

## Research Article

# Combining the Multi-Genetic Algorithm and Support Vector Machine for Fault Diagnosis of Bearings

Jianbin Xiong <sup>1</sup>, Qinghua Zhang <sup>2,3</sup>, Qiong Liang,<sup>4</sup> Hongbin Zhu,<sup>2,3</sup> and Haiying Li<sup>5</sup>

<sup>1</sup>School of Automation, Guangdong Polytechnic Normal University, Guangzhou 510665, China

<sup>2</sup>Guangdong Provincial Key Laboratory of Petrochemical Equipment Fault Diagnosis, Maoming 525000, China

<sup>3</sup>School of Computer and Electronic Information, Guangdong University of Petrochemical Technology, Maoming 525000, China

<sup>4</sup>School of Computer, Guangdong Polytechnic Normal University, Guangzhou 510665, China

<sup>5</sup>School of Computer, Jiaying University, Meizhou 514015, China

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

Received 7 January 2018; Revised 29 June 2018; Accepted 5 August 2018; Published 23 October 2018

Academic Editor: Paola Forte

Copyright © 2018 Jianbin Xiong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Overstudy or understudy phenomena can sometimes occur due to the strong dependence of support vector machine (SVM) algorithms on particular parameters and the lack of systems theory relating to parameter selection. In this paper, a parameter optimization algorithm for the SVM is proposed based on multi-genetic algorithm. The algorithm optimizes the correlation kernel parameters of the SVM using evolutionary search principles of multiple swarm genetic algorithms to obtain a superior SVM prediction model. The experimental results demonstrate that by combining the genetic algorithm and SVM algorithm, fault diagnosis can be effectively realized for bearings of rotating machinery.

## 1. Introduction

Rotating machinery has a wide range of applications in the modern industry, within the petrochemical, motor vehicle, power, metallurgy, manufacturing, and other important engineering fields [1–5]. Bearings are one of the most important components of rotating machinery and are also relatively easy to damage, which makes the technology for diagnosing bearing faults an important scientific research. According to statistics, approximately 30% of rotating machinery faults are caused by bearing faults since the running state of the bearing can directly affect the performance of the machine, and furthermore, the faults can result in violent vibrations of the mechanical equipment, leading to damage [5]. Moreover, the working status of machinery affects not only its operation but can also adversely affect follow-up production [6], causing significant losses to the national economy as well as posing a tremendous threat to the lives and safety of personnel working with the equipment. Therefore, the discussion and investigation into fault diagnosis technologies for rotating machinery is of critical

importance [1, 2, 7, 8]. To enable large rotating machinery to work safely and reliably, timely and accurate diagnosis of faults is required. Traditional diagnostic methods include touching, hearing, and seeing; however, these sensory methods overly rely on personal experience and lack any scientific basis. In addition, for more complex conditions, traditional diagnostic methods are unable to meet the needs of equipment fault diagnosis and maintenance. At present, research in the field of fault diagnosis for rolling bearings is booming, and the theory and technology are developing rapidly [9]. Important research is being carried out in the field of rotating machinery fault diagnosis using algorithms. In the field of mechanical fault diagnosis, the neural network has some value in certain applications and has been widely used since it offers major advantages in solving complex nonlinear problems. Chen et al. previously applied neural networks in fault diagnosis and achieved positive results [10, 11]. However, slow convergence, local minima, overlearning, and underlearning still exist, and the neural network algorithm requires a large amount of fault data, thus restricting further application and development of neural

networks for intelligent fault diagnosis. Accordingly, Xie et al. instead proposed a support vector machine (SVM) for fault diagnosis, making use of advantages such as strong generalization ability and ease of adapting to nonlinear problems. While some promising results were achieved, no specific method was given for parameter selection [12, 13]. Jiang et al. used support vectors to distinguish different faults by analyzing twelve time-domain features and found the method to have positive effects on a number of classifications, including chipped teeth, missing teeth, inner-race defects, and outer-race defects. However, the classifications of the twelve time-domain features of different fault types are different, leading to only approximately 50% fault classification and reaching 90% accuracy. Moreover, it is not possible to determine which of the twelve time-domain features to use as a basis for trouble shooting, in actual fault diagnosis [3]. Xiong et al. proposed a fault diagnosis method based on dimensionless data fusion. The method first deals with fault data using dimensionless data to obtain five dimensionless indexes, then uses the SVM to classify indicators and obtain the fault type. The method was able to improve classification. However, further improvements are still needed, as well as additional studies on the selection of nuclear parameters. In the early days, experience was relied upon to repeatedly debug methods to obtain parameter values when using SVM to solve problems. This approach is tedious, and moreover, it is difficult to find the global optimal value of the function, which thus fails to achieve the expected accuracy and satisfactory result. In practice, cross testing and grid search methods have been used to determine the important parameters of the SVM, although results have been unsatisfactory [14]. An SVM method based on the genetic algorithm has also been proposed in the reference [15], and the genetic algorithm was used to select the kernel function parameters and soft-edge values as well as the penalty parameters, needed to construct the SVM model. To a certain extent, the algorithm achieves better classification results; however, the genetic algorithm is still in early development and often falls into the pitfall of local convergence, thus cannot be used to optimize parameters [16]. To solve the problems outlined above, this paper proposes a fault diagnosis method based on the multipopulation genetic algorithm and SVM. In contrast to the neural network, the SVM can be used for small sample sizes, avoids the pitfall of overfitting, and is more generalized. Furthermore, the SVM can maximize extractions using sample data classification characteristics with limited information [17]. Therefore, the SVM is widely used in various fields, and the main factor determining performance is the selection of the radial basis function (RBF) kernel parameters, including the penalty factor  $C$  and nuclear parameter  $\gamma$ . Here, we used the global optimization ability of the multipopulation genetic algorithm to optimize the kernel parameters of the SVM to obtain an SVM model with high classification accuracy. The experimental results show that the proposed method can quickly obtain a superior SVM model, thereby effectively improving fault diagnosis accuracy.

The rest of this paper is arranged as follows. An introduction of the theory is presented in Section 2, followed

by the experimental setup and procedure in Section 3, and discussion of the experimental results in Section 4. Finally, a summary of the paper is given in Section 5.

## 2. Correlation Theory

*2.1. Genetic Algorithm and Multipopulation Genetic Algorithm.* The genetic algorithm (GA) is a type of intelligent algorithm which simulates natural genetic mechanisms and biological evolution as first described by Professor Holland of the University of Michigan [18]. The algorithm draws upon natural selection and genetic mechanisms of the biological world, using rules of probability transition to guide the search direction, and is well suited to parameter selection and optimization [19].

The genetic algorithm, based on Darwin's biological evolution theory and Mendel's genetics [18], consists of six main elements including chromosome coding, population initialization, individual fitness function, genetic operators, setting the operating parameters, and setting the operational termination loop conditions. For the fitness function, individuals are screened by repeated iterative evolution through three operations of gene duplication, selection, crossover, and mutation, and individuals with high fitness are retained while those with poor fitness are eliminated. Therefore, the new group inherits the information from this generation and is better than the previous generation. The operations above are repeated until a particular set of conditions are met. Finally, the optimal individuals in the population are decoded as the optimal solutions or the approximate optimal solutions, for which the iteration met the termination condition. Due to its advantages of global search capability, strong robustness, and parallel processing of data, genetic algorithms have been widely used in various fields, for applications such as optimization control, pattern recognition, and machine learning [20].

*2.1.1. Population Size.* Population size directly affects the performance of genetic algorithms. In general, when the population size is between 20 and 200, an adequate trade-off between population diversity and algorithm complexity can be achieved.

*2.1.2. Fitness Function.* The fitness function is the criterion used to evaluate the degree of adaptation of each code string to the problem and is also known as the optimal judge function. The fitness function is the only information guiding the search type in genetic algorithms, and its quality will affect the attributes of the algorithm. The definition of fitness in this article is as follows:

$$\text{fitness} = \frac{\text{cross accuracy}}{\text{number of crosses}} \quad (1)$$

*2.1.3. Genetic Operators*

*(1) Selection Operator.* It copies parent chromosomes (individuals) of high fitness (good performance) to the next

generation. Supposing a population size of  $M$ , the probability defined as  $P$ , that the individual defined as  $w$ , will be selected to inherit the next generation is

$$P = \frac{\text{fit}(\omega)}{M}. \quad (2)$$

As seen from the above equation, the higher the fitness of the individual, the greater the probability the individual will be copied to the next generation.

(2) *Cross Operator*. A group of individuals are randomly selected from the group, and the cross operation with probability  $P_c$  is implemented to attain a pair of new individuals. From reference [21], a value of crossover probability from 0.5 to 1.0 can be obtained.

(3) *Artificial Selection Operator*. It randomly changes the value of a genetic gene with a small probability. In the binary-coded chromosome algorithm, the mutation operator randomly changes one of the chromosomal genes from 1 to 0 or 0 to 1 which ensures the diversity of genotypes in the population, as well as preventing search stagnation. Usually, the mutation probability from 0.01 to 0.05 can be obtained [21].

The flow diagram of the genetic algorithm is shown in Figure 1.

**2.2. Multipopulation Genetic Algorithm.** In theory, the genetic algorithm forms a set of relatively complete algorithm systems [22]. However, with more extensive research, many shortcomings have come to light, such as the premature problems [23] related to the following aspects:

- (1) Extraordinary individuals influence each generation by controlling them leading to a stagnant population.
- (2) Evolutionary search results are sensitive to the values of  $P_c$  and  $P_M$ , and different values lead to different calculation results.
- (3) For the standard genetic algorithm (SGA), evolution is terminated when the number of iterations reaches the maximum genetic algebra set by humans. However, if the number of iterations is too small, the evolution will be insufficient, which will lead to immature convergence.

Accordingly, this paper proposes a multipopulation genetic algorithm using the idea of a parallel genetic algorithm to divide the groups into subgroups. Each subgroup evolves independently according to a certain pattern, and the best individual of each evolutionary generation is kept, to maintain the diversity of the population and inhibit premature convergence.

Compared to the SGA, MPGA has made the following improvements.

- (1) Multipopulation genetic algorithm (MPGA) introduces multiple populations to optimize the search at the same time. Different populations are assigned with different crossover and mutation probabilities to achieve different search objectives.

- (2) Each population is contacted by migration operators to achieve the coevolutionary of multiple groups. The optimal solution is the result of multiple population coevolutionary.
- (3) The optimal individual in each evolutionary generation of each group is preserved by the artificial selection operator, which is used as the basis for judging the convergence of the algorithm.

Related algorithms used in multipopulation genetic algorithms are the following:

- (1) *Elite population*: in each generation of evolution, the optimal population of other species is selected by an artificial selection operator to form the elite population. At the same time, the essence population is also the basis for judging the termination of the algorithm
- (2) *Immigration operator*: the migration operator periodically introduces the optimal individuals of various groups into other populations to exchange information among the populations. The specific operation rule is that the worst individuals in the target population are replaced by the optimal individuals of the source population
- (3) *Artificial selection operator*: the function of the artificial selection operator is to select the best individual in all kinds of groups and put it into the essence one to preserve the best individual so as to ensure that the best individual produced by various groups is not destroyed and lost [23]

The principles of the multipopulation genetic algorithm process are shown in Figure 2.

**2.3. Support Vector Machines.** The support vector machine (SVM) is a new machine-learning method with foundations in Vapnik statistical theory (statistical learning, STL) [24, 25]. The principle of minimizing structural risk is adopted to minimize the sampling error, and at the same time, it improves the ability to generalize the model as well as allowing unlimited data dimensions. In the case of linear classification, the classification surface is taken as the largest distance between two samples. When the classification is nonlinear, a high-dimensional space transformation is adopted to transform nonlinear classification into a linear classification problem in high-dimensional space [26, 27]. In addition, the method has the advantages of strong adaptability and ability to forecast, global optimization, high training efficiency, and the ability to generalize models, which solves the problems of non-linearity, small sample size, and high-dimensional pattern recognition.

As this paper only touches on SVM classification, herein, we introduce the basic ideas and principles behind linear SVM.

Assuming the linear separable sample sets as  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n), x \in R^u, y \in \{-1, +1\}$ , where  $\mu$  present the number of dimensions; when  $x$  is represented by

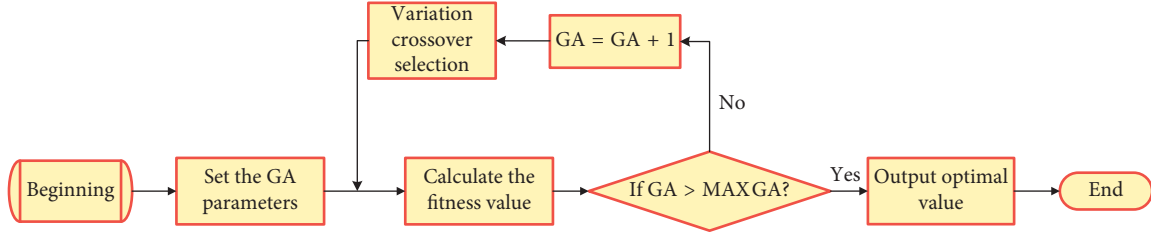


FIGURE 1: Flow chart of genetic algorithm.

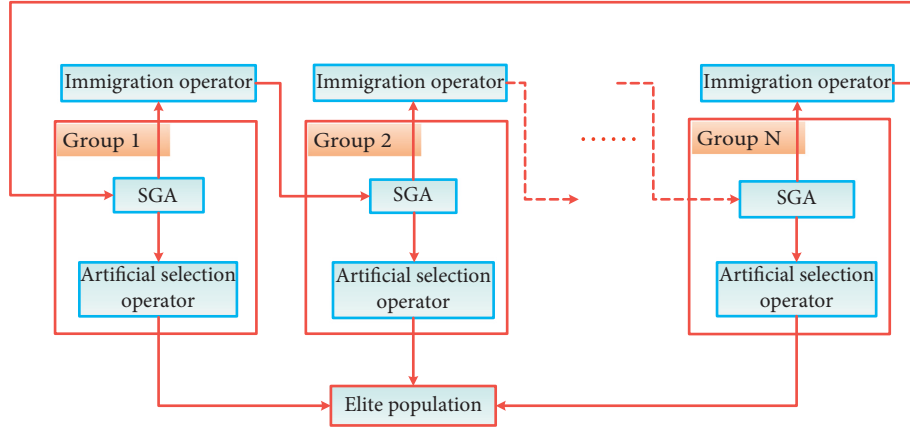


FIGURE 2: Flow chart of multipopulation genetic algorithm.

$y = +1$ , it belongs to the first category, and when it is represented by  $y = -1$ , it belongs to the second category.

*Hypothesis 1.* There exists a classification hyperplane:

$$w^T x + b = 0. \quad (3)$$

The above sample points can be divided into two categories using Equation (3), and the same category can be clustered on the same side of the hyperplane. Then,

$$\begin{cases} w^T x_i + b \geq 1, & y_i = 1, \\ w^T x_i + b \leq -1, & y_i = -1, \end{cases} \quad i = 1, 2, \dots, n. \quad (4)$$

Equation (4) can be normalized as  $y_i (w^T x + b) - 1 \geq 0$ .

Considering Figure 3, the distance between two categories of a plane is

$$\delta = \frac{|b + 1 - (b - 1)|}{\|w\|} = \frac{2}{\|w\|}. \quad (5)$$

To eliminate the problem of extracting of the norm, thereby making the problem more convenient to solve, Equation (5) can be converted to

$$\max \frac{2}{\|w\|} \Rightarrow \min \|w\| \Rightarrow \min \frac{\|w\|^2}{2}. \quad (6)$$

To obtain the maximum distance on the classification plane, this can be transformed into the following optimization problem:

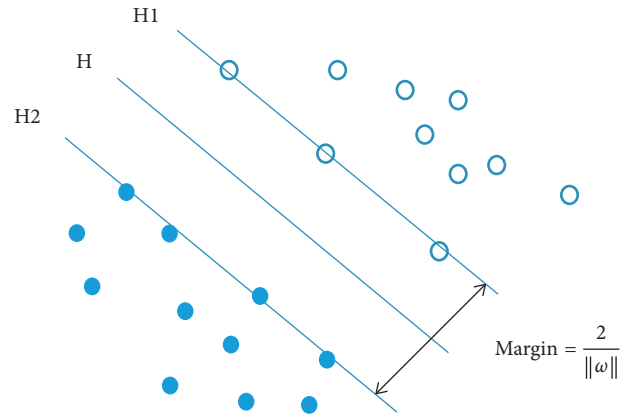


FIGURE 3: Diagram of optimal hyperplane.

$$\begin{cases} \min & \frac{\|w\|^2}{2}, \\ \text{s.t.} & y_i (w^T x + b) - 1 \geq 0. \end{cases} \quad (7)$$

Equation (7) is a convex optimization problem with a special substructure and can therefore be solved using the Lagrangian multiplier:

$$L(w, b, \lambda_i) = \frac{\|w\|^2}{2} - \sum_{i=1}^n \lambda_i [y_i (w^T x_i + b) - 1]. \quad (8)$$

Here,  $\lambda_i > 0, i = 1, 2, 3, \dots, n$ , is the Lagrange coefficient.

Due to the complexity of calculations, this cannot be solved directly, but on the basis of the Lagrange duality theory, Equation (8) can be converted to the dual problem:

$$\left\{ \begin{array}{l} \max \quad \sum_{i=1}^n \lambda_i - \frac{1}{2} \left( \sum_{i=1, j=1}^n \lambda_i \lambda_j x_i x_j \right), \\ \text{s.t.} \quad \left\{ \begin{array}{l} \sum_{i=1}^n \lambda_i y_i = 0, \\ \lambda_i \geq 0. \end{array} \right. \end{array} \right. \quad (9)$$

There is only one variable in Equation (9), which can be solved using the quadratic programming method.

If the data set is mostly separable, only a few sample points fail to construct the classification hyperplane. In such cases, we need to add a slack variable to Equation (7).

$$\left\{ \begin{array}{l} \min \quad \frac{\|\omega\|^2}{2} + C \sum_{i=1}^n \xi_i, \\ \text{s.t.} \quad \left\{ \begin{array}{l} y_i (\omega^T x + b) \geq 1 - \xi_i, \\ \xi_i > 0, \end{array} \right. \quad i = 1, 2, \dots, n. \end{array} \right. \quad (10)$$

Here, a penalty factor  $C$  is added at the same time to control sample parameters of the wrong degree.

In the vast majority of practical problems, data are nonlinear. To classify data, the usual method is to build a classification hyperplane in high-dimensional space via nonlinear mapping,  $\Psi: R^d \rightarrow H$ .

In general, nonlinear separable samples are sufficiently nonlinearly mapped to be able to find a hyperplane that separates the different types of data sets; however, the complexity of the calculation increases as the number of dimensions increases. Therefore, if there is only a singular or linear inseparable point in a certain dimensional space, we also need to add the relaxation variable  $\xi$  in high-dimensional space, as described above.

Considering the increase in the amount of computation when calculating the dot product by mapping it in the high-dimensional feature space, Vapnik et al. proposed a kernel function  $K(x_i, x_j)$  satisfying the Mercer condition, instead of the dot product operation, which greatly reduces the computational load and complexity [28]. That is,

$$K(x_i, x_j) = \Phi(x_i) \Phi(x_j). \quad (11)$$

Therefore, after mapping the sample to the high-dimensional feature space, the corresponding dual problem is

$$\left\{ \begin{array}{l} \max \quad W(\lambda) = \sum_{i=1}^n \lambda_i - \frac{1}{2} \left( \sum_{i=1, j=1}^n \lambda_i \lambda_j K(x_i, x_j) \right), \\ \text{s.t.} \quad \left\{ \begin{array}{l} \sum_{i=1}^n \lambda_i y_i = 0, \\ 0 \leq \lambda_i \leq C, \end{array} \right. \quad i = 1, 2, \dots, n. \end{array} \right. \quad (12)$$

If  $\lambda^* = [\lambda_1^*, \lambda_2^*, \dots, \lambda_n^*]^T$  is the solution of Equation (4), then

$$\omega^* = \sum_{i=1}^n \lambda_i^* y_i \Phi(x_i). \quad (13)$$

The final optimization classification function is:

$$\begin{aligned} f(x) &= \text{sgn}(\omega^* \Phi(x) + b^*) \\ &= \text{sgn} \sum_{i=1}^n \lambda_i^* y_i \Phi(x_i) \Phi(x) + b^* \\ &= \text{sgn} \sum_{i=1}^n \lambda_i^* y_i K(x_i, x_j) + b^*. \end{aligned} \quad (14)$$

If the SVM takes the RBF kernel as the kernel function, the optimization problem of Equation (12) can be transformed as follows:

$$\left\{ \begin{array}{l} \max \quad \sum_{i=1}^n \lambda_i - \frac{1}{2} \left( \sum_{i=1, j=1}^n \lambda_i \lambda_j y_i y_j \exp(-\gamma \|x_i - x_j\|^2) \right), \\ \text{s.t.} \quad \left\{ \begin{array}{l} \sum_{i=1}^n \lambda_i y_i = 0, \\ 0 \leq \lambda_i \leq C, \end{array} \right. \end{array} \right. \quad (15)$$

where  $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2 / 2\sigma^2)$ ,  $\sigma > 0$ , is the expression of the RBF core.

According to Equation (15), the performance of the SVM based on the RBF kernel function depends mainly on the selection of the penalty parameters  $C$  and  $\gamma$ . If the value of  $C$  is small, the training error will increase or otherwise lead to the problems of overlearning and weak generalization ability. Moreover, the value of  $\gamma$  also influences the results. If the value of  $\gamma$  is too small, overlearning will occur, otherwise, it will lead to other learning problems. Therefore, selecting appropriate values for  $C$  and  $\gamma$  is necessary [29].



### 3. Experiment

**3.1. Experimental Setup.** Experimental data were collected from a large rotating machine, belonging to Guangdong Provincial Petrochemical Equipment Fault Diagnosis Key Laboratory, using a multistage centrifugal fan fault diagnosis set. The set consisted of 11 kW five-stage centrifugal fans with gearboxes, frequency conversion motors, and a variety of fault gears and bearing components, and it was able to simulate common faults of a multistage centrifugal fan set. We collected five types of data sets including normal bearing, short of ball bearing, outer-ring wear, inner-ring wear, and large and small gears teeth deficiency. Pictures of part of the normal and faulty components are shown in Figure 4.

During the data collection process, various types of data sets for the machine operation status were captured using the ETM390 data collector. Firstly, we found a position from chassis and marked it. Secondly, we placed the data acquisition probe of the sensor EMT390 at the marked position, denoted with label "5" as shown in Figure 5 (right). Finally, by using EMT390, we could obtain the instantaneous values of acceleration. In our experiment, the sampling frequency was 1024 Hz, and the total sampling time in each data set was 410 seconds. The real time data were stored using their data management system and processed with the algorithm.

To establish the SVM optimization model, we used the average accuracy rate, obtained by cross validation, as the fitness value of the genetic algorithm. Then, the global search ability of the GA was used to search for the ideal SVM parameters  $C$  and  $\gamma$ , to obtain the best SVM model. The specific process is shown in Figure 6 and Figure 7.

To implement the algorithm proposed in this paper, MATLAB Software with the LIBSVM toolbox, written by Mr. Lin Zhiren, was used.

#### 3.2. Experimental Procedure

##### 3.2.1. Data Collection and Processing

- (a) The various types of acceleration data sets for the machine operation status were collected at the same position on the chassis via the EMT390 data collector. Data were acquired twice for each type with 49 sets collected in total.
- (b) The collected data were exported using the EMT390 data management system and saved as a mat file.
- (c) An individual mutual dimensionless process was performed on the various machine operation status data sets in MATLAB to obtain five mutual dimensionless indicators, stored in a matrix and numbered according to the type of the machine operation status [3].
- (d) Fifty data sets from a total of 98 sets of different machine operation status data were extracted to form the training set with 250 data sets, and 50% of the data were randomly selected as the model training set via the random function in MATLAB,

while the remaining data were used to form the test set.

##### 3.2.2. Steps to Implement the Algorithm

- (a) The relevant parameters of the multipopulation genetic algorithm were set, including the maximum genetic conservation algebra  $M$ , crossover probability  $P_c$ , and mutation probability  $P_M$ , as  $M = 10$ ,  $P_c = [0.7, 0.9]$ , and  $P_M = [0.001, 0.05]$ , respectively.
- (b) The SVM parameters  $C$  and  $\gamma$  were binary encoded to generate the genetic algorithm, which is the initial population where  $C, \gamma \in [0, 10]$ .
- (c) The randomly generated data were imported into the SVM and the average accuracy, obtained by cross-validating the SVM, was set as the objective function of the genetic algorithm.
- (d) The populations were processed using genetic manipulations including crossover, mutation, selection, and the artificial selection operator.
- (e) The fitness value was tested to determine whether it satisfied the optimization condition that is, keeping generation  $M$  unchanged, and the optimal result was output, otherwise, step 4 was repeated.
- (f) The optimized values of  $C$  and  $\gamma$  were obtained and imported into the SVM to create the optimal SVM model for bearing fault diagnosis.

**3.3. Fault Vibration Signal Analysis.** At the same driver motor speed, the vibrational signal spectrum of five states including normal, short of ball outer-ring wear, inner-ring wear, and big gear tooth deficiency was obtained and is shown in Figures 8–12.

It can be seen from the time-domain waveforms that the time-domain signals of the various machine operation status are clearly different. However, we cannot identify the specific type of the machine operation status [30]. Therefore, only by analyzing and processing the signal of machine operation status and extracting the feature vectors of various machine operation status can we accurately determine the specific type of bearing failure. Accordingly, in this paper, we present a dimensionless algorithm to handle the collected machine operation status data. First, the method obtains five dimensionless indexes such as the waveform index, peak index, pulse index, margin-degree indicator, and kurtosis indicator by processing the vibration acceleration values using the dimensionless algorithm. Then, the reciprocal dimension is obtained from the dimensionless index, which can reduce the distance between the internal structures of each dimensionless index, thereby reducing coincidence of the same dimensionless indexes. Evidence bodies formed by this heterogeneous dimension data set can improve the accuracy of fault diagnosis [31].

In the experiment, 5 different types of machine operation status data sets were selected. The types of machine operation status and number of each corresponding data sets are shown in Table 1.

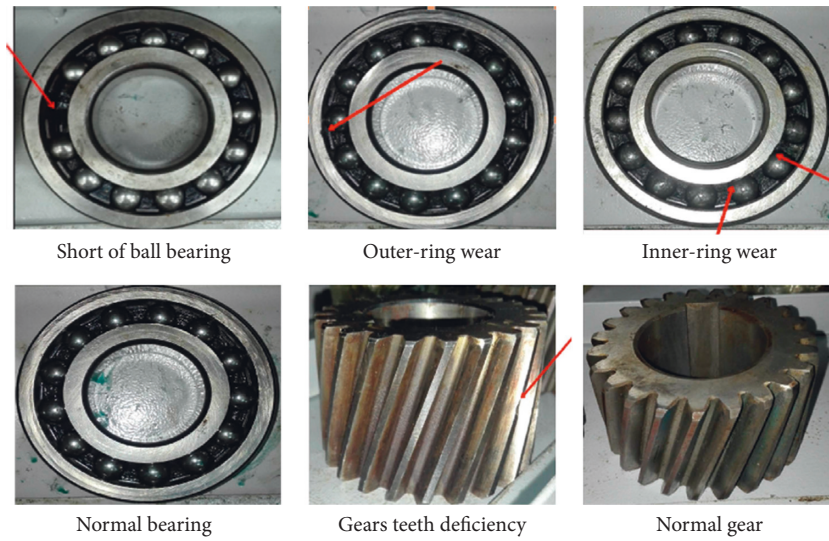


FIGURE 4: Photographs of various fault components and normal components.

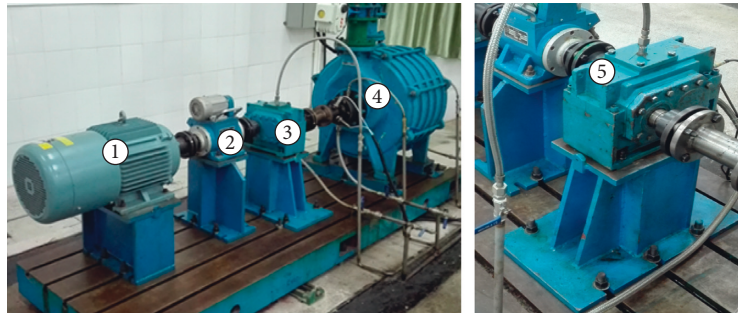


FIGURE 5: Test platform used for fault diagnosis of large-scale rotating equipment used in petrochemical industry (left). The machine mainly consists of (1) the electric motor, (2) gearbox, (3) transmission, (4) fan, and (5) data acquisition probe placement.

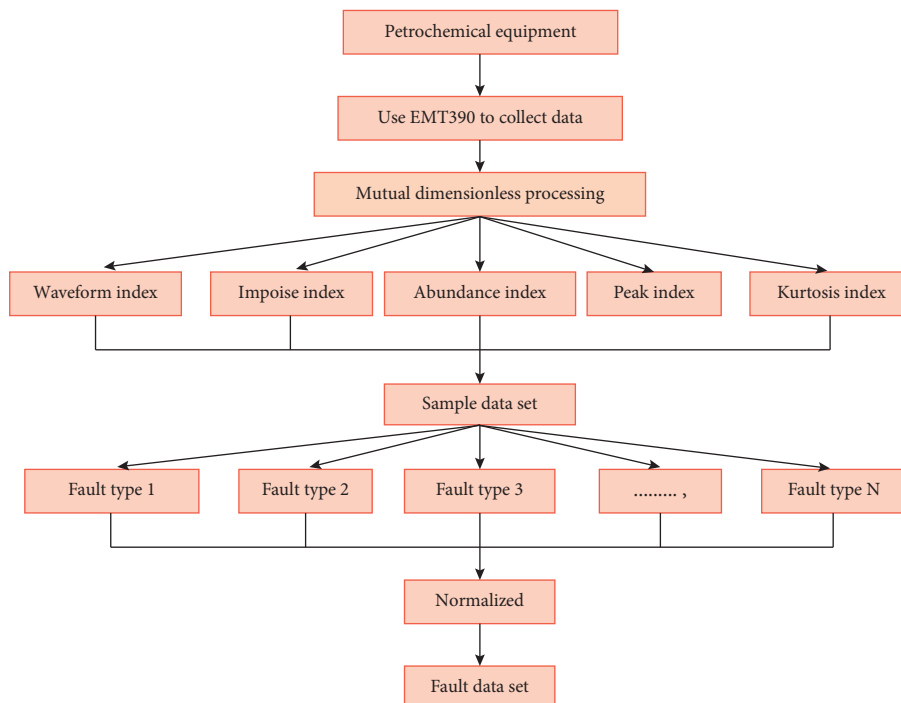


FIGURE 6: Flow chart of data processing steps.

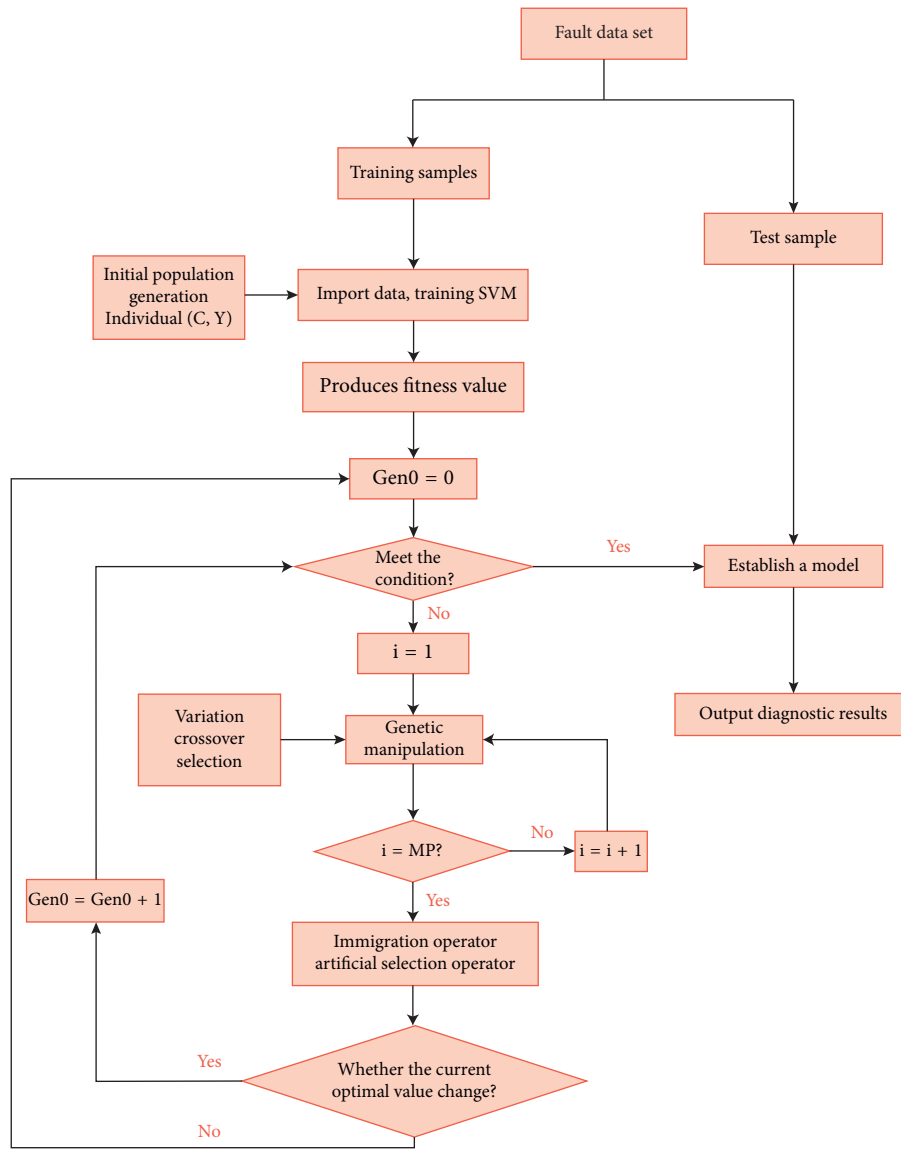


FIGURE 7: Flow chart of combined algorithm of multigroup genetic algorithm and SVM.

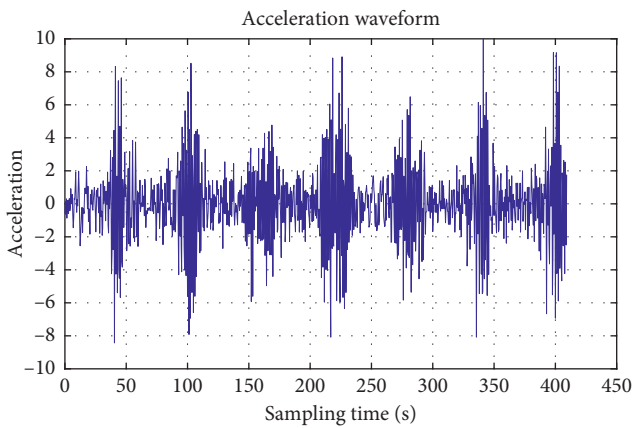


FIGURE 8: Time-domain waveform of normal bearing.

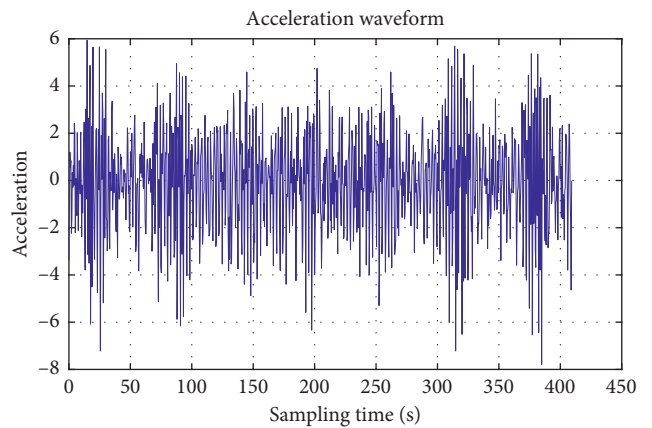


FIGURE 9: Time-domain waveform of outer-ring wear.



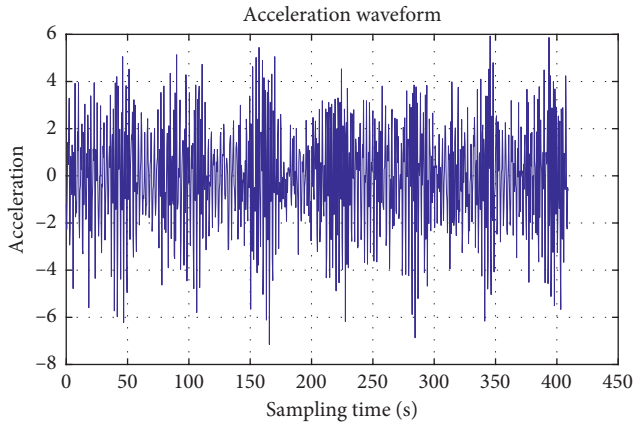


FIGURE 10: Time-domain waveform of inner-ring wear.

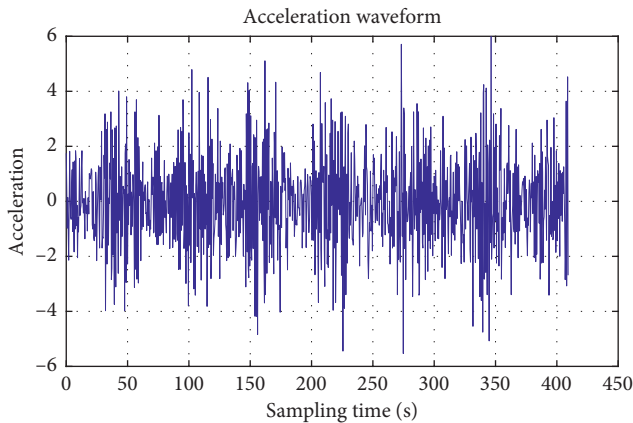


FIGURE 11: Time-domain waveform of short of ball bearing.

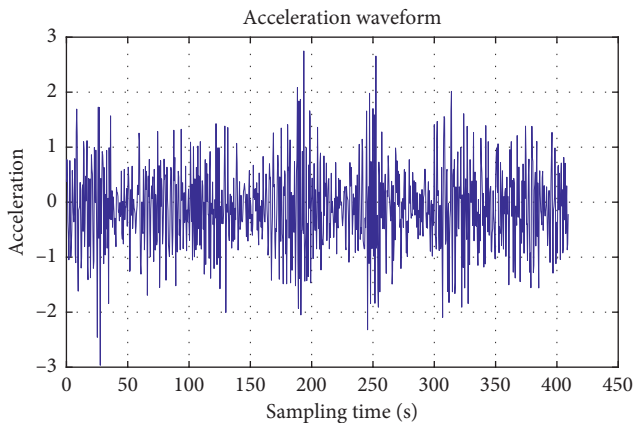


FIGURE 12: Deficiency time-domain waveform of large and small gear teeth.

#### 4. Simulation and Results

Simulations of the running states of a normal rolling bearing, outer-ring wear, inner-ring wear, lack of ball bearings, and large and small gear with five kinds of tooth fault were

performed. For each state, 50 sets of data were randomly selected with a total sample of 250 groups. The `randperm` function in MATLAB was used to randomly generate data and 50% of the total data were used as training data, while the remaining 50% were used as the test data. On each run, different results were generated by the program since the training set and testing set were randomly generated. Figure 13 corresponds to the result of one random experiment with the accuracy rate of 86.4%. The “True label” represents the true type of the unit, and “Predict label” represents the type predicted by this algorithm.

As seen in Table 2, the normal bearing data can be identified extremely well using the algorithm proposed in this paper, with an accuracy of 100%, while the accuracy of inner-ring wear fault data reached 88.9%. However, the bearing outer-ring wear and missing ball bearing were not well identified.

To test the practicality and feasibility of the algorithm for bearing fault diagnosis in engineering applications, the traditional cross-validation algorithm and SVM algorithm based on the genetic algorithm were compared. First, 80 groups were randomly selected for each type from the measured data. The types of machine operation status and corresponding groups are shown in Table 3. Then, the random function in MATLAB was used to generate three data sets from which half of the data were taken as training data, and the remaining were taken as test data. Finally, the simulation was performed, and the accuracies were obtained. The comparing results are shown in Tables 4–6.

The results of the comparison show that the time required by each of the three algorithms increases with the number of training samples. However, compared with the SGA-SVM and MPGA-SVM, the increase is more evident in time required for training the cross-validation algorithm. For the first group, No. 1 of Tables 4–6, corresponding to the machine operation status of only the normal bearings and bearing outer-ring wear, we see that the cross-validation algorithm, SGA-SVM, and MPGA-SVM can correctly classify the data; however, for other fault types, the diagnosis accuracy declines. Observing the accuracy rates of the second and third groups, the MPGA-SVM and cross-validation algorithms are similar and higher than using the SGA-SVM algorithm.

In addition, we compared the algorithm in this paper with the PPMCC fuse nondimension algorithm that was proposed in reference [32]. Six types of online data sets for petrochemical units, including lack of bearing ball, wear of bearing’s inner ring, wear of bearing’s outer ring, and missing teeth of the large gear and bent shaft were compared. In this comparison, the two algorithms used the same test data sets and training data sets. The comparison result is shown in Table 7.

In Table 7, we can find that the accuracy got from MPGA-SVM is higher than that from the PPMCC in different machine operation status. That is, to say, the algorithm presented in this paper is obviously similar to that proposed by reference [32].

TABLE 1: Machine operation status and number for each category of bearing.

Label	Label 1	Label 2	Label 3	Label 4	Label 5
Machine operation status	Normal bearing	Outer-ring wear	Inner-ring wear	Short of ball bearing	Large and small gears teeth deficiency
Number	1–50	51–100	101–150	150–200	201–250

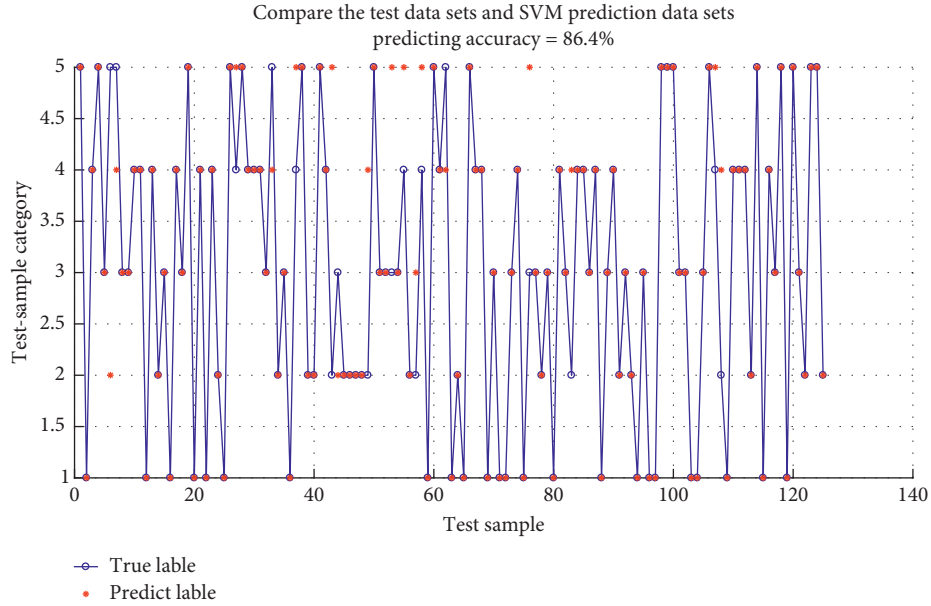


FIGURE 13: Distribution map of prediction sample and real sample.

TABLE 2: Result of machine operation status prediction.

Type	Label 1	Label 2	Label 3	Label 4	Label 5
Prediction number	24	22	27	29	23
Accuracy	100%	77.3%	88.9%	82.8%	82.6%

TABLE 3: Types of machine operation status corresponding to each group.

Number	Label 1	Label 2	Label 3	Label 4
No. 1	80	80	0	0
No. 2	80	80	80	0
No. 3	80	80	80	80

TABLE 4: Comparison of different algorithms.

Method	Cross-validation SVM				SGA-SVM				MPGA-SVM			
	Accuracy (%)	Time (s)	C	$\gamma$	Accuracy (%)	Time (s)	C	$\gamma$	Accuracy (%)	Time (s)	C	$\gamma$
Number												
No. 1	100	13.62	0.036	0.574	100	1.62	6.776	6.752	100	3.23	4.310	6.031
No. 2	93.33	32.76	0.047	2.297	90	3.56	5.626	7.803	91.67	5.02	0.851	3.184
No. 3	82.5	64.42	0.660	16	80.63	5.899	7.007	6.147	82.5	9.88	1.160	9.567

TABLE 5: Comparison of different algorithms.

Method	Cross-validation SVM				SGA-SVM				MPGA-SVM			
	Accuracy (%)	Time (s)	C	$\gamma$	Accuracy (%)	Time (s)	C	$\gamma$	Accuracy (%)	Time (s)	C	$\gamma$
Number												
No. 1	100	13.55	0.047	0.435	100	1.22	2.205	0.599	100	2.67	0.548	0.553
No. 2	93.33	36.47	0.054	1.741	91.67	3.244	2.128	0.057	94.17	5.19	0.605	2.584
No. 3	84.38	64.05	2.639	12.126	83.75	6.40	5.831	9.863	83.75	10.7449	2.863	9.989

TABLE 6: Comparison of different algorithms.

Method	Cross-validation SVM				SGA-SVM				MPGA-SVM			
	Accuracy (%)	Time (s)	C	$\gamma$	Accuracy (%)	Time (s)	C	$\gamma$	Accuracy (%)	Time (s)	C	$\gamma$
Number												
No. 1	100	14.13	0.041	1	100	1.44	8.808	9.168	100	3.09	7.393	5.949
No. 2	90	33.97	0.082	3.482	90.83	2.30	0.294	1.454	93.33	3.91	9.421	0.878
No. 3	84.38	61.86	18.379	1.516	83.13	4.68	9.362	1.115	84.38	7.75	3.283	2.798

TABLE 7: Comparison of different algorithms.

Machine operation status	MPGA-SVM model (%)	PPMCC model (%)
Normal	92.5	80
Lack of ball bearing	41.25	27.5
Wear of bearing's outer ring	40	27.5
Wear of bearing's inner ring	68.75	13.75
Missing teeth of the large gear	25	13.5
Bent shaft	60	18.75

## 5. Discussion

As shown in Table 4, the training times are clearly different between algorithms. The multigroup genetic algorithm can search for the global optimal value in a short time due to implicit parallelism and powerful global search capabilities for optimizing results. This has also been seen in practical engineering applications, whereby if the number of training samples is too large, the time required for cross-validation algorithms is so long that it is difficult to achieve. The classification accuracy rates of SVM trained by cross-validation algorithms and multiple swarm genetic algorithms are higher than genetic algorithms, so it is clear by comparison that the combination of multipopulation genetic algorithms and SVM is a compromise between time and accuracy.

In addition, the objective function of the genetic algorithm must meet the accuracy of the K-CV of the SVM, which enables the model to avoid overlearning [33]. If the accuracy rate of a single training of an SVM is used as the fitness value for multiple genetic algorithms without using the cross-validation algorithm, the classification accuracy rate of the prediction will have large disparities, as high as 40%, leading to overlearning and loss of generalization.

## 6. Conclusions

To solve the difficulties of parameter selection for SVM, a method of optimizing the SVM parameters using

multipopulation genetic algorithm was proposed in this paper. The algorithm uses the average accuracy rate as the objective function in the K-CV method of the SVM and then selects the important parameters using a multiple group genetic algorithm. The two main advantages of the algorithm are that it can achieve the ideal effect and requires shorter training times. Our simulation results show that the algorithm is effective, providing a feasible method for parameter optimization of the SVM, and the experimental results demonstrate the practical values of the algorithm for bearing fault diagnosis. However, in Tables 2, 4, 5 and 7, we can find that a relatively low-diagnostic accuracy rate still exists for some types of fault, and the diagnosis accuracy declines with the increase of types of machine operation status, and our experiments have shown the method proposed in the paper is not perfect.

Therefore, our future research will focus on adjusting the algorithm in terms of the following aspects: the  $P_c$  and  $P_M$  value ranges, at this stage, were only set by experts' suggestions, as well as the  $C$  and  $\gamma$  ranges. The appropriate range of  $C$  and  $\gamma$  can not only save time for the algorithm to find parameter values but also make it probable to find better parameter values. Thus, a better fault diagnosis model will be obtained.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant no. 61473331, in part by the Science and Technology Plan of Guangdong Province of China under Grant no. 2017A070712024, in part by the Sail Plan Training High-Level Talents of Guangdong

Province of China, in part by the Introduction of Talents' Project of Guangdong Polytechnic Normal University of China under Grant no. 991512203, in part by the 2016 Annual Scientific and Technological Innovation Special Fund to Foster Students' Projects of Guangdong under Grant no. pdjh2016b0341, in part by the Guangdong University of Petrochemical Technology College Students Innovation Incubation Project under Grant no. 2015pyA006, in part by the Science and Technology Project of Guangzhou under Grant nos. 201604010099, 2016B030306002, and 2016B030308001, in part by the Fundamental Research Funds for the Central Universities under Grant no. x2jqD2170480, in part by Guangdong Province Science and Technology Major Special Projects (Grant no. 2017B030305004), Guangdong Province Science and Technology Application of Major Special Projects (Grant no. 2016B020243011), and Major Provincial Scientific Research Projects of Guangdong Normal Universities (Grant no. 2017KZDXM052).

## References

- [1] A. Ashari, R. Nikoukhah, and S. Campbell, "Active robust fault detection in closed-loop systems: quadratic optimization approach," *IEEE Transactions on Automatic Control*, vol. 57, no. 10, pp. 2532–2544, 2012.
- [2] A. Grastien and A. Anbulagan, "Diagnosis of discrete event systems using satisfiability algorithms: a theoretical and empirical study," *IEEE Transactions on Automatic Control*, vol. 58, no. 12, pp. 3070–3083, 2013.
- [3] J. B. Xiong, Q. H. Zhang, Z. P. Peng, G. X. Sun, and Y. D. Cai, "Double sample data fusion method based on combination rules," *IEEE Access*, vol. 4, pp. 7487–7499, 2016.
- [4] H. Zuo, J. M. Gu, and J. L. Yu, "Fault diagnosis of rolling bearings based on supporting vector machine," *Bearing*, vol. 8, pp. 36–39, 2008.
- [5] A. S. Qin, Q. H. Zhang, Q. Hu, G. X. Sun, J. He, and S. Q. Lin, "Remaining useful life prediction for rotating machinery based on optimal degradation indicator," *Shock and Vibration*, vol. 2017, Article ID 6754968, 12 pages, 2017.
- [6] F. Dalvand, S. Dalvand, F. Sharafi, and M. Pecht, "Current noise cancellation for bearing fault diagnosis using time shifting," *IEEE Transactions on Industrial Electronics*, 2017, in press.
- [7] D. H. Zhou, Y. Liu, and X. Hw, "Review on fault diagnosis techniques for closed-loop systems," *Acta Automatica Sinica*, vol. 39, no. 11, pp. 1933–1943, 2013.
- [8] C. Shen, Q. He, F. Kong, and P. W. Tse, "A fast and adaptive varying-scale morphological analysis method for rolling element bearing fault diagnosis," *Proceedings of Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 227, no. 6, pp. 1362–1370, 2013.
- [9] Y. Chen, *The Research of Fault Diagnosis of Rolling Bearing Based on Mbration Signal Analysis Method*, Lanzhou University of Technology, Lanzhou, China, 2014.
- [10] X. H. Chen and X. Y. Xiong, "Fault diagnosis of ball bearing based on wavelet and radial basis function neural networks," *Mechanical Engineering and Automation*, vol. 1, pp. 13–15, 2006.
- [11] P. J. Gui and J. Z. Lin, "Fault diagnosis of rolling bearings based on artificial neural network," *Machinery*, vol. 31, no. 10, pp. 58–60, 2004.
- [12] F. Na, "Fault diagnosis which based on support vector machine," *Industrial Control Computer*, vol. 23, no. 12, pp. 45–46, 2010.
- [13] F. F. Xie, J. Zhang, and J. Zheng, "Support vector machine for diagnosis of mechanical fault," *Sensor World*, vol. 12, no. 4, pp. 42–44, 2006.
- [14] D. W. Zhang, Z. M. Duan, P. Li, and X. H. Zhang, "Fault diagnosis method based on genetic algorithm and LSSVM," *Computer Measurement and Control*, vol. 18, no. 8, pp. 1745–1747, 2010.
- [15] J. X. Yun, "Fault diagnosis for rolling bearing on genetic-SVM classifier," *Advanced Materials Research*, vol. 199–200, pp. 620–624, 2011.
- [16] Y. N. Zhu, G. H. Wu, and J. Y. Du, "Facial expression recognition based on the multi-species genetic algorithm," *Journal of Hangzhou Dianzi University*, vol. 33, no. 1, pp. 44–47, 2013.
- [17] J. Wang and H. Y. Liu, "Support vector machine and genetic algorithm-based tool failure diagnosis," *Modular Machine Tool and Automatic Manufacturing Technique*, vol. 1, pp. 74–76, 2013.
- [18] J. H. Holland, *Adaptation in Natural and Artificial System: An Introduction with Application to Biology, Control and Artificial Intelligence*, University of Michigan Press, Ann Arbor, MI, USA, 1975.
- [19] Y. H. Jiang, G. M. Cheng, J. W. Kan, Z. Y. Xuan, J. J. Ma, and Z. H. Zhang, "Rolling bearing fault diagnosis based on NGA optimized SVM," *Chinese Journal of Scientific Instrument*, vol. 34, no. 12, pp. 2684–2689, 2013.
- [20] Q. Y. Wang, Y. Q. He, and X. L. Peng, "Parameters optimization of support vector machine based on improved genetic algorithm," *Computer and Modernization*, vol. 3, pp. 33–36, 2015.
- [21] M. Srinivas and L. Patnaik, "Adaptive probabilities of crossover and mutation in genetic algorithms," *IEEE Transactions on Systems Man and Cybernetics*, vol. 24, no. 4, pp. 656–667, 2002.
- [22] D. P. Liu, G. L. Shan, Q. L. Zhang, and X. S. Duan, "Parameters optimization of support vector machine based on improved genetic algorithm," *Microcomputer Applications*, vol. 31, no. 5, pp. 11–15, 2010.
- [23] L. Yu, F. Shi, H. Wang, and F. Hu, *30 Analysis of Case Algorithms*, Beihang University Press, Beijing, China, 2015.
- [24] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag New York, Inc, New York, NY, USA, 1995.
- [25] V. Vapnik, *Statistical Learning Theory: DBLP*, John Wiley & Sons Inc., New York, 2010.
- [26] O. Chapelle, P. Haffner, and V. N. Vapnik, "Support vector machines for histogram-based image classification," *IEEE Transactions on Neural Networks*, vol. 10, no. 5, pp. 1055–1064, 1999.
- [27] G. Isabelle, W. Janson, B. Stephen et al., "Gene selection for cancer classification using support vector machines," *Machine Learning*, vol. 46, pp. 389–422, 2002.
- [28] Y. Q. Gen, *Research on Locomotive Bearing Fault Diagnosis System Based on Support. Vector Machine*, Central South University, Changsha, China, 2008.
- [29] H. Z. Wang, K. Liu, and H. Y. Zhu, "Load forecasting of power system based on cloud model and support vector machine with RBF," *Journal of Lanzhou University of Technology*, vol. 42, no. 4, pp. 85–88, 2016.
- [30] Y. Y. Ye, *Fault Diagnosis of Rolling Bearing Based on Optimized Support Vector Machine*, Fuzhou University, Fuzhou, China, 2014.

- [31] J. B. Xiong, Q. H. Zhang, J. F. Wang, L. Liang, P. H. Cheng, and Q. Liang, "Data fusion method based on mutual dimensionless," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 2, pp. 506–517, 2017.
- [32] J. B. Xiong, Q. Liang, J. F. Wan, Q. H. Zhang, X. H. Chen, and R. B. Ma, "The order statistics correlation coefficient and PPMCC fuse non-dimension in fault diagnosis of rotating petrochemical unit," *IEEE Sensors Journal*, vol. 18, no. 11, pp. 4704–4714, 2018.
- [33] K. Wang, Z. R. Hou, and C. L. Wang, "Intrusion detection based on cross-validation SVM," *Journal of Test and Measurement Technology*, vol. 24, no. 5, pp. 419–423, 2010.





**Hindawi**

Submit your manuscripts at  
[www.hindawi.com](http://www.hindawi.com)

