore, a precise condition diagnosis of bearing system is essential to detect defects before they lead to failures.

A precise condition diagnosis can be achieved by good condition monitoring. Condition monitoring of a bearing is reflected by vibration signal data. The vibration signal data is captured by accelerometers which record the condition of the bearing continuously. Vibration signal data is commonly used for bearing condition diagnosis due to its intrinsic advantage of revealing bearing failure [3–5]. Vibration signal under different condition will show different pattern [6], as can be seen in Figure 1. Figure 1 shows that the vibration signal data of normal bearing has different pattern from faulty bearing. Vibration signal of faulty bearing has higher amplitude compared with the normal bearing. However, multiple bearing system with one faulty bearing may not be visually different on the represented vibration signal data compared to all bearings being normal. Therefore, it is important to have a technique which is able to accurately diagnose the system condition based on the continuously monitored vibration signals.

It is important to extract features from vibration signal data, since the vibration signal data captured from mechanical system such as bearing are complex in nature and some of the useful information is corrupted [5, 7]. In order to capture
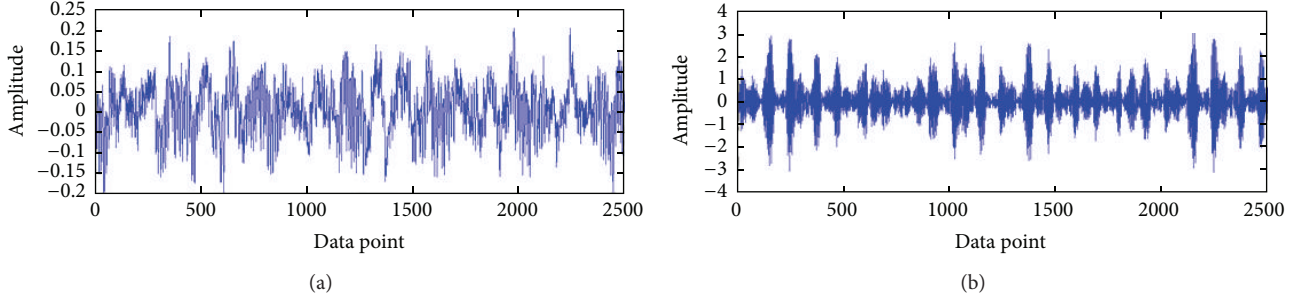
(a)

(b)

FIGURE 1: Vibration signal data of normal bearing (a) and faulty bearing (b).

TABLE 1: Multiple bearings specifications [22].

Bearing	Inside diameter (inches)	Outside diameter (inches)	Thickness (inches)	Ball diameter (inches)	Pitch diameter (inches)
DE bearing	0.9843	2.0472	0.5906	0.3126	1.537
FE bearing	0.6693	1.5748	0.4724	0.2656	1.122

the diagnostic information from the vibration signal data, it is appropriate to compute as many features as possible. In this paper, we extract ten features from vibration signal data of the bearing system. Those features are standard deviation, skewness, kurtosis, the maximum peak value, absolute mean value, root mean square value, crest factor, shape factor, impulse factor, and clearance factor [8]. These features are effective and practical for condition diagnosis due to their relative sensitivity to early faults and robustness to various loads and speeds [9]. But the choice of features is often arbitrary, which will lead the situation where some features contain the same information and the others contain no useful information at all [8]. Therefore, a technique to determine which are the dominant features for multiple bearing diagnoses is important. One of the simplest ways to find the dominant parameters is by trying and combining the features continuously up to the desired result achieved. However, this method will spend much time and memory to obtain the dominant features. Grey relational analysis (GRA) is one of the methods which are commonly used to find the dominant features. It is used as feature selection method to remove irrelevance and redundant factors that affect the results [10]. The essential thing of GRA is that it can be used to describe the relation among features [11]. GRA can be employed to explain the complicated interrelationship among the data when the trends of their development are either homogeneous or heterogeneous [12]. In this study, we used grey relational analysis (GRA) to determine the dominant features which contain useful information in multiple bearings condition diagnosis. The dominant features from GRA will be used as the input in condition diagnosis algorithm.

Diagnosis algorithm had been proposed by many researchers, some of them using individual metaheuristic techniques such as the genetic algorithms (GAs) and fuzzy and neural networks (NNs) [13–17]. However, individual metaheuristic techniques suffer from their own drawbacks, which can be overcome by forming a hybrid approach combining the advantages of each technique [18]. Hence, researchers have recently started to propose hybrid metaheuristic
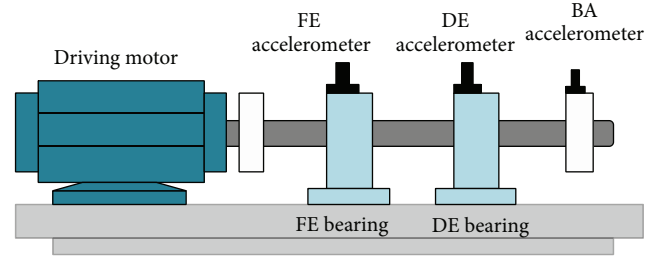

FIGURE 2: Bearing and accelerometer structure.

techniques to improve the performance of condition diagnosis. Wulandhari et al. [19–21] improved the condition diagnosis work in specific type of fault for multiple bearing using hybrid genetic algorithms and back propagation neural networks (GAs-BPNNs) approach. In this paper, we propose an improvement of GA-BPNNs by applying adaptive methods to GAs and back propagation neural networks (AGAs-BPNNs) and using GRA to identify and select the dominant features for AGAs-BPNNs algorithm in order to obtain a precise condition diagnosis for the multiple bearing systems.

2. Bearing Vibration Signal Data

In this paper, we use the vibration signals data from the Case Western Reserve University Bearing Data Center [22]. The vibration signals data were captured from a two-bearing system, which consists of Drive End bearing (DE) and Fan End bearing (FE), with various combinations of the bearing conditions. The specifications of the bearing are given in Table 1. For the purpose of capturing the vibration signals data, three accelerometers were attached on the bearings and the baseline (BA) as shown in Figure 2. Bearing vibration data were collected under seven different conditions: (1) FE and DE Normal, (2) FE Normal and DE Inner Race Fault (DE-IRF), (3) FE Normal and DE Ball Fault (DE-BF), (4) FE

Normal and DE Outer Race Fault (DE-ORF), (5) FE Inner Race Fault (FE-IRF) and DE Normal, (6) FE Ball Fault (FE-BF) and DE Normal, and (7) FE Outer Race Fault (FE-ORF) and DE Normal. The example of the data captured is presented in Table 2.

From Table 2, we can see that each condition has three streams of data as captured by the three accelerometers; thus, each feature will be extracted from three accelerometers. Based on the available data, generally only seven condition classes of bearing can be specified as the output of the diagnosis. In this paper, we expanded the condition classes from seven to sixteen classes by combining and mixing the available data. For the FE-IRF and DE-IRF class, for instance, its BA data was set or obtained from the average of BA accelerometer in FE-IRF and DE-IRF condition, the FE data was obtained from FE accelerometer in FE-IRF condition, and the DE data was obtained from DE accelerometer in DE-IRF condition. The expansion of condition classes was done to obtain more specific condition diagnosis for each bearing so that any action to each bearing can be specifically carried out. The advantage of this expansion is that, here, we can identify the condition of DE and FE bearing simultaneously. In the seven classes case, we can only identify the condition of either one. The sixteen classes of the bearing conditions are presented in Table 3.

The classes of multiple bearing conditions are influenced by ten features extracted from the data. The values of the features lie within the interval which is the lower and upper bound of the data extraction. The interval of the features values is presented in Table 4.

3. The Proposed Algorithm

This paper proposed a hybrid method of GRA, AGA, and BPNN to diagnose the condition of multiple bearing systems. This hybridization applies GRA to determine the dominant features by analyzing the relation between each feature and its ideal values. The algorithm is started by initializing the features which are extracted from the vibration signal data and then calculate the grey relational coefficient (GRC) followed by calculating the grey relational grade (GRG) which both of them including to the GRA process. The results from the GRA are the dominant features which will be used as the input of the AGAs-BPNNs. The framework of the GRA-AGA-BPNNs is shown in Figure 3.

3.1. Grey Relational Analysis (GRA). A system for which the relevant information is completely known is called a white system, while a system for which the relevant information is completely unknown is a black system. Any system between these limits is a grey system which has poor and limited information [23]. In multiple bearing systems, any information about the condition of the bearing is not completely revealed by the features extracted from the vibration signal data. This unclear condition of data can be overcome using GRA which was proposed by Deng [24] in 1982. GRA utilizes the mathematical method to analyzing correlation between the references series which is the ideal value of features

and the alternatives series [25]. It firstly normalizes the features extracted and then translates the performance of all alternatives into a comparability sequence with the ideal value called grey relational generating [26], followed by the calculation of grey relational coefficient between all comparability sequences and the references sequences. Finally, the grey relational grade between the reference sequence and every comparability sequence is calculated based on the grey relational coefficient. The highest grey relational grade of the alternatives features indicates that the features have dominant influence to the condition diagnosis. The procedures of GRA are shown in Figure 4.

The GRA starts by normalizing the features extracted, for the simplification; we omit the units of all parameters. This can be done easily by multiplying the certain parameter, say p_i, with 1/the unit of parameter p_i. Let x_{ij} be the ith sample of the jth feature extracted where $i = 1, 2, \ldots, 240$ and $j = 1, 2, \ldots, 10$:

$$x_{\max} = \max\left(x_{ij} \mid i = 1, 2, \ldots, 240; j = 1, 2, \ldots, 10\right),$$
$$x_{\min} = \min\left(x_{ij} \mid i = 1, 2, \ldots, 240; j = 1, 2, \ldots, 10\right). \tag{1}$$

Then, the normalized features x_{ij}^* can be obtained by the following equation [27]:

$$x_{ij}^* = \frac{x_{ij} - x_{\min}}{x_{\max} - x_{\min}}. \tag{2}$$

Next, the grey relational generating is conducted by determining the reference or ideal values of the features extracted. Let P_j be the features extracted, with $j = 1, 2, \ldots, 10$. C_s is the condition of the bearing systems with $s = 1, 2, \ldots, 16$. It is noticed that we used 240 pieces of sample data and 16 classes for bearing conditions where each class C_s consists of 15-item data which fall in the condition class s. The ideal values of the features extracted, say z_{sj}, are the average of jth feature extracted value in sth condition which can be written as follows:

$$z_{sj} = \frac{1}{15} \sum_{i=15s-14}^{15s} x_{ij}^*. \tag{3}$$

For example, z_{11} is the ideal value for the first condition class (class of FE and DE Normal) and the parameter p_1. Regarding obtaining the next step, namely, grey relational coefficient, we need to determine the comparability sequences of the alternatives values and the ideal values. It is noticed that we used 240 pieces of sample data and 16 classes for bearing conditions where each class C_s consists of 15-item data which fall in the class s. Then, we define

$$w_{ij} = z_{1j} \quad \text{for } j = 1, \ldots, 10; \ i = 1, \ldots, 15,$$

$$w_{ij} = z_{2j} \quad \text{for } j = 1, \ldots, 10; \ i = 16, \ldots, 30,$$

$$w_{ij} = z_{3j} \quad \text{for } j = 1, \ldots, 10; \ i = 31, \ldots, 45,$$

$$\vdots$$

$$w_{ij} = z_{nj} \quad \text{for } j = 1, \ldots, 10; \ i = 15(n-1)+1, \ldots, 15n, \tag{4}$$

TABLE 2: Example of bearing vibration data.

| Number	Normal bearing		Drive end bearing fault data									Fan end bearing fault data								
			Inner race fault			Ball fault			Outer race fault			Inner race fault			Ball fault			Outer race fault		
	DE	FE	BA	DE	FE	BA	DE	FE	BA	DE	FE	BA	DE	FE	BA	DE	FE	BA	DE	FE
1	0.053	0.146	0.065	−0.083	−0.402	0.016	−0.003	−0.247	0.000	0.009	−0.407	0.098	−0.025	−0.051	0.017	−0.168	0.320	−0.031	−0.134	0.127
2	0.089	0.098	−0.023	−0.196	−0.005	0.017	−0.096	0.143	0.069	0.424	0.263	0.042	−0.029	−0.192	−0.004	0.181	0.326	−0.120	0.003	−0.259
3	0.100	0.055	−0.089	0.233	−0.107	−0.036	0.114	0.003	0.031	0.013	0.495	−0.042	−0.046	0.051	−0.169	0.044	−0.260	−0.006	−0.027	−0.060
4	0.059	0.037	−0.094	0.104	−0.074	−0.045	0.257	−0.107	−0.037	−0.265	−0.423	0.081	0.001	0.151	−0.069	−0.270	0.031	0.060	−0.184	0.454
5	−0.005	0.054	−0.076	−0.181	0.209	0.008	−0.058	0.136	−0.116	0.237	−0.307	0.059	−0.037	−0.095	0.090	−0.138	0.447	−0.131	−0.203	0.075
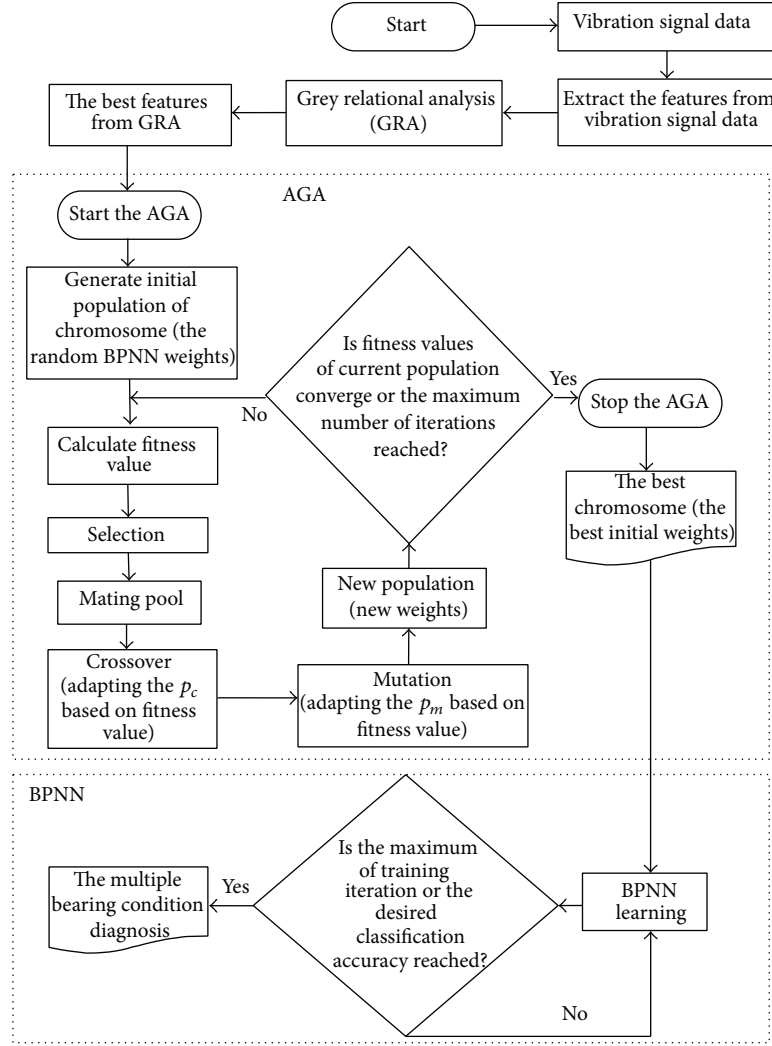
FIGURE 3: The framework of GRA-AGAs-BPNNs.

where $n = 1, \ldots, 16$; then, the comparability of the alternatives and ideal values can be calculated as

$$\Delta w_{ij} = \left| x_{ij}^* - w_{ij} \right|, \tag{5}$$

where Δw_{ij} is the comparability of alternatives and ideal values, x_{ij}^* is the alternatives value which is normalized features extracted from the vibration signal data, and w_{ij} is the ideal values which is defined based on the condition classes. Based on (4) and (5), the next step of GRA procedures, namely, grey relational coefficient calculation between comparability sequences and ideal sequences, is written as follows:

$$\mathrm{GRC}_{ij} = \frac{\min_k \left(\min_l \left(\Delta w_{kl} \right) \right) + \xi \max_k \left(\max_l \left(\Delta w_{kl} \right) \right)}{\left(\Delta w_{ij} \right) + \xi \max_k \left(\max_l \left(\Delta w_{ij} \right) \right)} \tag{6}$$

$$\text{for } k = 1, 2, \ldots, 240; \; l = 1, 2, \ldots, 10,$$

where GRC_{ij} is the grey relational coefficient value of ith samples and jth feature and ξ is the distinguishing coefficient which is defined in the range $0 \leq \xi \leq 1$. Then, the grey

relational grade (GRG) is determined by averaging the GRC to each feature and is represented as

$$\mathrm{GRG}_j = \frac{1}{m} \sum_{i=1}^{m} \mathrm{GRC}_{ij}, \tag{7}$$

where m is the number of samples. Equation (7) is used to find the GRG of accelerometers BA, DE, and FE. The final grey relational grade on the feature j, say GRG_j, is the average of GRG_j^*s from BA, DE, and FE, respectively. The results of the experiment using GRA are presented in Section 4.

3.2. The Proposed GRA-AGAs-BPNNs Algorithm. The proposed GRA-AGAs-BPNNs algorithm is the hybrid algorithm which combines the advantages of GRA, GAs, and BPNNs to one algorithm for the condition diagnosis of the multiple bearing systems. The dominant features from GRA are used in AGAs-BPNNs algorithm to classify the condition effectively. Adaptive operator probabilities techniques in GAs are proposed to obtain better initial weights for BPNNs training.

TABLE 3: Sixteen classes of bearing conditions.

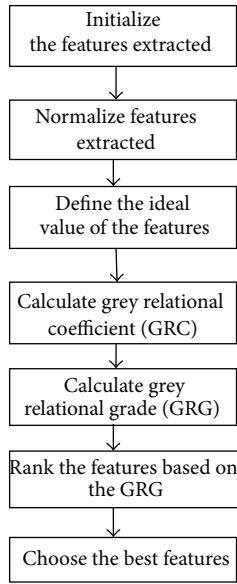| Number	Condition
C1	FE and DE Normal
C2	FE Normal and DE-IRF
C3	FE Normal and DE-ORF
C4	FE Normal and DE-BF
C5	FE-IRF and DE Normal
C6	FE-ORF and DE Normal
C7	FE-BF and DE Normal
C8	FE-IRF and DE-IRF
C9	FE-IRF and DE-ORF
C10	FE-IRF and DE-BF
C11	FE-ORF and DE-IRF
C12	FE-ORF and DE-ORF
C13	FE-ORF and DE-BF
C14	FE-BF and DE-IRF
C15	FE-BF and DE-ORF
C16	FE-BE and DE-BF

FIGURE 4: Grey relational analysis procedures.

The adaptive technique is applied to maintain the diversity of the population by varying the probabilities of crossover (p_c) and mutation (p_m) as, for example, [28–31]. The algorithm for the proposed AGAs-BPNNs is as follows.

(1) Let (I_k, T_k) be the kth input and target pair of the problem to be solved by BPNN, with $k = 1, 2, \ldots, N_{in}$, and N_{in} is the number of paired data.

(2) Let N_{pop}, N_{chro}, p_{c0}, p_{m0}, and N_{iter} be the number of populations, number of chromosomes, initial crossover probability, initial mutation probability, and maximum number of iterations, respectively. Initialize p_{c0}, p_{m0}, R_{pc}, and R_{pm} where R_{pc} are random vectors of numbers which generated in range [0, 1]

with size $1 \times N_{chro}/2$ and R_{pm} are random vectors of numbers which generated in range [0, 1] with size $1 \times N_{chro}$. Set $i = 0$.

(3) Determine the number of dominant features chosen.

(4) Determine the BPNN architecture in terms of the number of input neuron, hidden layer, hidden neuron and output neuron, and the activation functions based on the number of dominant features chosen.

(5) Generate an initial population of chromosomes Q_0. Each chromosome contains genes which correspond to BPNN random weights, and the number of genes in a chromosome is equal to the number of BPNN weights.

(6) Calculate the fitness value $F(i, j)$ of the jth chromosome in population Q_i using

$$F(i, j) = \frac{1}{E(i, j)}, \quad j = 1, 2, \ldots, N_{chro}, \tag{8}$$

where $E(i, j)$ is the mean square error (MSE) of the jth chromosome in the population Q_i. It is calculated based on the selected BPPN architecture as follows:

$$E(i, j) = \frac{1}{2} \sum_{k=1}^{N_{in}} \left(T_{kj} - O_{kj}^i \right)^2, \tag{9}$$

where T_{kj} is the target of the kth input in the jth chromosome and O_{kj}^i is the output of the kth input in the jth chromosome of the population Q_i based on the selected BPNN architecture.

(7) Generate the mating pool by selecting the best chromosomes using roulette selection methods.

(8) Select parent pairs of population Q_i, say $(\phi_{1s}^i, \phi_{2s}^i)$, from the mating pool for crossover mechanism where $s = 1, 2, \ldots, S$ and $S = \lfloor N_{chro}/2 \rfloor$.

(9) Calculate the crossover probability of the sth parents pairs in the population Q_i [32]:

$$p_c\left(i, \phi_{1s}^i, \phi_{2s}^i\right)$$
$$= \begin{cases} p_{c0} \dfrac{\left(F_{max}(i) - F'(i, s)\right)}{\left(F_{max}(i) - \overline{F}(i)\right)} & \text{if } F'(i, s) > \overline{F}(i) \\ p_{c0} & \text{otherwise,} \end{cases} \tag{10}$$

where

$$F'(i, s) = \begin{cases} F\left(\phi_{1s}^i\right) & \text{if } F\left(\phi_{1s}^i\right) > F\left(\phi_{2s}^i\right) \\ F\left(\phi_{2s}^i\right) & \text{otherwise.} \end{cases}$$

$F(\phi_{1s}^i)$ and $F(\phi_{2s}^i)$ are the fitness values of parent 1 and parent 2, respectively; $F_{max}(i)$ is the maximum fitness value of the population Q_i; $\overline{F}(i)$ is the average fitness value of the population Q_i.

TABLE 4: The proposed feature interval of the sixteen condition classes.

Number	Condition classes	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}
C1	DE and FE Normal	[0.064, 0.088]	[−0.166, 0.336]	[2.361, 3.421]	[0.185, 0.291]	[0.054, 0.073]	[0.005, 0.008]	[29.942, 47.827]	[0.083, 0.114]	[2.966, 4.524]	[3.471, 5.279]
C2	FE Normal and DE IRF	[0.087, 0.094]	[−0.135, 0.247]	[2.831, 3.497]	[0.236, 0.330]	[0.068, 0.075]	[0.008, 0.009]	[30.717, 38.750]	[0.110, 0.119]	[3.465, 4.586]	[4.182, 5.487]
C3	FE Normal and DE ORF	[0.078, 0.093]	[−0.256, 0.009]	[2.636, 3.390]	[0.223, 0.368]	[0.062, 0.074]	[0.006, 0.009]	[29.835, 45.121]	[0.097, 0.117]	[3.482, 5.063]	[4.123, 6.018]
C4	FE Normal and DE BF	[0.033, 0.039]	[−0.131, 0.171]	[2.688, 3.380]	[0.099, 0.139]	[0.026, 0.031]	[0.001, 0.002]	[76.682, 104.733]	[0.192, 0.230]	[2.609, 5.978]	[2.975, 13.998]
C5	FE IRF and DE Normal	[0.063, 0.092]	[−0.191, 0.153]	[2.202, 5.258]	[0.198, 0.276]	[0.036, 0.077]	[0.004, 0.009]	[24.479, 55.206]	[0.103, 0.113]	[2.609, 5.978]	[2.975, 13.998]
C6	FE ORF and DE Normal	[0.083, 0.085]	[−0.114, 0.035]	[2.383, 2.672]	[0.199, 0.238]	[0.068, 0.070]	[0.006, 0.007]	[27.955, 33.393]	[0.101, 0.104]	[2.878, 3.458]	[3.358, 4.032]
C7	FE BF and DE Normal	[0.117, 0.121]	[−0.139, 0.038]	[1.978, 2.342]	[0.239, 0.319]	[0.099, 0.104]	[0.014, 0.015]	[17.123, 22.767]	[0.137, 0.144]	[2.343, 3.213]	[2.613, 3.671]
C8	FE IRF and DE IRF	[0.077, 0.092]	[−0.149, −0.014]	[2.579, 4.377]	[0.228, 0.286]	[0.054, 0.075]	[0.006, 0.009]	[28.882, 46.978]	[0.107, 0.114]	[3.173, 5.282]	[3.726, 9.742]
C9	FE IRF and DE ORF	[0.078, 0.093]	[−0.175, −0.014]	[2.441, 4.231]	[0.232, 0.309]	[0.051, 0.075]	[0.005, 0.009]	[29.705, 47.740]	[0.101, 0.114]	[3.280, 5.038]	[3.833, 9.406]
C10	FE IRF and DE BF	[0.046, 0.064]	[−0.151, 0.089]	[2.460, 4.083]	[0.157, 0.195]	[0.032, 0.053]	[0.003, 0.005]	[53.629, 70.318]	[0.073, 0.077]	[3.244, 4.872]	[3.747, 9.239]
C11	FE ORF and DE IRF	[0.085, 0.089]	[−0.081, 0.099]	[2.672, 3.084]	[0.227, 0.284]	[0.068, 0.072]	[0.007, 0.008]	[30.872, 36.071]	[0.106, 0.111]	[3.323, 4.022]	[3.895, 4.759]
C12	FE ORF and DE ORF	[0.081, 0.089]	[−0.150, 0.009]	[2.510, 2.961]	[0.221, 0.296]	[0.066, 0.072]	[0.006, 0.008]	[36.222, 37.993]	[0.100, 0.110]	[3.301, 4.154]	[3.855, 4.860]
C13	FE ORF and DE BF	[0.059, 0.061]	[−0.090, 0.074]	[2.637, 2.986]	[0.159, 0.175]	[0.047, 0.050]	[0.004, 0.0043]	[54.090, 67.421]	[0.073, 0.077]	[3.331, 3.920]	[3.874, 4.588]
C14	FE BF and DE IRF	[0.103, 0.107]	[−0.099, 0.079]	[2.474, 2.803]	[0.256, 0.321]	[0.085, 0.089]	[0.011, 0.012]	[24.912, 29.619]	[0.125, 0.131]	[3.062, 3.770]	[3.574, 4.393]
C15	FE BF and DE ORF	[0.098, 0.106]	[−0.155, −0.012]	[2.342, 2.731]	[0.255, 0.231]	[0.082, 0.088]	[0.010, 0.011]	[24.737, 32.152]	[0.118, 0.128]	[3.092, 3.805]	[3.576, 4.517]
C16	FE BF and DE BF	[0.076, 0.080]	[−0.111, 0.054]	[2.408, 2.767]	[0.175, 0.220]	[0.063, 0.067]	[0.007, 0.008]	[47.319, 61.536]	[0.091, 0.096]	[3.083, 3.805]	[3.553, 4.468]

[a, b] is the close interval between a and b.

TABLE 5: The GRG and sequence of features extracted.

Number	GRG ξ = 0.3	GRG ξ = 0.5	GRG ξ = 0.6	GRG ξ = 1	Features
1	0.912	0.945	0.954	0.972	Root mean sq. value
2	0.912	0.945	0.953	0.971	Standard deviation
3	0.909	0.943	0.952	0.970	Abs. mean value
4	0.868	0.916	0.929	0.956	Skewness
5	0.864	0.910	0.923	0.951	Max peak value
6	0.761	0.833	0.855	0.907	Shape factor
7	0.746	0.830	0.854	0.904	Kurtosis
8	0.480	0.604	0.646	0.752	Crest factor
9	0.412	0.536	0.580	0.695	Impulse factor
10	0.367	0.488	0.532	0.652	Clearance factor

(10) Calculate mutation probability of the jth chromosome in the population Q_i [32]:

$$p_m(i, j) = \begin{cases} p_{m0} \dfrac{(F_{\max}(i) - F(i, j))}{(F_{\max}(i) - \overline{F}(i))} & \text{if } F(i, j) > \overline{F}(i) \\ p_{m0} & \text{otherwise,} \end{cases}$$

(11)

where $F(i, j)$ is the fitness value of the jth chromosome in the population Q_i.

(11) Set $i = i + 1$.

Generate Q_i by applying crossover and mutation mechanism based on the following rules.

   (a) For $s = 1 : S$,
   if $p_c(i, \phi_{1s}^i, \phi_{2s}^i) > R_{pc}(s)$, do a crossover between ϕ_{1s}^i and ϕ_{2s}^i. Otherwise, copy ϕ_{1s}^i and ϕ_{2s}^i as offspring.
   (b) For $j = 1 : N_{\text{chro}}$,
   if $p_m(i, j) > R_{pm}(j)$, do a mutation of the jth chromosome. Otherwise, the jth chromosome is kept unchanged.

(12) If Q_i converge or i is equal to N_{iter}, then the best chromosome is obtained and used as the initial weights for BPNN learning. Otherwise, go to step (6).

4. Experimental Evaluation and Discussion

In this section, we will describe and discuss the result of the experiment from GRA-AGAs-BPNNs in classifying the condition of the multiple systems. For the GRA process, we tried several values of ξ, namely, 0.3, 0.5, 0.6, and 1. Based on (7), GRG value for each was calculated. The sequence of the features extracted based on GRG value is given in Table 5.

Table 5 shows that if the distinguishing coefficient ξ is closer to 0, then the GRG of the feature will have a range wider than if the distinguishing coefficient ξ is closer to 1. For $\xi = 0.3$, the GRG range is around 0.545 while the $\xi = 1$ GRG is around 0.320 as shown in Figure 5. From Table 5, we can see, however, that the ξ values are varied and the sequence of
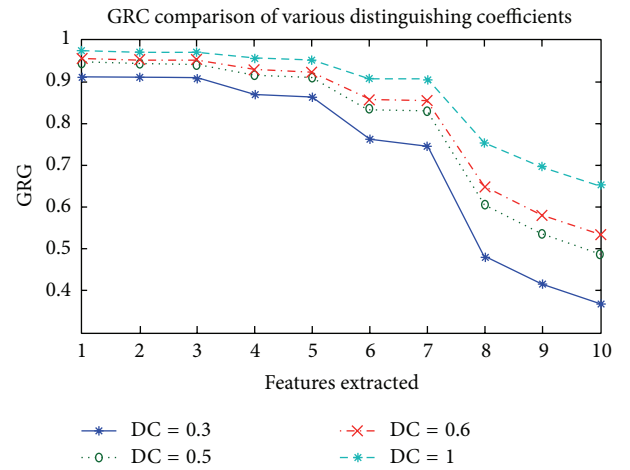

FIGURE 5: GRG comparison of various distinguishing coefficients.

features extracted is the same. Root mean square value is the features with the highest GRG value.

In this paper, we conducted experiments of AGAs-BPNNs using one, three, five, and seven of the best dominant features based on the GRG values in Table 5. In AGAs-BPNNs techniques, we set AGAs features as follows: 100 chromosomes of each population, initial crossover probability of 0.6, and initial mutation probability of 0.01. For BPNNs, we use three hidden layers and refer to m-l_1-l_2-l_3-n as m neurons input, l_1 neurons in the first hidden layers, l_2 neurons in the second hidden layers, l_3 neurons in the third hidden layers, and n neurons output. As stated in Section 2, vibration signals data is recorded from three accelerometers; thus, features are extracted from three accelerometers which cause the number of neuron inputs in BPNNs to be equal to $3 \times$ number of features. For example, if we choose five dominant features, then the neuron input contains 15 neuron inputs, where each neuron contains 240 samples of data, while each hidden layer contains 15 neurons as well and 16 neurons in output layer. The 240 samples of data are split randomly into three sets: 80% for training, 10% for validation, and 10% for testing. BPNN uses *logsig* and *purelin* as the activation function, where learning rate and momentum are 0.05 and 0.25, respectively.

TABLE 6: Comparison of performance between the AGAs-BPNNs and GRA-AGAs-BPNNs.

Number of dominant features selected	BPNNs topology	Training	Validation	Testing	CPU time
10 features	**30-30-30-30-16**	**99.3%**	**91.7%**	**92.4%**	**2905.0**
1 feature	3-3-3-3-16	41.5%	38.8%	36.2%	992.2
3 features	9-9-9-9-16	83.6%	80.4%	80.4%	1116.8
5 features	**15-15-15-15-16**	**99.4%**	**97.1%**	**96.7%**	**1279.3**
7 features	**21-21-21-21-16**	**100.0%**	**94.6%**	**92.5%**	**1468.2**
1 feature (the 10th feature)	3-3-3-3-16	28.7%	26.7%	26.3%	982.6
2 features (the 1st and the 10th features)	6-6-6-6-16	59.9%	52.9%	55.4%	1074.8

TABLE 7: Percentage of increased accuracy and reducing time compared to AGAs-BPNNs without GA.

Number of dominant features	Training (%)	Validation (%)	Testing (%)	Time reduced (%)
5 features	0.1%	5.4%	4.3%	56%
7 features	0.7%	2.9%	0.1%	49.5%

We conducted ten times the experiments of the GRA-AGAs-BPNNs for each one, three, five, and seven dominant features combination. We also conducted experiments using the lowest GRG features and combination of the highest and the lowest GRG features in AGA-BPNN to obtain condition diagnosis of multiple bearing systems. These experiments were carried out to see the influence of the selection of dominant features combination on condition diagnosis algorithm performance for multiple bearing systems. The performance of the algorithm is characterized based on the accuracy in classifying the condition of the bearings. The classification accuracy is computed by

$$\text{classification accuracy} = \frac{\text{total true predicted class}}{\text{total output}} \times 100\%. \tag{12}$$

The experiments were executed using a computer with Intel Core 2 Quad processor Q8200, 2.33 GHz and 1.96 GHz, and RAM 3.46 GB. The result of the GRA-AGAs-BPNNs is given in Table 6.

Table 6 shows the comparison of the classification accuracy between GRA-AGAs-BPNNs and AGA-BPNN algorithm without GRA (which involve ten features). We can see that the classification accuracy of GRA-AGA-BPNN using five and seven dominant features with topologies 15-15-15-15-16 and 21-21-21-21-16, respectively, has higher accuracy than AGA-BPNN with topology 30-30-30-30-16 without GRA and using ten features. From the experiment results, we obtain that root mean square value, standard deviation, absolute mean value, skewness, and maximum peak value can give good diagnosis of multiple bearing conditions. We notice that, by applying GRA in AGA-BPNN algorithm, we can achieve higher classification accuracy with shorter time. GRA is capable of determining which features can give dominant contribution in condition diagnosis of the multiple bearing system.

5. Conclusion

In this paper, we introduced a new hybrid technique of grey relational analysis, adaptive operator probabilities in genetic algorithms (AGAs), and back propagation neural networks (BPNNs), called GRA-AGAs-BPNNs for condition diagnosis of multiple bearing systems. We used grey relational analysis (GRA) to determine the dominant features which contain useful information of multiple bearings condition. GRA determines that mean square value, standard deviation, absolute mean value, skewness, and maximum peak value can give good diagnosis of the multiple bearings condition. The features from GRA are used in AGA-BPNNs to diagnose the condition effectively. We exploited the strong capability in optimization of genetic algorithms, which here have been further improved by varying the mutation and crossover operators probabilities, for searching the best initial weights for BPNNs, and the strong capability in classification of BPNNs to classify or diagnose the condition of a multiple bearing system. The AGAs strengthen the BPNNs to achieve the higher classification accuracy in shorter CPU time compared to the standard BPNN or the hybrid GAs-BPNNs.

Experimental results showed that the GRA is capable of improving the classification accuracy of AGAs-BPNNs in shorter time. The accuracy achieves 100%, 94.6%, and 92.5% for the training, validation, and testing, respectively, for 7 dominant features and 99.4%, 97.1%, and 96.7% for the training, validation, and testing, respectively, for 5 dominant features. The accuracy is increased and the processing time is reduced compared to the AGAs-BPNNs without GRA as shown in Table 7. This achievement provides the benefits for condition diagnosis in the real case, since we require a precise and quick process to diagnose the condition of multiple bearing in order to avoid total breakdown.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] P. V. J. Rodríguez and A. Arkkio, "Detection of stator winding fault in induction motor using fuzzy logic," *Applied Soft Computing*, vol. 8, no. 2, pp. 1112–1120, 2008.

[2] S. Nandi and H. A. Toliyat, "Fault diagnosis of electrical machines-a review," in *Proceedings of the IEEE International Electric Machines and Drives Conference (IEMDC '99)*, pp. 219–221, Seattle, Wash, USA, May 1999.

[3] G. Wang, Z. Luo, X. Qin, Y. Leng, and T. Wang, "Fault identification and classification of rolling element bearing based on time-varying autoregressive spectrum," *Mechanical Systems and Signal Processing*, vol. 22, no. 4, pp. 934–947, 2008.

[4] M. Xia, F. Kong, and F. Hu, "An approach for bearing fault diagnosis based on PCA and multiple classifier fusion," in *Proceedings of the 6th IEEE Joint International Information Technology and Artificial Intelligence Conference (ITAIC '11)*, pp. 321–325, Chongqing, China, August 2011.

[5] W. Y. Liu, J. G. Han, and J. L. Jiang, "A novel ball bearing fault diagnosis approach based on auto term window method," *Measurement*, vol. 46, no. 10, pp. 4032–4037, 2013.

[6] W. Li, F. Jiang, Z. Zhu, G. Zhou, and G. Chen, "Fault diagnosis of bearings based on a sensitive feature decoupling technique," *Measurement Science and Technology*, vol. 24, no. 3, Article ID 035602, 2013.

[7] R. Tiwari, V. K. Gupta, and P. K. Kankar, "Bearing fault diagnosis based on multi-scale permutation entropy and adaptive neuro fuzzy classifier," *Journal of Vibration and Control*, 2013.

[8] W. Li, T. Shi, G. Liao, and S. Yang, "Feature extraction and classification of gear faults using principal component analysis," *Journal of Quality in Maintenance Engineering*, vol. 9, no. 2, pp. 132–143, 2003.

[9] Y. Lei, Z. He, and Y. Zi, "Application of an intelligent classification method to mechanical fault diagnosis," *Expert Systems with Applications*, vol. 36, no. 6, pp. 9941–9948, 2009.

[10] A. M. Deris, A. M. Zain, and R. Sallehuddin, "Hybrid GR-SVM for prediction of surface roughness in abrasive water jet machining," *Meccanica*, vol. 48, no. 8, pp. 1937–1945, 2013.

[11] T. Xuerui and L. Yuguang, "Using grey relational analysis to analyze the medical data," *Kybernetes*, vol. 33, no. 2, pp. 355–362, 2004.

[12] T. V. Sibalija and V. D. Majstorovic, "Novel approach to multi-response optimisation for correlated responses," *FME Transactions*, vol. 38, pp. 39–48, 2010.

[13] B. Li, M.-Y. Chow, Y. Tipsuwan, and J. C. Hung, "Neural-network-based motor rolling bearing fault diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 47, no. 5, pp. 1060–1069, 2000.

[14] B. Sreejith, A. K. Verma, and A. Srividya, "Fault diagnosis of rolling element bearing using time-domain features and neural networks," in *Proceedings of the IEEE Region 10 Colloquium and 3rd International Conference on Industrial and Information Systems (ICIIS '08)*, pp. 1–6, Kharagpur, India, December 2008.

[15] H. Wang and P. Chen, "Fault diagnosis for a rolling bearing used in a reciprocating machine by adaptive filtering technique and fuzzy neural network," *WSEAS Transactions on Systems*, vol. 7, no. 1, pp. 1–6, 2008.

[16] L. Zhang, G. Xiong, H. Liu, H. Zou, and W. Guo, "Bearing fault diagnosis using multi-scale entropy and adaptive neuro-fuzzy inference," *Expert Systems with Applications*, vol. 37, no. 8, pp. 6077–6085, 2010.

[17] Y. Zhang and R. B. Randall, "Rolling element bearing fault diagnosis based on the combination of genetic algorithms and fast kurtogram," *Mechanical Systems and Signal Processing*, vol. 23, no. 5, pp. 1509–1517, 2009.

[18] P. Jayaswal, S. N. Verma, and A. K. Wadhwani, "Methodology and theory application of ANN, fuzzy logic and wavelet transform in machine fault diagnosis using vibration signal analysis," *Journal of Quality in Maintenance Engineering*, vol. 16, pp. 190–213, 2010.

[19] L. A. Wulandhari, A. Wibowo, and M. I. Desa, "Hybrid neural network-genetic algorithms approach for fault diagnosis of bearing system," in *Proceedings of the 1st International Conference on Industrial Engineering and Service Science (IESS '11)*, Solo, Indonesia, 2011.

[20] L. Wulandhari, A. Wibowo, and M. Desa, "Condition diagnosis of multiple bearings using adaptive operator probabilities in genetic algorithms and back propagation neural networks," *Neural Computing and Applications*, 2014.

[21] L. A. Wulandhari, A. Wibowo, and M. I. Desa, "Condition diagnosis of bearing system using multiple classifier of artificial neural networks and adaptive probabilities of crossover/mutation in genetic algorithms," *WSEAS Transactions on Systems and Control*, 2014.

[22] K. A. Loparo, http://csegroups.case.edu/bearingdatacenter/pages/welcome-case-western-reserve-university-bearing-data-center-website.

[23] S. Balasubramanian and D. S. Ganapathy, "Grey relational analysis to determine optimum process parameters for wire electro discharge machining (WEDM)," *International Journal of Engineering Science and Technology*, vol. 3, no. 1, pp. 95–101, 2011.

[24] J. L. Deng, "Control problems of grey systems," *Systems & Control Letters*, vol. 1, no. 5, pp. 288–294, 1982.

[25] C.-H. Li and M.-J. Tsai, "Multi-objective optimization of laser cutting for flash memory modules with special shapes using grey relational analysis," *Optics and Laser Technology*, vol. 41, no. 5, pp. 634–642, 2009.

[26] Y. Kuo, T. Yang, and G.-W. Huang, "The use of grey relational analysis in solving multiple attribute decision-making problems," *Computers and Industrial Engineering*, vol. 55, no. 1, pp. 80–93, 2008.

[27] J. L. Lin and C. L. Lin, "The use of the orthogonal array with grey relational analysis to optimize the electrical discharge machining process with multiple performance characteristics," *International Journal of Machine Tools and Manufacture*, vol. 42, no. 2, pp. 237–244, 2002.

[28] K. L. Mak, Y. S. Wong, and X. X. Wang, "Adaptive genetic algorithm for manufacturing cell formation," *International Journal of Advanced Manufacturing Technology*, vol. 16, no. 7, pp. 491–497, 2000.

[29] J. Aihong and Y. Lizhe, "Fault diagnosis based on adaptive genetic algorithm and BP neural network," in *Proceedings of the 2nd International Conference on Computer Engineering and Technology*, vol. 6, pp. 427–430, 2010.

[30] S. Blum, R. Puisa, J. Riedel, and M. Wintermantel, "Adaptive mutation strategies for evolutionary algorithms," in *The Annual Conference: EVEN at Weimarer Optimierungsund Stochastiktage*, 2001.

[31] V. Snášel, P. Kromer, and J. Platos, "Benchmarking hybrid selection and adaptive genetic operators," in *Znalosti 2008*, V. Snášel, Ed., pp. 224–233, FIIT STU Bratislava, Ústav Informatiky a Softvérového Inzinierstva, 2008.

[32] M. Srinivas and L. M. Patnaik, "Adaptive probabilities of crossover and mutation in genetic algorithms," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 24, no. 4, pp. 656–667, 1994.