
Backtracking Search Optimization Algorithm and its Application to Roller Bearing Fault Diagnosis

HungLinh Ao, T. Nguyen Thoi, V. Ho Huu, Linh Anh-Le and TrangThao Nguyen

Division of Computational Mathematics and Engineering (CME), Institute for Computational Science (INCOS), Ton Duc Thang University, Ho Chi Minh City, 700000, Vietnam.

Faculty of Civil Engineering, Ton Duc Thang University, Ho Chi Minh City, 700000, Vietnam.

Minh Quang Chau

Faculty of Mechanical Engineering, Industrial University of Ho Chi Minh City, Ho Chi Minh City, 700000, Vietnam.

(Received 27 November 2014; accepted 10 June 2016)

It is clearly known that support vector machine (SVM) parameters have significant effects on the accurate rate of classification result. Adjusting the SVM parameters improves its effectiveness and accuracy, which is always a challenge. On the other font, the Backtracking Search Optimization Algorithm (BSOA), an evolutionary algorithm for solving optimization problems, is proposed and proven to be effective through various benchmark problems. This paper proposes an optimization method for the SVM parameters based on BSOA. For convenience, the proposed method has been named BSOA-SVM. This method is tested with some real-world benchmark data sets to verify its robustness and effectiveness. Then, BSOA-SVM is applied for diagnosing roller bearing fault, which is a real world problem. In this diagnosing process, the original acceleration vibration signals are first decomposed into product function (PFs) by using the local mean decomposition (LMD) method. Next, initial feature matrices are extracted from PFs by singular value decomposition (SVD) techniques to give single values. Finally, these values serve as input vectors for the BSOA-SVM classifier. The results from the problem show that the combination of the BSOA-SVM classifiers obtains higher classification accuracy with a lower cost time compared to other methods.

1. INTRODUCTION

Optimization of SVM parameters has always been a complex task for researchers since it was developed. In recent years, many algorithms were employed to handle this task, such as the trial and error procedures,¹ the grid algorithm,² the cross-validation method,³ the generalization error estimation method,⁴ the gradient descent method,⁵ and so on. Unfortunately, these methods still contain some drawbacks that hamper the effectiveness of SVM. For example, the grid method requires complex computations and is time consuming while the cross-validation method also requires long and complicated calculations.² The heuristic algorithms, such as the genetic algorithm (GA), the particle swarm optimization (PSO),⁶ and the ant colony optimization (ACO)⁷ were also used to optimize SVM parameters. However, PSO is easily trapped into the local optimization areas⁸ while GA has an expensive computational cost.⁹

Recently, Pinar Civicioglu developed the Backtracking Search Optimization Algorithm (BSOA), which is an evolutionary algorithm (EA) for solving optimization problems. The BSOA method could solve real-valued numerical optimization problems for a short time and the search result was better than other EAs.¹⁰ Unlike other methods in the EA group, BSOA has only one control parameter in the algorithm. This makes the method much simpler to use. Therefore, in this paper, BSOA is combined with the SVM to give a so called BSOA-SVM for solving classification problems. BSOA-SVM was applied to diagnose the fault of roller bearing. In this diagnosing process, the original acceleration vibration signals were first decomposed into product function (PFs) by using the LMD method. Next, initial feature matrices were extracted from PFs by singular value decomposition (SVD) techniques

to give single values. Finally, these values served as input vector for the BSOA-SVM classifier. The classification results of the proposed method show a higher accuracy and lower cost time compared with the GA-SVM, the PSO-SVM, and the CMAES-SVM methods.

The rest of this paper is organized as follows: in Section 2, the BSOA method is briefly reviewed. In Section 3, the parameter optimization algorithm based on the BSOA method is addressed. The fault diagnosis method based on LMD-SVD and BSOA-SVM, in which initial feature matrices extracted from a number of PFs are used as input vectors of BSOA-SVM, is presented in Section 4. In Section 5, the fault diagnosis method is used to diagnose the condition of actual roller bearings and is compared with the GA, the PSO, and the CMAES methods. Finally, the paper is concluded in Section 6.

2. BACKTRACKING SEARCH OPTIMIZATION ALGORITHM

BSOA is an adaptive search algorithm that uses three basis genetic operators including selection, mutation, and crossover to generate trial individuals. The principle of BSOA, which consists of six steps,¹⁰ is presented in the flow chart in Fig. 1. More details of the steps are presented in following sections.

2.1. Define the Problem and Algorithm Parameter

The mathematical formulation of a typical optimization can be written as:

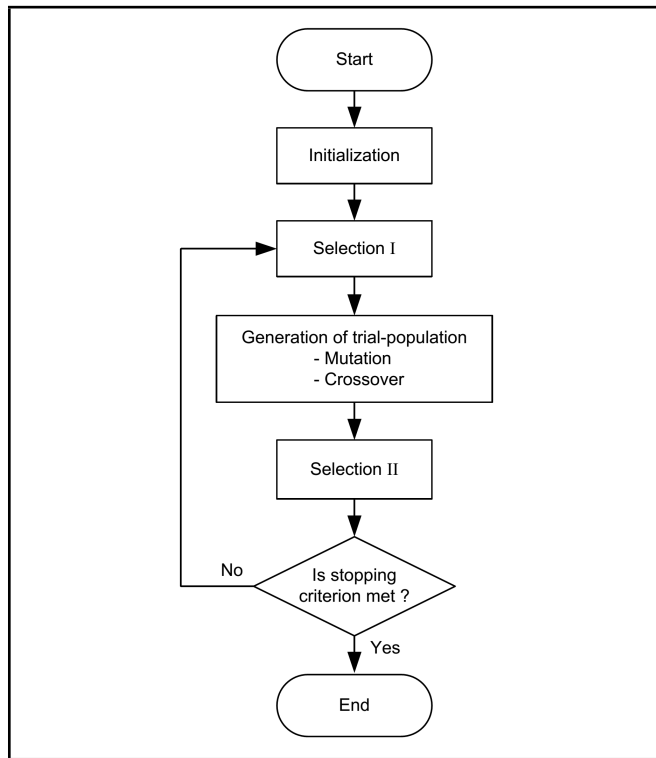


Figure 1. The flow chart for the Backtracking Search Optimization Algorithm (BSOA).

$$\begin{aligned} &\text{Minimize } f(x) \\ &\text{subject to } x_i \in D_i = [l_i, u_i], \quad i = \{1, 2, \dots, N\}. \end{aligned} \quad (1)$$

where $f(x)$ is a fitness function, $x = (x_1, x_2, \dots, x_N)$ is the vector of decision variables, N is the number of decision variables, and D_i is the range of feasible values for the i -th decision variable, where l_i and u_i are the lower and upper bounds of the i -th decision variable, respectively.

2.2. Initialization

In this phase, the setting of the algorithm was initialized and the values of algorithmic parameters were assigned. Population P was initialized as follows:

$$P_{i,j} \sim W(l_i, u_i); \quad (2)$$

for $i = 1, 2, 3, \dots, D$ and $j = 1, 2, 3, \dots, N$, where N was the population size (*PopSize*), D was the dimension of the problem, W was the uniform distribution, and each P_j was a target individual in the population P .

2.3. Selection-I

This stage was aimed to determine the previous population P' ,

$$P'_{i,j} \sim W(l_i, u_i). \quad (3)$$

Based on 'if-then' rule, the option of redefining P' at the beginning of iteration could be expressed as follows:

$$\text{If } a < b \text{ then } P' := P \mid a, b \sim W(0, 1); \quad (4)$$

where ' $:=$ ' was the update operation and a and b were random numbers. Next,

$$P' := \text{permuting}(P'); \quad (5)$$

where the permuting function was a random shuffling function.

2.4. Mutation

This process generated the initial form of the trial population, M_u , as follows:

$$M_u = P + F(P' - P); \quad (6)$$

where F was the control parameter that controlled the amplitude of the search-direction matrix $(P' - P)$. The value of this parameter was selected as per the following equation:¹⁰

$$F = 3r_n; \quad (7)$$

where $r_n \sim M(0, 1)$, M was the standard normal distribution, and F was the controlled parameter.

2.5. Crossover

This process generated the final form of the trial population T , which was updated with:

$$T_{i,j} := P_{i,j} \text{ if } \text{map}_{i,j} = 1; \quad (8)$$

where $i \in 1, 2, 3, \dots, D$ and $j \in 1, 2, 3, \dots, D$.

The crossover strategy of BSOA used the ceiling function to define $\text{rnd} \sim W(0, 1)$. Furthermore, the number of elements of individuals was controlled by a mix rate parameter (*mixrate*). These individuals were mutated in a trial by using the ceiling function $\lceil \text{mixrate} \cdot \text{rnd} \cdot D \rceil$.

2.6. Selection-II

Based on greedy selection, the T_j s that had better fitness values than the corresponding P_j s, were used to update the P_j s.

In this stage, the boundary control mechanism was also used to form the limit search space.¹⁰

3. PARAMETER OPTIMIZATION OF SVM BASED ON BSOA

3.1. The Support Vector Machine (SVM)

SVM is a kind of machine learning techniques that is based on the statistical learning theory. The basic idea of SVM is mapping the training samples from the input space into a higher-dimensional feature space by using a mapping function ϕ .⁷ Suppose that there was a given training sample set $G = \{(x_i, y_i), i = 1, 2, \dots, l\}$, where each sample $x_i \in R^d$ belonged to a class by $y \in \{+1, -1\}$ and the training data was not linearly separable in feature space, then the target function could be expressed as follows:¹¹

$$\begin{aligned} &\text{Minimize } \phi(\omega) = \frac{1}{2} \langle \omega, \omega \rangle + C \sum_{i=1}^l \xi_i \\ &\text{subject to } y_i (\langle \omega, \phi(x_i) \rangle + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \\ & \quad \quad \quad i = \{1, 2, \dots, l\}. \end{aligned} \quad (9)$$

where ω was the normal vector of the hyperplane, C was the penalty parameter, b was the bias, ξ_i were nonnegative slack variables, and $\phi(x)$ was the mapping function.

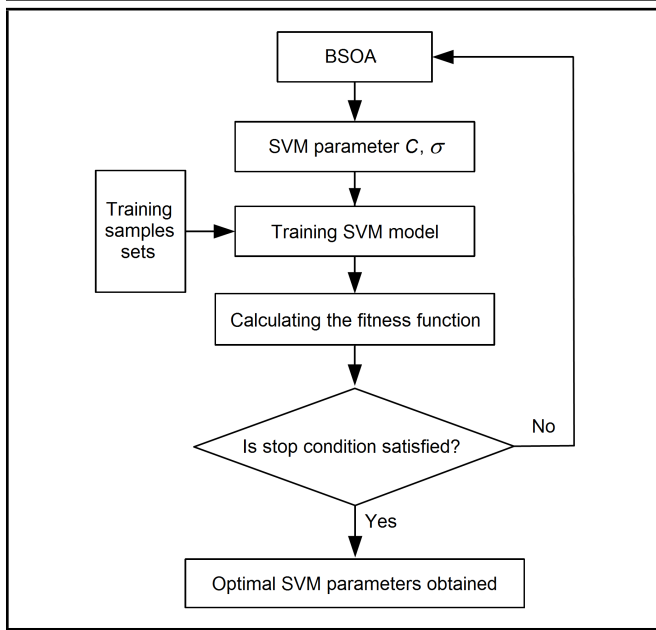


Figure 2. The parameter optimization flowchart of SVM based on BSOA.

By introducing a set of Lagrange multipliers $\alpha_i \geq 0$, the optimization problem could be rewritten as:

$$\begin{aligned} \text{Maximize } L(\omega, b, \alpha) &= \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\ \text{subject to } 0 \leq \alpha_i \leq C, \quad &\sum_{i=1}^l \alpha_i y_i = 0. \end{aligned} \quad (10)$$

The decision function can be obtained as:⁷

$$f(x) = \text{sgn} \left[\sum_{i=1}^l \alpha_i y_i K(x_i, x) + b \right]. \quad (11)$$

The most common kernel function used in the SVM method is the radial basis function kernel, as shown in the following equation:¹¹

$$K(x, x_i) = \exp(-\|x - x_i\|^2 / 2\sigma^2); \quad (12)$$

where σ is the kernel parameter.

3.2. Parameter Optimization of SVM Based on BSOA

It is widely known that the performance of SVM is significantly affected by its parameters. The parameters need to be chosen as a penalty factor C and the kernel parameter σ in the Gaussian kernel function. Selecting these parameters was not an easy task. Generally, C and σ are selected based on experience. Therefore, in this paper, BSOA was used to optimize the parameters of SVM. As a result, C and σ became the variables and the test error of SVM became the fitness function in the optimization problem. The test error of SVM was given as follows:

$$\text{fitness}(x) = \text{Test_Error}_{\text{SVM}}(x); \quad (13)$$

where $x = (C, \sigma)$ and the test error of SVM was defined as:

$$\text{Test_Error}_{\text{SVM}} = \frac{\text{Number of incorrect classification in test samples}}{\text{Total number of samples in test set}}. \quad (14)$$

The flow chart of BSOA-SVM is shown in Fig. 2.

Table 1. Properties of the problems.

Name	Data	Train	Test	Input	Class
Iris	150	105	45	4	3
Thyroid	215	151	64	5	3
Seed	210	147	63	7	3
Wine	178	125	53	13	3
<i>E. coli</i>	327	229	98	7	5
Glass	214	149	65	9	6

3.3. Experimental Results

To evaluate the performance of the proposed BSOA-SVM method, six common benchmark data sets from the University of California Irvine (UCI) benchmark were used, including the Iris, Thyroid, Seed, Wine, *Escherichia coli* (*E. coli*), and Glass data sets. Table 1 gives the sizes of the training and test sets. Each sample set was divided into two sub-sets: one for training the SVM and one for testing the obtained model. The proportion of the training and test sets were 70% and 30% of the total samples, respectively. This proportion was chosen based on trial and error so that the performance of the obtained SVM was optimal with the available samples.

The BSOA-SVM, the GA-SVM, the PSO-SVM, and the CMAES-SVM methods were used to classify these data sets. To make a fair comparison, the values of these four methods were chosen to be the same (e.g., *iteration* = 30 and *PopSize* = 30). For the PSO, the parameters were fixed with the values given in the literature¹² (i.e., $W = 0.9$, $c_1 = 0.5$, and $c_2 = 1.25$). For CMAES, the parameters were fixed with the values given in the literature¹³ (i.e., $\sigma' = 0.25$ and $\mu = 4 + 3 \log(N)$). The testing results of each method were taken as the averaged value of 30 runs. The training data and test data were both mixed and randomly divided, as seen in Table 1.

According to Lin et al.,¹⁴ the lower and upper bounds of C were given in [0.01, 35000] and σ in [0.01, 32] for the BSOA-SVM, GA-SVM, PSO-SVM, and CMAES-SVM classifiers. Each search method gave the values of C and σ in order to give the smallest value of the classification error. These results are shown in Table 2.

The detailed classification results of the each data set are provided in Tables 2 to 7 so as to illustrate the effectiveness of the proposed method. The Iris, Thyroid, Seed, and Wine data sets included three classes, so we needed two SVM classifiers. The *E. Coli* data set included five classes, so we needed four classifiers. The Glass data set included six classes, so we needed five classifiers. These tables show the optimal parameters (C and σ), the average test error, and the average cost time done by different algorithms.

Table 8 gives the classification results of the Iris, Thyroid, Seed, Wine, *E. Coli*, and Glass data sets, respectively. It can be seen from Table 8 that the test error and the cost time of BSOA-SVM was lower than the one by the GA-SVM, the PSO-SVM, and the CMAES-SVM methods. According to Civicioglu,¹⁰ BSOA used a mutation mechanism with one individual and a complex crossover mechanism. Furthermore, BSOA took advantage of the experiences that were obtained from previous generations by using its memory. From the tables, it can be seen that the BSOA-SVM classifier obtained a higher classification accuracy in a shorter amount of time compared to other methods. The BSOA-SVM method was next applied to a roller bearing fault diagnosis problem.

Table 2. The identification results of the Iris data set.

Method	Training samples	Test samples	Optimal C	Optimal σ	Average cost time (s)	Average test error (%)
BSOA-SVM1	105	45	229.32	26.11	9.21	0.000
PSO-SVM1	105	45	214.60	3.99	22.56	0.215
GA-SVM1	105	45	306.82	18.47	17.17	0.242
CMAES-SVM1	105	45	299.34	7.67	15.28	1.411
BSOA-SVM2	75	30	265.76	21.68	7.41	3.112
PSO-SVM2	75	30	320.75	5.55	15.77	4.118
GA-SVM2	75	30	169.36	31.72	13.37	4.202
CMAES-SVM2	75	30	206.49	11.94	13.57	4.811

Table 3. The identification results of the Thyroid data set.

Method	Training samples	Test samples	Optimal C	Optimal σ	Average cost time (s)	Average test error (%)
BSOA-SVM1	151	64	29300.97	0.36	17.28	2.344
PSO-SVM1	151	64	21607.25	10.44	39.31	2.358
GA-SVM1	151	64	6880.24	0.75	32.18	2.442
CMAES-SVM1	151	64	24468.10	10.54	30.13	2.922
BSOA-SVM2	45	20	32570.13	24.75	16.24	0.000
PSO-SVM2	45	20	8896.08	4.74	37.54	0.089
GA-SVM2	45	20	23584.73	29.99	29.18	0.265
CMAES-SVM2	45	20	10571.69	27.29	27.75	0.253

Table 4. The identification results of the Seed data set.

Method	Training samples	Test samples	Optimal C	Optimal σ	Average cost time (s)	Average test error (%)
BSOA-SVM1	147	63	16099.27	20.95	24.53	2.910
PSO-SVM1	147	63	27681.87	13.98	48.82	2.918
GA-SVM1	147	63	8338.74	19.4	39.28	4.589
CMAES-SVM1	147	63	18234.38	27.26	39.21	2.997
BSOA-SVM2	98	42	20976.92	24.35	16.97	0.000
PSO-SVM2	98	42	34276.51	5.6	44.98	0.098
GA-SVM2	98	42	27888.75	7.39	35.17	1.546
CMAES-SVM2	98	42	28738.10	17.54	36.19	0.282

Table 5. The identification results of the Wine data set.

Method	Training samples	Test samples	Optimal C	Optimal σ	Average cost time (s)	Average test error (%)
BSOA-SVM1	124	54	9411.199	5.867	14.92	0.0617
PSO-SVM1	124	54	32665.452	4.422	33.02	0.0617
GA-SVM1	124	54	0.704	2.792	22.11	2.1711
CMAES-SVM1	124	54	22099.552	4.286	24.31	0.1852
BSOA-SVM2	84	35	18028.496	30.345	14.61	1.3836
PSO-SVM2	84	35	26695.179	29.837	33.89	1.5723
GA-SVM2	84	35	2.925	2.145	30.91	2.5786
CMAES-SVM2	84	35	4441.430	2.473	26.32	2.3899

Table 6. The identification results of the *E. coli* data set.

Method	Training samples	Test samples	Optimal C	Optimal σ	Average cost time (s)	Average test error (%)
BSOA-SVM1	229	98	7920.09	23.81	26.73	4.341
PSO-SVM1	229	98	6174.79	28.60	59.43	4.364
GA-SVM1	229	98	290.97	17.86	44.78	4.797
CMAES-SVM1	229	98	8213.08	31.02	42.52	7.814
BSOA-SVM2	130	54	25949.05	14.46	25.21	10.200
PSO-SVM2	130	54	19606.28	12.53	53.24	10.200
GA-SVM2	130	54	19860.53	12.83	42.91	10.892
CMAES-SVM2	130	54	337.55	32.00	40.41	11.925
BSOA-SVM3	75	32	33525.82	8.92	23.38	7.742
PSO-SVM3	75	32	33965.60	6.56	47.33	7.777
GA-SVM3	75	32	8421.75	13.26	39.83	7.972
CMAES-SVM3	75	32	31350.95	26.08	38.22	8.728
BSOA-SVM4	50	22	165.11	25.69	17.21	6.315
PSO-SVM4	50	22	10.77	5.89	44.25	6.382
GA-SVM4	50	22	1.01	1.45	37.81	7.111
CMAES-SVM4	50	22	331.07	4.83	37.37	7.647

Table 7. The identification results of the Glass data set.

Method	Training samples	Test samples	Optimal C	Optimal σ	Average cost time (s)	Average test error (%)
BSOA-SVM1	149	65	19822.58	27.91	64.12	0.620
PSO-SVM1	149	65	34301.94	29.24	139.92	0.626
GA-SVM1	149	65	10226.04	3.67	134.82	0.854
CMAES-SVM1	149	65	14153.06	18.37	91.15	0.462
BSOA-SVM2	100	44	504.61	19.37	62.45	2.080
PSO-SVM2	100	44	839.26	29.64	137.53	2.084
GA-SVM2	100	44	11808.45	6.84	110.35	2.365
CMAES-SVM2	100	44	2391.84	24.94	85.83	2.137
BSOA-SVM3	48	20	31452.79	8.63	58.61	0.000
PSO-SVM3	48	20	32529.36	3.23	134.84	0.000
GA-SVM3	48	20	27770.96	11.83	102.26	0.000
CMAES-SVM3	48	20	26416.18	12.41	84.72	0.000
BSOA-SVM4	36	15	34617.43	2.74	57.41	3.050
PSO-SVM4	36	15	34939.12	5.61	131.56	3.065
GA-SVM4	36	15	20462.37	26.62	98.67	6.492
CMAES-SVM4	36	15	31631.04	9.56	80.21	6.321
BSOA-SVM5	27	11	26017.53	1.47	55.62	0.110
PSO-SVM5	27	11	21049.01	6.19	128.82	0.110
GA-SVM5	27	11	28350.18	19.44	86.19	14.247
CMAES-SVM5	27	11	26899.62	27.46	77.24	1.981

Table 8. The average test error and cost time of the proposed BSOA-SVM compared with the GA-SVM, the PSO-SVM, and the CMAES-SVM (%).

Data	Average of error and cost time	Method			
		BSOA-SVM	PSO-SVM	GA-SVM	CMAES-SVM
Iris	Error (%)	1.556	2.167	2.222	3.111
	Time (s)	8.31	19.17	15.27	14.43
Thyroid	Error (%)	1.172	1.224	1.354	1.588
	Time (s)	16.76	38.43	30.68	28.94
Seed	Error (%)	1.455	1.508	3.068	1.640
	Time (s)	20.75	46.90	37.23	37.70
Wine	Error (%)	0.722	0.817	2.369	1.287
	Time (s)	29.53	66.91	53.02	50.63
<i>E. coli</i>	Error (%)	7.150	7.181	7.693	9.029
	Time (s)	23.13	52.06	41.33	39.63
Glass	Error (%)	1.172	1.177	4.792	2.180
	Time (s)	59.64	134.53	106.46	83.83

4. BSOA-SVM AND LMD-SVD FOR ROLLER FAULT DIAGNOSIS PROBLEM

4.1. The Local Mean Decomposition (LMD) Method

The LMD method was developed from the simple assumption that any complicated signal consists of several product functions (PF s).¹⁵ In this way, each signal could be decomposed into a number of PF s and a residue $r_n(t)$:

$$x(t) = \sum_{p=1}^n PF_p(t) + r_n(t); \quad (15)$$

where p is the number of the product function, r_n is a monotonic function.¹⁵

4.2. The Singular Value Decomposition (SVD) Technique

The SVD technique is aimed to decompose a matrix into three matrices: singular values and singular vectors of initial matrix. Assuming that there was a matrix Σ , which had $M \times N$ dimension, and was expressed in the form of

$$\Sigma = E\Delta V^T; \quad (16)$$

where $E = [e_1, e_2, e_3, \dots, e_n] \in R^{N \times N}$, $E^T E = I$, $V = [v_1, v_2, v_3, \dots, v_n] \in R^{M \times M}$, $V^T V = I$, $\Delta \in R^{N \times M}$, $\Delta = [\text{diag}\{\sigma_1, \dots, \sigma_p\} : 0]$, $p = \min(N, M)$, and $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0$. The i -th left and right singular vectors of matrix Σ

were vectors e_i and v_i , respectively. The values of σ_i were the singular values of the matrix Σ .

After the roller bearing signals were decomposed into PF s by the LMD method, all of the PF s were divided into two initial feature vector matrices X and Y

$$X = \begin{bmatrix} PF_1 \\ PF_2 \\ \vdots \\ PF_J \end{bmatrix}, \quad Y = \begin{bmatrix} PF_{J+1} \\ PF_{J+2} \\ \vdots \\ PF_n \end{bmatrix}; \quad (17)$$

where $J = n/2$ (when n is an even number) and $J = (n+1)/2$ (when n is an odd number). The characteristic of the roller bearing vibration signal $x(t)$ could be extracted from the initial feature vector matrices X and Y . In addition, the singular values that reflect the nature characteristics of the vector matrices X and Y as well as the roller bearing vibration signal can be used as fault feature vectors. After extracting fault feature vectors, the BSOA-SVM classifier could be employed to identify the working condition and fault pattern of roller bearing.

The flow chart of the roller bearing fault diagnosis method based on LMD-SVD and BSOA-SVM is shown in Fig. 3. It can be seen from the flowchart that the roller bearing fault diagnosis process included six main parts:

- (1) Select sampling frequency f_s under three conditions of the roller bearing (i.e., normal, outer-race fault, and inner-race fault).
- (2) Sample M times at this frequency. And the $3M$ signals were taken as samples that were divided into two subsets: the training samples and testing samples.

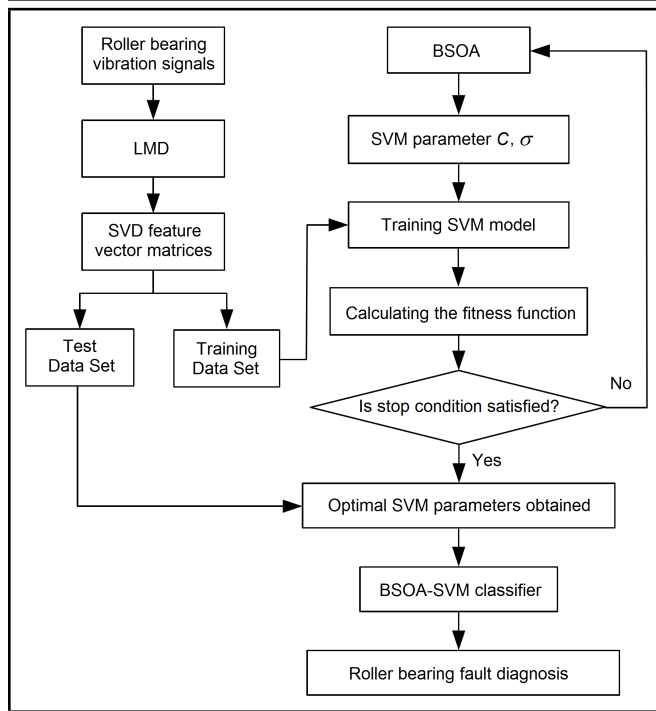


Figure 3. The flow chart of the roller bearing fault diagnosis based on BSOA-SVM and LMD-SVD.

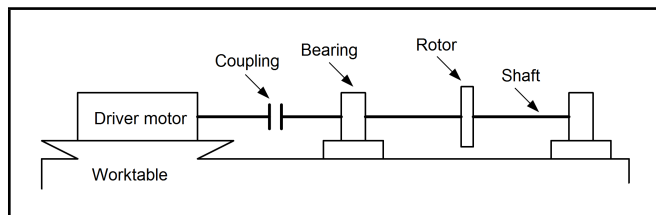


Figure 4. The test rig.

- (3) Decompose each sample signal by the LMD method.
- (4) Denote different amount of PF s in various signals by n_1, n_2, \dots, n_{3M} , and let $n = \max(n_1, n_2, \dots, n_{3M})$. If for some signals the amount of PF components, n_k , was less than n : $n_k < n$ ($k = 1, 2, \dots, 3M$), then it could be padded with zero to n components: $PF_1(t), PF_2(t), \dots, PF_n(t)$. That is $PF_i(t) = 0$ for $i = n_k + 1, n_k + 2, \dots, n$.
- (5) Create initial feature vector matrices X and Y to each roller bearing vibration signal corresponding to different working conditions according to Eq. (17). The singular values of initial feature vector matrices X and Y can be obtained as follows:

$$\sigma_{X,j} = [\sigma_{X,j}^1, \sigma_{X,j}^2, \dots, \sigma_{X,j}^J]; \quad (18)$$

$$\sigma_{Y,j} = [\sigma_{Y,j}^{J+1}, \sigma_{Y,j}^{J+2}, \dots, \sigma_{Y,j}^n]; \quad (19)$$

where $\sigma_{X,j}^1 \geq \sigma_{X,j}^2 \geq \dots \geq \sigma_{X,j}^J$, $\sigma_{Y,j}^{J+1} \geq \sigma_{Y,j}^{J+2} \geq \dots \geq \sigma_{Y,j}^n$, $j = 1, 2, 3$ denoted the normal condition, outer-race fault and inner-race fault, respectively.

- (6) Construct and train the BSOA-SVM classifiers with the training samples, which were obtained from singular values of the initial feature vector matrices. The fitness function was given by Eq. (13). The obtained values of C and σ were inputted into the BSOA-SVM classifier. Next, the

fault feature vectors of testing samples were inputted into the trained the BSOA-SVM classifier and then the working condition was given by the output of the BSOA-SVM classifiers.

5. ROLLER BEARING FAULT DIAGNOSIS APPLICATION

5.1. Data Acquisition

First, data acquisition was carried out on the small test rig, as shown in Fig. 4. This is a popular procedure for testing imbalances, misalignment, and various types of bearing faults. This test rig included a motor, a coupling, a rotor, and a shaft with two roller bearings. The roller bearings are the 6311 type. The shaft rotational frequency was 25 Hz, and the rotor's polar moment of inertia was 0.03 kgm^2 . By experimentation, the first three resonance frequencies of the roller bearing were determined to be 420 Hz, 732 Hz, and 1016 Hz, respectively. So the sampling frequency could be taken as 4096 Hz. The vibration signals were collected from the acceleration sensor, which had been mounted on a bearing seat at a steady frequency of shaft. Because the roller bearing was usually turned at a constant speed, the starting and stopping processes could be ignored. The fault was created by laser cutting slots that had a width and depth of 0.15 mm and 0.13 mm, respectively. Three conditions of roller bearings (normal, inner-race fault (IR fault), and outer-race fault (OR fault)) were tested. There were 45 vibration signals from the bearings in each condition that were obtained, from which 30 groups were selected at random as the training data.

Second, the data of the Case Western Reserve University Bearing Data Center Website (CWRUBDCW)¹⁶ was used with the permission of Professor K. A. Loparo. The test stand included a 2 hp Reliance Electric motor, a torque transducer/encoder, a dynamometer, and control electronics. The sample frequency was 485063 Hz and the motor speed was 1772 rpm. The deep groove ball bearing manufactured by SKF was used in this test stand. The drive end bearings are of the 6205-2RS JEM type. The test bearings of electro-discharge machining with fault diameters of 0.007 inches were selected. The roller bearings with the four conditions (normal, inner-race fault, outer-race fault, and ball fault) were tested, and 80 vibration signals from the bearings in each condition were obtained, from which 56 groups were selected at random as the training data.

5.2. Application

First, the roller bearing vibration signals are decomposed into a number of PF s by the LMD method. It's noticed by the analysis that the fault information of roller bearing was mainly included in the first five PF components. Therefore, the initial feature vector matrix X only was established by the first five PF components.

Second, the corresponding singular value σ_X of initial feature vector matrix X was extracted by applying the SVD. Then, these values were put into the BSOA-SVM classifier.

Third, in order to define the condition of roller bearing, SVM1 was first used to separate the normal condition from another condition by setting the normal condition as $y = +1$ and the other conditions as $y = -1$. Next, SVM2 was used to separate the outer-race fault from other condition by setting

Table 9. The identification results of the author’s data obtained by LMD-SVD-BSOA-SVM, LMD-SVD-GA-SVM, LMD-SVD-PSO-SVM, and LMD-SVD-CMAES-SVM methods.

Method	Training samples	Test samples	Optimal C	Optimal σ	Average cost time (s)	Average test error (%)
BSOA-SVM1	45	15	1173.746	0.434	4.347	1.667
PSO-SVM1	45	15	2365.376	0.098	9.733	1.905
GA-SVM1	45	15	26190.945	0.602	7.885	1.905
CMAES-SVM1	45	15	1422.560	32.000	7.330	2.381
BSOA-SVM2	20	10	30226.276	2.761	4.084	3.000
PSO-SVM2	20	10	16244.680	22.746	9.630	6.000
GA-SVM2	20	10	27911.125	6.460	7.689	7.667
CMAES-SVM2	20	10	31117.544	32.000	7.264	11.333

Table 10. The identification results of the author’s data based on the LMD-SVD and BSOA-SVM methods.

Test samples	Singular value of fault feature $\sigma X, x$					BSOA-SVM1 classifier	BSOA-SVM2 classifier	Identification results
(1) Normal	74.51	31.85	18.25	13.09	7.78	(+1)		Normal
(2) Normal	55.80	8.98	4.10	2.40	1.76	(+1)		Normal
(3) Normal	74.16	39.58	28.44	26.14	22.51	(+1)		Normal
(4) Normal	67.47	15.54	6.03	4.24	3.08	(+1)		Normal
(5) Normal	65.56	15.99	11.34	9.09	7.53	(+1)		Normal
(6) OR fault	139.94	91.39	50.12	27.04	25.94	(-1)	(+1)	OR fault
(7) OR fault	149.93	123.32	86.34	74.89	56.82	(-1)	(+1)	OR fault
(8) OR fault	203.29	68.58	45.97	39.68	33.83	(-1)	(+1)	OR fault
(9) OR fault	75.58	59.36	42.15	36.11	29.43	(-1)	(+1)	OR fault
(10) OR fault	77.70	42.70	19.70	17.10	7.00	(-1)	(+1)	OR fault
(11) IR fault	283.48	184.78	98.54	82.4	77.59	(-1)	(-1)	IR fault
(12) IR fault	199.92	140.14	70.77	54.52	29.69	(-1)	(-1)	IR fault
(13) IR fault	205.00	108.00	81.00	54.00	31.00	(-1)	(-1)	IR fault
(14) IR fault	196.49	158.08	73.12	52.14	24.01	(-1)	(-1)	IR fault
(15) IR fault	173.73	127.74	76.84	53.05	38.86	(-1)	(-1)	IR fault

Table 11. The identification results of CWRUBDCW data obtained by the LMD-SVD-BSOA-SVM, LMD-SVD-GA-SVM, LMD-SVD-PSO-SVM, and LMD-SVD-CMAES-SVM methods.

Method	Training samples	Test samples	Optimal C	Optimal σ	Average cost time (s)	Average test error (%)
BSOA-SVM1	56	24	1093.473	18.344	4.0155	1.250
PSO-SVM1	56	24	3434.477	12.445	9.3462	1.389
GA-SVM1	56	24	579.592	14.623	7.5844	1.389
CMAES-SVM1	56	24	25825.370	32.000	6.9994	1.528
BSOA-SVM2	42	18	24476.678	4.646	3.8741	0
PSO-SVM2	42	18	30047.771	1.602	9.1629	0
GA-SVM2	42	18	19252.667	8.882	7.3246	0
CMAES-SVM2	42	18	23781.571	32.000	6.8474	0
BSOA-SVM3	28	12	17539.863	7.110	3.7015	0.278
PSO-SVM3	28	12	4435.299	19.081	8.6774	0.278
GA-SVM3	28	12	0.010	0.663	7.5844	0.278
CMAES-SVM3	28	12	18325.339	32.000	6.4860	0.556

Table 12. The identification results of CWRUBDCW data based on the LMD-SVD and BSOA-SVM methods.

Test samples	Singular value of fault feature $\sigma X, x$				BSOA-SVM1 classifier	BSOA-SVMr classifier	BSOA-SVM3 classifier	Identification results
(1) IR fault	8.5153	2.9152	1.6010	0.6691	(+1)			IR fault
(2) IR fault	8.5559	3.2202	1.8215	0.6273	(+1)			IR fault
(3) IR fault	8.4412	3.1352	1.9311	0.5246	(+1)			IR fault
(4) IR fault	8.5901	2.8227	1.8962	0.7208	(+1)			IR fault
(5) IR fault	8.9155	3.0961	1.5089	0.9228	(+1)			IR fault
(6) IR fault	8.8645	3.3804	2.4951	1.9901	(+1)			IR fault
(7) OR fault	7.9041	2.6419	2.0240	0.7417	(-1)	(+1)		OR fault
(8) OR fault	6.6604	2.5858	1.7798	0.8708	(-1)	(+1)		OR fault
(9) OR fault	6.9439	2.4304	1.3590	0.8068	(-1)	(+1)		OR fault
(10) OR fault	7.8690	2.2100	1.8414	0.6937	(-1)	(+1)		OR fault
(11) OR fault	6.7675	2.4369	1.419	0.9283	(-1)	(+1)		OR fault
(12) OR fault	7.4424	3.4101	2.4794	1.4612	(-1)	(+1)		OR fault
(13) Ball fault	4.5775	1.2130	0.8730	0.3686	(-1)	(-1)	(+1)	Ball fault
(14) Ball fault	4.4927	0.8078	0.7951	0.4279	(-1)	(-1)	(+1)	Ball fault
(15) Ball fault	4.6239	0.9917	0.7678	0.6424	(-1)	(-1)	(+1)	Ball fault
(16) Ball fault	4.5416	1.1273	0.7654	0.3606	(-1)	(-1)	(+1)	Ball fault
(17) Ball fault	4.4950	0.8587	0.6016	0.5237	(-1)	(-1)	(+1)	Ball fault
(18) Ball fault	4.2611	0.8124	0.6932	0.4347	(-1)	(-1)	(+1)	Ball fault
(19) Normal	2.8399	1.7668	1.6482	1.3742	(-1)	(-1)	(-1)	Normal
(20) Normal	1.9474	1.3361	1.1629	1.0016	(-1)	(-1)	(-1)	Normal
(21) Normal	1.6105	1.1714	0.9231	0.8979	(-1)	(-1)	(-1)	Normal
(22) Normal	1.3062	0.8826	0.7353	0.6410	(-1)	(-1)	(-1)	Normal
(23) Normal	2.2264	1.1755	1.1589	1.0573	(-1)	(-1)	(-1)	Normal
(24) Normal	2.0903	1.3878	1.191	0.8797	(-1)	(-1)	(-1)	Normal

outer-race fault as $y = +1$ and the other condition as $y = -1$. Because the first data set had only got three conditions that needed to be identified, the rest was inner-race fault. The identification results of the same testing samples are shown in Tables 9 and 10.

Table 9 shows that accuracy of LMD-SVD-BSOA-SVM is higher than LMD-SVD-PSO-SVM, LMD-SVD-GA-SVM, and LMD-SVD-CMAES-SVM while the computational time was lower. Table 10 shows detail identification result of the author's data based on LMD-SVD-BSOA-SVM.

The CWRUBDCW data set included four conditions of roller bearing so three SVM classifiers were used. The identification results of the same testing samples are shown in Tables 11 and 12.

It can be seen from Table 11 that the BSOA-SVM classifier gave more accurate results with less computational time than that by other methods. With IR fault, the BSOA-SVM method obtained the best identification result compared with other methods. With the OR fault, the classification success rate of the four methods was 100%. With ball fault, the identification result of BSOA-SVM, PSO-SVM, and GA-SVM was higher than that of CMAES-SVM. Table 12 shows the identification results of the CWRUBDCW data based on the LMD-SVD and BSOA-SVM methods.

6. CONCLUSIONS

In this paper, an optimal algorithm for SVM parameter based on BSOA is proposed. The experimental problems are solved to demonstrate the effectiveness of BSOA-SVM. The testing results of some real-world benchmark data sets show that the BSOA-SVM classifier has a high accuracy with low computational cost time. These results thus prove that the BSOA-SVM classifier gave more accurate results in a shorter time compared to the GA-SVM, PSO-SVM, and CMAES-SVM methods. Furthermore, the BSOA-SVM method is applied to diagnose the roller bearing fault by combining it with the LMD-SVD method. SVD is applied to extract the singular values of matrices from the initial feature vector matrices of *PFs*. The results indicate that a combination of the BSOA-SVM classifier and the LMD-SVD method can effectively reduce the test error and cost time.

ACKNOWLEDGEMENTS

This research was funded by the Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 107.99-2014.11.

REFERENCES

- ¹ Imbault, F. and Lebart, K. A stochastic optimization approach for parameter tuning of support vector machines, *International Conference on Pattern Recognition*, (2004). <http://dx.doi.org/10.1109/icpr.2004.1333843>
- ² Friedrichs, F. and Igel, C. Evolutionary tuning of multiple SVM parameters, *Neurocomputing*, **64**, 107–117, (2005). <http://dx.doi.org/10.1016/j.neucom.2004.11.022>
- ³ Abe, S. *Advances in pattern recognition*, Springer-Verlag, London, (2005). <http://dx.doi.org/10.1007/1-84628-219-5>
- ⁴ Keerthi, S. S. Efficient tuning of SVM hyperparameters using radius/margin bound and iterative algorithms, *IEEE Transactions on Neural Networks*, **13**, 1225–1229, (2002). <http://dx.doi.org/10.1109/tnn.2002.1031955>
- ⁵ Chapelle, O., Vapnik, V., Bousquet, O., and Mukherjee, S. Choosing multiple parameters for support vector machines, *Machine Learning*, **46**, 131–159, (2002). <http://dx.doi.org/10.1023/A:1012450327387>
- ⁶ Kennedy, J. and Eberhart, R. C. Particle swarm optimization, *Proceedings of the 1995 IEEE International Conference on Neural Networks*, **4**, 1942–1948, (1995). <http://dx.doi.org/10.1109/ICNN.1995.488968>
- ⁷ Zhang, X., Chen, X. and He, Z., An ACO-based algorithm for parameter optimization of support vector machines, *Expert Systems with Applications*, **37**, 6618–6628, (2010). <http://dx.doi.org/10.1016/j.eswa.2010.03.067>
- ⁸ Li, H., and Yan, X. A new optimization algorithm for weight optimization, *Proceedings of the Advances in Computation and Intelligence: Third International Symposium, ISICA 2008*, Wuhan, China, 723–730, (2008). http://dx.doi.org/10.1007/978-3-540-92137-0_79
- ⁹ Eberhart, R. and Shi, Y. Comparison between genetic algorithms and particle swarm optimization, *Evolutionary Programming VII*, V. W. Porto, N. Saravanan, D. Waagen, and A. E. Eiben (Eds.), Springer Berlin Heidelberg, 611–616, (1998). <http://dx.doi.org/10.1007/BFb0040812>
- ¹⁰ Civicioglu, P. Backtracking search optimization algorithm for numerical optimization problems, *Applied Mathematics and Computation*, **219**, 8121–8144, (2013). <http://dx.doi.org/10.1016/j.amc.2013.02.017>
- ¹¹ Vapnik, V. N. *The nature of statistical learning theory*, Springer, New York, (1995). <http://dx.doi.org/10.1007/978-1-4757-2440-0>
- ¹² Ebbesen, S., Kiwiz, P., and Guzzella, L. A generic particle swarm optimization Matlab function, *Proceedings of the American Control Conference (ACC)*, 1519–1524, (2012). <http://dx.doi.org/10.1109/acc.2012.6314697>
- ¹³ Wang, Y., Cai, Z., and Zhang, Q. Differential evolution with composite trial vector generation strategies and control parameters, *IEEE Transactions on Evolutionary Computation*, **15**, 55–66, (2011). <http://dx.doi.org/10.1109/tevc.2010.2087271>
- ¹⁴ Lin, S. W., Ying, K. C., Chen, S. C., and Lee, Z. J. Particle swarm optimization for parameter determination and feature selection of support vector machines, *Expert Systems with Applications*, **35**, 1817–1824, (2008). <http://dx.doi.org/10.1016/j.eswa.2007.08.088>
- ¹⁵ Smith, J. S. The local mean decomposition and its application to EEG perception data, *Journal of The Royal Society Interface*, **2**, 443–454, (2005). <http://dx.doi.org/10.1098/rsif.2005.0058>
- ¹⁶ Loparo, K. A. Bearings vibration data set-case Western Reserve University, (2003), Retrieved from <http://csegroups.case.edu/bearingdatacenter/pages/welcome-case-western-reserve-university-bearing-data-center-website> (Accessed September 14, 2014).