The International Conference on Electronics & Oil: From Theory to Applications
March 05-06, 2013, Ouargla, Algeria

Autoregressive Modeling preprocessing with SVM Optimization based on BFA for Bearing Fault Diagnosis

T. Thelaidjia (1) and S. Chenikher (2)

(1) *Department of Electrical Engineering, Tebessa University, Tebessa, 12002, Algeria*
t.tladjia@mail.univ-tebessa.dz
(2) *Department of Electrical Engineering, Tebessa University, Tebessa, 12002, Algeria*
s.chenikher@mail.univ-tebessa.dz.

Abstract— **As an effective tool in pattern recognition and machine learning, support vector machine (SVM) has been adopted abroad. In developing a successful SVM classifier, extracting feature is very important. This paper proposes the application of Autoregressive Modeling to SVM for feature extraction. According to the fact that parameter selection of support vector machine(SVM) for fault diagnosis is difficult, a new method based on bacterial foraging algorithm(BFA) for support vector machine parameter optimization was proposed , then the faster optimization of the parameters "c" and kernel parameter "σ" was performed. The results have shown feasibility and effectiveness of the proposed approach.**

Key-Words— **Bacterial Foraging Algorithm, Autoregressive Modeling, Support Vector Machine, Wavelet Packet, Fault Diagnosis, Roller Bearing, Machine Learning, Time Series Analysis, Vibration Measurement, Rotating Machines.**

I. INTRODUCTION

BEARINGS are frequently applied components in the vast majority of rotating machines. Their running quality influences the working performance of equipment. Statistically, 30% of rotational mechanical equipments malfunction is caused by the faults in bearings [1]. Therefore, many important researches had been done in the advanced field of bearing fault diagnosis [2]–[6]. Using the vibration signals of rolling bearings and components to monitor and diagnose their working state, is the common used method in the study of bearing fault diagnosis [7], [1] and [5]. Support Vector Machine (SVM) is a new machine learning method which was introduced by Vapnik on the foundation of statistical learning theory (SLT). However, since the middle of 1990s, the algorithms used for SVM started emerging with greater availability of computing power [8], [9]. The main difference between the known domain of artificial neural network (ANN) and SVM is in the principle of risk minimization (RM) [4], [5]. In case of SVM, structural risk minimization (SRM) principle is used to minimize an upper bound on the expected risk whereas in ANN, traditional empirical risk minimization (ERM) is used to minimize the error on the training data. The difference in RM leads to better generalization performance for SVM than ANN. According to the literature, SVM has been successfully applied to many applications, such as pattern identification, regression analysis, function approximating, etc [10]–[13]. The results give the evidence that the technique is not only quite satisfying from a theoretical point of view, but also can lead to high performance in practical applications.

Finding out good features is an important phase in distinguishing the different mechanical failure. Many approaches have been developed in recent years, such as Parts Principal Component Analysis, Empirical Mode Decomposition and so on [14] and [15] . As an interesting example, Wavelet Packet analysis has been utilized for impulse mechanical failure classification [1].

Parameter optimization is the key to perform SVM. At present, the widely used methods of parameter optimization for SVM are network search method, K-order cross-validation method, Leave-one-out method, etc. These algorithms have the disadvantage of huge amount of computation, and the calculated parameters are not always the best. In recent years, a series of intelligent bionic algorithms are proposed based on the biological behavior study in the natural, such as genetic algorithm (GA) and particle swarm optimization (PSO) [16]–[18].

In 2002, Kevin M. Passino has proposed a new bionic optimization method based on phagocytosis behavior study of Escherichia coli :bacterial foraging algorithm(BFA) [19],[20].This intelligent algorithm is becoming a hot topic because of its advantages of swarm intelligence, parallel search, easy to jump out of local minima, rapid convergence velocity. D. H. Kim, C.H. Cho have applied it to Neural Network Fuzzy Learning [21], C.Ying, S.Zibo, M.Hua, W.Qinghua have applied it to image compression combining with BP network [22], W. Xuesong, C. Yuhu, H. Minglin have applied it to predictive control and obtained good result [23], S. Mishra proposed a hybrid least square-fuzzy bacterial foraging strategy for harmonic estimation [24]. In this paper, an effective method is proposed by applying BFA to determine SVM parameters, and carry on fault feature recognition of bearing, which can improve the accuracy of fault diagnosis effectively.

II. FEATURE EXTRACTION

The fault diagnosis is essentially a problem of pattern recognition, of which, an important step is feature extraction. In this study, two types of feature extraction are applied :Wavelet packet transform and Autoregressive modeling.

A. Wavelet packet algorithm

The step of feature extraction based on three layer wavelet packet is given as follows:
Firstly, The vibration signal x(t) was decomposed by a mother wavelet, the signal features in eight frequency bands from low to high were extracted in the third layer.
Secondly, The signal in each frequency band is extracted and the wavelet packet decomposition coefficient was reconstructed. D_{30} presents the reconstructed signal of d_{30}, D_{31} presents the reconstructed signal of d_{31}, and so on. The composed signal is defined as :

$$D = D_{30} + D_{31} + D_{32} + D_{33} + D_{34} + D_{35} + D_{36} + D_{37} \quad (1)$$

Finaly, The signal energy of each frequency band is calculated as:

$$E_{3j} = \sum_{k=1}^{n} |d_{jk}|^2 \quad (2)$$

where d_{jk} : is the wavelet transform coefficients.

Normalized, Let, $T = \sum_{j=0}^{7} E_{3j}$.
Eigenvector E was constructed based on each frequency band energy:

$$E = \left[\frac{E_{30}}{T}, \frac{E_{31}}{T}, \frac{E_{32}}{T}, \frac{E_{33}}{T}, \frac{E_{34}}{T}, \frac{E_{35}}{T}, \frac{E_{36}}{T}, \frac{E_{37}}{T}\right] \quad (3)$$

B. LEAST-SQUARES METHOD FOR AR PARAMETER ESTIMATION

In this section, we derive a method of AR estimator, witch based on a least-squares (LS) minimization criterion using the time-domain relation A(z)y(t) = e(t) [25], [26]. Let x(n) be an AR process of order p. Then x(n) satisfies:

$$e(n) = x(n) + \sum_{K=1}^{p} \alpha_k x(n-k) = x(n) + \hat{x}(n) \quad (4)$$

We interpret $\hat{x}(n)$ as a linear prediction of x(n). from the n previous samples x(n − 1), ..., x(n − p), and we interpret e(n) as the corresponding prediction error.
The vector $\alpha = [\alpha_1, \dots, \alpha_p]'$ that minimizes the prediction error variance $\rho = E\{|e(n)|^2\}$ is the AR coefficient vector,we have:

$$\rho = E\{|e(n)|^2\} = E\{|\hat{x}(n) - x(n)|^2\} = r_{xx}(0) + \alpha^H r + r^H \alpha + \alpha^H R \alpha \quad (5)$$

where α, R and r are defined by:

$$\alpha = \alpha_1, \dots, \alpha_p]' \quad (6)$$

$$r = (r_{xx}(1), r_{xx}(2), \dots, r_{xx}(p))' \quad (7)$$

$$R = \begin{pmatrix} r_{xx}(0) & r_{xx}(-1) & \dots & r_{xx}(-p+1) \\ r_{xx}(1) & r_{xx}(0) & \cdots & r_{xx}(-p+2) \\ \vdots & \vdots & \vdots & \vdots \\ r_{xx}(p-1) & r_{xx}(p-2) & \cdots & r_{xx}(0) \end{pmatrix} \quad (8)$$

The vector α that minimizes (5) is given by:

$$\alpha = -R^{-1} r \quad (9)$$

With corresponding minimum prediction error:

$$\rho = r_{xx}(0) - r^H R^{-1} r \quad (10)$$

The least-squares AR estimation method is based on a finite sample approximate solution of the above minimization problem. Given a finite set of measurements $\{(x(n))\}_{n=1}^{N}$ we approximate the minimization of $\rho = E\{|e(n)|^2\}$ by the finite sample cost function:

$$f(\alpha) = \sum_{n=n_0}^{n_1} |e(n)|^2 = \sum_{n=n_0}^{n_1} \left|x(n) + \sum_{k=1}^{p} \alpha_k x(n-k)\right|^2 \quad (11)$$
$$= \|h + X\alpha\|^2 \quad (12)$$

Such that:

$$h = \begin{pmatrix} x(n_0) \\ x(n_0+1) \\ \vdots \\ \vdots \\ x(n_1) \end{pmatrix};$$

$$X = \begin{pmatrix} x(n_0-1) & \cdots & x(n_0-p) \\ x(n_0) & \cdots & x(n_0+1-p) \\ x(n_1-1) & \cdots & x(n_1-p) \end{pmatrix};$$

where we assume x(n) = 0 for n < 1 and n > N, The vector α that minimizes f (α) is given by:

$$\alpha = -(X * X)^{-1}(X * h) \quad (13)$$

where, as seen from (12) the definitions of X and h depend on the choice of (n_0, n_1), when $n_0 = p + 1$ and $n_1 = N$ this choice is often named **the covariance method.**

III. SUPPORT VECTOR MACHINE

The support vector machine (SVM) is a supervised learning method that generates input-output mapping functions from a set of labeled training data. For classification, nonlinear kernel functions are often used to transform input data to a high-dimensional feature space in which the input data become more separable compared to the original input space [27].

Support vector machine (SVM) based on statical learning theory is proposed according to optimal hyperplane in the

case of linear separable [1].
If the hyperplane separate all samples correctly, it must satisfy the following condition [6]:

$$y_k(\langle w;x\rangle - \lambda_0) \geq +1, \forall k \in \{1,\dots,n\} \quad (14)$$

In order to find the optimal hyperplane, we need to minimize the following functional [6]:

$$\varphi(\omega) = \frac{1}{2}\|\omega\|^2 \quad (15)$$

Solution of the optimal problem is given by the saddles of Lagrange function as below:

$$L(\omega,\lambda_0,\alpha) = \frac{\|\omega\|^2}{2} - \sum_{k=1}^{n}\alpha_k[y_k(\langle w;x\rangle - \lambda_0) - 1] \quad (16)$$

where $\alpha = (\alpha_1,\dots,\alpha_n)$ is the Lagrange coefficient; $\alpha_i \geq 0, \forall i$.

The original problem can be transferred to the dual problem as below:

$$W(\alpha) = \sum_{k=1}^{n}\alpha_k - \frac{1}{2}\sum_{k,k'=1}^{n}\alpha_k\,\alpha_{k'}\,y_k y_{k'}\langle x_k;x_{k'}\rangle \quad (17)$$

subject to:

$$\sum_{k=1}^{n}\alpha_k y_k = 0 \; et \; \alpha_k \geq 0$$

If α^* is the optimal solution, then:

$$\langle w^*;x\rangle = \sum_{k=1}^{n}\alpha_k^* y_k\langle x_k;x\rangle \quad (18)$$

It means that the weight coefficients of the optimal hyper-plane is the linear combination of the training sample vector. According to the Kuhn −Tucker condition, the solution of optimal problem must satisfy:

$$\alpha_*^k[y_k(\langle w^*;x\rangle - \lambda_0^*) - 1] = 0 \quad (19)$$

were λ_0^* is given by:

$$\lambda_0^* = \frac{1}{N_{sv}}\sum_{s=1}^{N_{sv}}(y_s - x_s^T\omega^*) \quad (20)$$

where N_{sv}: nomber of support vectors.
After solving the above problem, we can get the optimal classification function as below:

$$D(x) = sgn\left[\sum_{k\in N_{sv}}\alpha^* y_k x_i^T x - \lambda_0^*\right] \quad (21)$$

The nonseparable problem can be solved by soft-margin SVM [9], [28], [29].
If we used the inner $\kappa(x_k,x)$ substitute for the inner of the optimal hyperplane, the original feature space is mapped to new feature space [30]. And the optimal function can be formulated as below:

$$W(\alpha) = \sum_{k=1}^{n}\alpha_k - \frac{1}{2}\sum_{k,k'=1}^{n}\alpha_k\,\alpha_{k'}\,y_k y_{k'}\kappa(x_k,x_{k'})$$

subject to:

$$\sum_{k=1}^{n}\alpha_k y_k = 0 \; and \; 0 \leq \alpha_k \leq c$$

The corresponding decision function is written as below:

$$D(x) = sgn\left[\sum_{k\in N_{sv}}\alpha^* y_k\kappa(x_i^T x) - \lambda_0^*\right] \quad (22)$$

here, $\kappa(x_k,x)$ is called kernel function.
Usually, the kernel function can be expressed as below [9], [10], [31].

Polynomial:

$$\kappa(x_1,x_2) = (1 + \langle x_1,x_2\rangle)^q \quad (23)$$

where "q" is the degree of the polynomial.

Radial basis function (RBF):

$$\kappa(x_1,x_2) = \exp\left(-\|x_1-x_2\|^2/2\delta^2\right) \quad (24)$$

where δ^2 is the variance of the Gaussian function.

Sigmoid:

$$\kappa(x_1,x_2) = \tanh(\alpha_0\langle x_1;x_2\rangle + \beta_0) \quad (25)$$

where α_0 and β_0 are the parameters of kernel function. The classification performances of SVM are affected by three techniques, i.e., the selecting of the kernel, the choosing of the kernel parameters, and the choosing of the regularization parameter "c" [6].

Most of cases in practical are multi-classed, such as in the rolling bearing classifying, it can be sorted into normal, outer race fault and inner race fault, etc. So, we have to design an approach to expend the application of SVM to a multi-classifying field because the SVM can deal with only two classes. The different combination principles constitute different classifying algorithm [10], [32], [33]. We employ the one-against-the-rest method to compose a multi-fault classifier. Since the SVM generalization performance heavily depends on the right setting of "c" and "σ", these two parameters need to be set properly by the user. According to the experience from numerical experiments [34], [35], "c" and "σ" exhibit a (strong) interaction. As a consequence, they should be optimized simultaneously, rather than separately.

IV. THE BACTERIAL FORAGING ALGORITHM

In the bacterial foraging algorithm, the optimization problem corresponds to the status of bacterium in the search space, i.e. the optimum value of fitness function. The original bacterial foraging optimization system consists of three principal mechanisms, namely, chemotaxis, reproduction, and elimination-dispersal.

Chemotaxis. In this mechanism, it simulates the movements of a bacterium, including "tumble" and "swim", as shown in Figure(1). Bacterium moves in any direction step by step is defined as tumble .After a Bacterium completes one tumble, if the fitness become better, it will continue to move a number of steps along the same direction. If the fitness does not become better, or complete max step number, it is finished. This process is named swim. Supposed $P^i(j,k,l)$ represents the bacterium's position at

j^{th} chemotactic, k^{th} reproductive, and l^{th} elimination dispersal step, the next position is defined as follows:

$$P^i(j+1,k,l) = P^i(j,k,l) + C(i)V(j) \quad (26)$$

Where, C(i) is the step length vector of the i^{th} bacterium, V (j) is the direction vector of the j^{th} chemotactic where is randomly generated.

Reproduction. In this mechanism, it simulates the selection process of bacterium survival for the fittest individuals.

When a bacterium's life will be over, it reaches a critical chemotactic step number, and the bacterium will reproduction. The reproduction follows the principle "survival of the fittest". Using the cumulative fitness of bacterium in chemotaxis as the standard, the lower half of the bacterium die, the higher half of the bacterium split into two sub bacterium.

The sub-bacterium inherits the parent bacterium's biological properties, such as the position, the step length and the fitness. In this process, the population size is maintained.

Supposed the population of bacterium is N , $F^i(j,k,l)$ is the fitness of the ith bacterium, sorted by descending, then the front N bacterium is replaced the behind N bacterium.

Elimination-dispersal. The chemotaxis provides a basis for local search, and the reproduction process speeds up the convergence simulated by the classical BFA. When a bacterium reproduction several times, it will be dispelled by certain probability to any position in search space. If a bacterium is chosen, it will be killed, and then a new bacterium is generated adding to the group.

The process of bacteria foraging algorithm for solving the optimization problem includes: (1) Encode the solutions for the problem; (2) Design evaluation function; (3) Generate initial solution population; (4) Optimize parameters by using the interaction between groups.

Supposed reN, cN , and edN denote maximum reproduction, maximum chemotaxis, and maximum elimination-dispersal respectively. Bacterial foraging algorithm contains following steps:

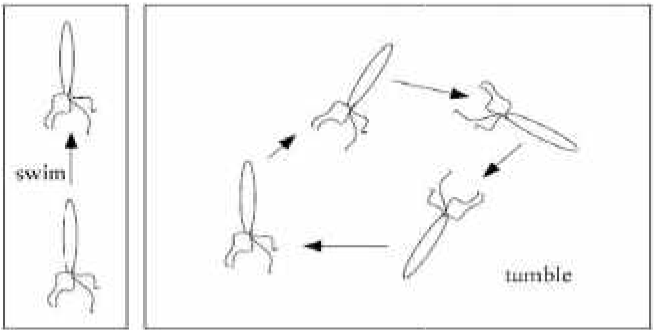

Fig. 1.Swim and tumble of a bacterium

- **Step1:** Initialize groups by using the evaluation function to assess the merits of the various individuals. Initialization l = 0, k = 0, j = 0,
- **Step2:** Elimination-dispersal loop, l = l + 1,
- **Step3:** Reproduction loop, k=k+1,
- **Step4:** Chemotaxis loop, j = j + 1,
- **Step5:** Executive chemotaxis operation ,
- **Step6:** If j< cN , go to step 4,
- **Step7:** Executive reproduction operation,
- **Step8:** If k< reN , go to step 3,
- **Step9:** Executive elimination-dispersal operation,
- **Step10:** If l< edN , go to step2, else finish.

V. EXPERIMENTAL SETUP AND VIBRATION DATA

In this section, an experimental dataset of a typical ballbearing is considered. These data are recorded by Tabaszewski and Cempel (1998). The ball bearing type that has been tested is 6402 (steel cage). The shaft rotational frequency and the sampling frequency of the analyzer for recording the acceleration (m/s^2) signal of the ball bearing are 24.5625 and 16,384 Hz, respectively. For the data acquistion, the B&K analyser is used [36].

VI. SVM PARAMETER O PTIMISTION BASED ON BFA

The performance of SVM is not only related to the type of kernel function, but also related to penalty factor and kernel parameters. The kernel function and parameters both determine the mapping of the original space to high dimensional space. The value of penalty factor can adjust the error and complexity. Taken the Gaussian kernel function as an example, the steps to select the penalty factor "c" and kernel parameter "σ" based on bacteria foraging algorithm are as follow:

- **Step1**. Read sample data S from file, then divide it into two groups S1 , S2 . Supposed the size of bacterium as N, and randomly generated N groups set of {c, σ} to initial location of the position of bacterium
- **Step2**. Design fitness evaluation function fitness = f (c,σ) . According to the value of "c" and σ , train the SVM model with group S1 , then consider opposite value of the testing accuracy with group S2 as the fitness.
- **Step3**. Execute the loop of chemotaxis, reproduction and elimination-dispersal.
- **Step4**. Encode, define the position of bacterium with best fitness as the optimal c^*or δ^*.

VII. ANALYSIS OF EXPERIMENTATION RESULTS

The database is composed from five different classes C1 to C5, including a normal bearing and four faults of roller bearing (outer race completely broken fault bearings, broken cage with one loose element fault bearings, damaged cage with four loose elements fault bearings and badly warned ball-bearings). Once the features are extracted by the AR method, the total database of bearing faults were divided in two sets: one for training (containing 60,71% of the samples), and the other for test (containing 39,29% of the samples).

Table I gives the classification results for this bearing fault classification problem based on :

- the proposed method (SVM-BFA based on Autoregressive Modeling feature extraction).
- SVM-BFA based on wavelet packet feature extraction, where Discrete Wavelet Packet transform was used to decompose the time signals into eight packets at level 3 via Daubechies-8

TABLE I.
CLASSIFICATION RESULTS IN VALIDATION AND TEST

SVM+	Kernel fonction	Opimal "c"	Opimal "σ"	Validation Rate%	Test Rate%
AR Modeling	Gaussian	85.8452	1.6709	97.06	100
Wavelet Packet	Polynomial	1465.2789	2.2246	85.29	63.64

As shown in table I:

- the classification accuracy is poor either in validation or test when we combine SVM with wavelet packet.
- the best classification result of bearing fault in the validation set: 97.06%, and in the test set: 100%, are obtained by using the proposed method of SVM-BFA based on Autoregressive Modeling feature extraction.

In order to select the optimal values of the parameter p (order of autoregressive modeling), for bearing fault classification, a series of experiments had been carried out by varying the values of this parameter. The important variation range of "p" is : **p ∈ [13, 20].**

The classification results and the the optimal values of "c" and "σ" obtained for different values of p are shown in table II.

TABLE II
SVM-BFA CLASSIFICATION RESULTS OF BEARING FAULT USING DIFFERENT VALUES OF "p".

Degree of AR modeling (p)	Optimal "c"	Optimal "σ"	Rate of validation %	Rate of test %
16	**85.8452**	**1.6709**	**97.06**	**100**
13	167.9000	-1.2894	91,18	100
14	287.9000	-0.7894	97.06	95.45
15	227.9513	1.5921	91,18	100
17	44.3725	2.2862	94.12	100
18	32.0312	3.0706	97.06	100
19	584.2063	-4.8548	94.12	100
20	253.8435	4.7371	94.12	100

As the results in Table II, The maximum correct classification result of bearing fault is obtained for **p = 16, c = 85.8452 and σ = 1.6709.**

A confusion matrix of dimension 5×5 is constructed to show the bearing fault classification performance . The diagonal elements represent the correctly classified bearing fault. The off-diagonal elements represent the misclassification of bearing faults.

As shown in table III, the best classification result of bearing fault in the validation set 97,06% and in the test set 100% is obtained by using the proposed method of SVM-BFA based

TABLE III.
CLASSIFICATION RESULT OF BEARING FAULT IN VALIDATION AND TEST.

SVM+	kernel fonction	validation result	test result	Validation Rate%	test Rate%
AR Modeling	Gaussian	70000 06000 00700 00160 00007	40000 03000 00500 00050 00005	97.06	100
Wavelet Packet	Polynomial	06000 00601 00232 00007	03000 00212 00005 00005	85.29	63.64

on Autoregressive Modeling feature extraction, where only 1 damaged cage with four loose elements fault bearing was judged to one loose elements fault bearing by error.

These results clearly show the high percentage of correct classification reached for the validation set and test set, which clearly shows the good generalization capacity of SVM-BFA based on Autoregressive Modeling for fault diagnosis of roller bearing.

VIII. CONCLUSION

For the Bearing faults diagnosis, input feature subset selection and the SVM parameters setting are crucial problems.

This paper presents a new method AR-SVM-BFA for bearing fault classification, AR modeling is utilized for feature extraction.

After feature extraction from bearing fault vibration signal, a Bacterial Foraging Algorithm is employed to simultaneously optimize the SVM kernel function parameter and the penalty parameter.

The proposed method can overcome the inefficiency for selecting reasonable parameters according to the experience in the traditional fault diagnosis. Compared with other methods, AR-SVM-BFA is simpler and easier to realize.

The combined AR modeling and SVM-BFA based technique is tested for bearing faults and provides satisfactory results.

REFERENCES

[1] M. Li and P. Zhao, "The application of wavelet packet and svm in

rolling bearing fault diagnosis," IEEE International Conference on Mechatronics and Automation, August 2008.

[2] L. Frosini, E. Bassi, A. Fazzi and C. Gazzaniga, "Use of the stator current for condition monitoring of bearings in induction motors," International Conference on Electrical Machines, September 2008.

[3] L. A. Escobar-Moreira, "An approach to ultrasonic fault machinery monitoring by using the wigner-ville and choi-williams distributions,"International Conference on Electrical Machines, September 2008.

[4] P. K. Kankar, C. S. Satish and S. P. Harsha, "Fault diagnosis of ball bearings using machine learning methods," ELSEVIER, 2011.

[5] L. Shuang and L. Meng, "Bearing fault diagnosis based on pca and svm," IEEE International Conference on Mechatronics and Automation, August 2007.

[6] G. Xian, "Mechanical failure classification for spherical roller bearing of hydraulic injection moulding machine using dwt-svm," ELSEVIER, August 2010.

[7] G. D. Vescovo, M. Paschero, A. Rizzi,R. Di Salvo and F.M.F. Mascioli, "Multi-fault diagnosis of rolling-element bearings in electric machines," International Conference on Electrical Machines, September 2010.

[8] V. N. Vapnik, The Nature of Statistical Learning Theory. Springer-Verlag, 1995.

[9] V. N. Vapnik, Statistical Learning Theory. Springer, 1998.

[10] C. Cortes and V. N. Vapnik, Support-vector networks.Machine Learning, 1995.

[11] C. J. Burges, "A tutorial on support vector machines for pattern recognition," Data Mining and Knowledge Discovery, vol. 2, pp. 121–167, 1998.

[12] S. R. Gunn, "Support vector machines for classification and regression," pp. 1–28, 1998.

[13] J. Christopher and C. Burges, "A tutorial on support vector machines for pattern recognition," Kluwer Academic Publishers, Bell Lab, Lucent Technologies, Boston, pp. 1–43, 1998.

[14] K. Chen and P. Fu, "The ppca a new method to select features in process of rolling bearing fault diagnosing," International Conference on Mechatronics and Automation, 2009.

[15] Q. Mian-hao and W. Zi-ying, "Fault diagnosis of bearing based on empirical mode decomposition and decision directed acyclic graph support vector machine," International Conference on Computational Intelligence and Natural Computing, 2009.

[16] R. Yuan and B. Guangchen, "Determination of optimal svm parameters by using ga/pso," Journal of computer, vol. 5, no. 8, August 2010.

[17] H. Cheng-Lung and D. Jian-Fan, "A distributed pso-svm hybrid system with feature selection and parameter optimization," ELSEVIER, 2008.

[18] X. Yun-Jie and X. Shu-Dong, "A new and effective method of bearing fault diagnosis using wavelet packet transform combined with support vector machine," Journal of computers, vol. 6, no. 11, November 2011.

[19] K. M. Passino, "Biomimicry of bacterial foraging for distributed optimization and control," IEEE Control Systems Magazine, vol. 22, pp.52–67, 2002.

[20] T. Datta and I. S. Misra, "Improved adaptive bacteria foraging algorithm in optimization of antenna array for faster convergence," Progress In Electromagnetics Research C, vol. 1, pp. 143– 157, 2008.

[21] D. H. Kim and C. H. Cho, "Bacterial foraging based neural network fuzzy learning," IICAI, pp. 2030–2036, 2005.

[22] C. Ying, S. Zibo, M. Hua and W. Qinghua, "An application of bacterial foraging algorithm in image compression," Journal of Shenzhen University Science and Engineering, vol. 25, pp. 153– 157, 2008.

[23] W. Xuesong, C. Yuhu and H. Minglin, "Estimation of distribution algorithm based on bacterialforaging and its application in predictive control," Acta Electronica Sinica, vol. 38, pp. 333–339, 2010.

[24] S. Mishra, "A hybrid least square-fuzzy bacteria foraging strategy forharmonic estimation," IEEE Transactions on Evolutionary Computation, vol. 9, pp. 61–73, 2005.

[25] P. Stoica and R. Moses, Spectral analysis of signal.Pearson Education, 2005.

[26] A. Gattal, S. Chenikher, K. M. Pekpe and J. P. Cassar, "Bearing faults diagnosis using artificial network with dimensionality reduction by pca," 1st International Conference on Electrical Engineering CIGET09, 2009.

[27] L. Wang, Support Vector Machines: Theory and Applications. Springer-Verlag Berlin Heidelberg, 2005.

[28] P. Clarkson and P. J. Moreno, "On the use of support vector machines for phonetic classification," IEEE Proceedings of the international conference acoustics, speech and signal processes, vol. 2, pp. 585–588, March 1999.

[29] P. Ding, Z. Chen, Y. Liu and B. Xu, "Asymmetrical support vector machines and applications in speech processing," IEEE Proceedings of the international conference acoustics, speech and signal processes, vol. 1, pp. 73–76, May 2002.

[30] C. Wang and Y. Song, "Support vector machine for mechanical faults diagnosis," IEEE International Conference on Measuring Technology and Mechatronics Automation, 2010.

[31] B. Scholkopf, A. Smola, R. C. Williamson and P. L. Bartlett, "New support vector algorithms," Neural Computation, vol. 12, pp. 1207–1245, 2000.

[32] M. XiaoXiao and H. XiYue, "2ptmc classification algorithm based on support vector machines and its application to fault diagnosis," Control and Decision, vol. 18(3), pp. 272–276, 2003.

[33] J. Weston and C. Watkins, "Multi-class support vector machines," European Symposium on Artificial Neural Networks, pp. 219–224, April 1999.

[34] X. F. Yuan and Y. N. Wang, "Parameter selection of svm for function approximation based on chaos optimization," Journal of Systems Engineering and Electronics, vol. 19, pp. 191–197, 2008.

[35] B. Ustun, W. J. Melssen , M. Oudenhuijzen and L. M. C. Buydens, "Determination of optimal support vector regression parameters by genetic algorithms and simplex optimization," Analytica Chimica Acta, vol. 54, pp. 292–305, 2005.

[36] M. Tabaszewski and C. Cempel, "Ball bearing dataset, structural integrity and damage assessment. network (sidanet.)." available at: www.sidanet.org, 1998. (consulted in July 2006).