International Journal of Computer Science, Communication & Information Technology (CSCIT)
Vol.7 pp. 37-42

Spectrum Analysis and Statistical Parameters Based Two Stage Support Vector Machine for Fault Severity Diagnosis of Roller Bearing

T. Thelaidjia, S. Chenikher, A. Moussaoui

***Abstract*—bearings are frequently applied components in the vast majority of rotating machines. The breakdown of one of their constitutive parts can stop processes and cause losses in terms of time and money. In this paper a two stage Support Vector Machine (SVM) approach is proposed for severity fault diagnosis in bearing , The method consists of two stages. Firstly, Frequency domain featues from the demodulated signal followed by support vector machine (SVM) is performed to detect different faults. At the second stage, statistical parameters are used as un input of four SVMs to classify faults severity for each class. To improve the classification accuracy for bearing fault prediction, Particle Swarm Optimisation (PSO) is employed to simultaneously optimize the SVMs parameter's. the proposed condition monitoring scheme is verified and demonstrated by the testing results.**

***Index Terms*—Condition monitoring; Fault Diagnosis; Machine learning; Particle Swarm Optimisation; Roller Bearing; Rotating machines; Spectral analysis; Statistical parameters; Support Vector Machine; Vibration measurement.**

I. INTRODUCTION

IN the context of the diagnosis of electrical systems, rotating machines occupy a predominant place. Statistical studies have shown that failures due to bearings are paramount regardless of the machines power range. Hence, the necessity of their monitoring and diagnosis in order to increase the service quality [1], [2]. Therefore, many important researches had been done in the advanced field of bearing fault diagnosis [1], [3], [4]. 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 [1], [5]. The key step of pattern classification and recognition is feature extraction from the vibration signal.
Frequency and Time-domain features are simple and effective features without heavy computations which especially are suitable for condition monitoring systems [6], [7], [8].
Several applications used the wavelet analysis as a powerful filtering tools [1], [5], [9], the discrete version , still called multiresolution analysis consists to double filtering the signal by a filter bank called wavelet. It allows a clear visualization of each part of the signal with a resolution adapted to its scale [10].
Support Vector Machine is a relatively new computational supervised learning method which was introduced by Vapnik. It uses structural risk minimization (SRM) principle to minimize an upper bound on the expected risk whereas in ANN, The error is minimized using empirical risk minimization (ERM) [11], [5]. They have been shown better generalization performance than artificial neural network (ANN), for this reason, the algorithms used for SVM started emerging with greater availability of computing power [12], [13] [14], [15]. 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 [16], [17].

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) [14], [18], [19].
PSO was proposed by Kennedy and Eberhart [20], [21]. And it is inspired by the social behavior of bird flocking, fish schooling and swarm theory, etc. The theoretical framework of PSO is very simple, and PSO possesses the properties of easy implementation and fast convergence [14], [22].

The proposed algorithm for automatic bearing condition classification has several steps. The first step is the digital processing of acquired data. Section 3 describes support vector machine.The PSO algorithm is discussed in section 4. The next section addresses the experimental setup . In section 6 the proposed methodology is discussed, The experimental results are presented in Section 7.

II. FEATURE EXTRACTION

The fault diagnosis is essentially a problem of pattern recognition, of which, an important step is feature extraction. The step of feature extraction is shown as follows:

- Firstly, the original signal is estimated by frequency domain features.
- Secondly, seven statistical feature parameters are extracted.

A. Frequency domain features

the frequency spectrum is extracted using the method proposed in [6], this approach consists on three steps:

This work was supported by Laboratory of Electrical Engineering, LABGET, Tebessa University, Algeria.
T. Thelaidjia is with Laboratory of Electrical Engineering, LABGET, Larbi Tebessi University, Tebessa, 12002, Algeria (bilel.toufik@yahoo.fr).
S. Chenikher is with with Laboratory of Electrical Engineering, LABGET, Larbi Tebessi University, Tebessa, 12002, Algeria (s.chenikher@mail.univ-tebessa.dz).
A. Moussaoui is with Department of Electrical Engineering, Guelma University, Guelma, Algeria .

ISSN 1737-930X

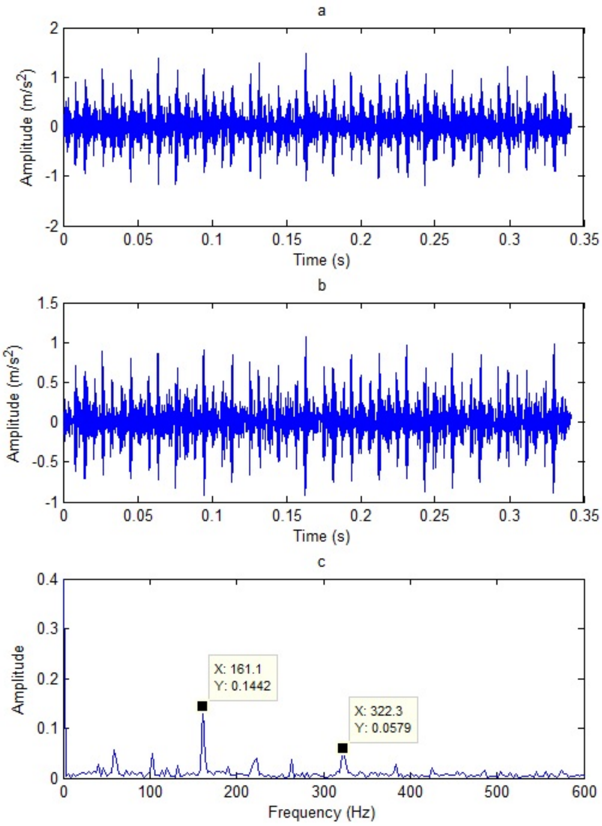

Fig. 1. **a** Measured signal, **b** Reconstructed signal and **c** Its envelope spectrum. Bearing with inner race defect.

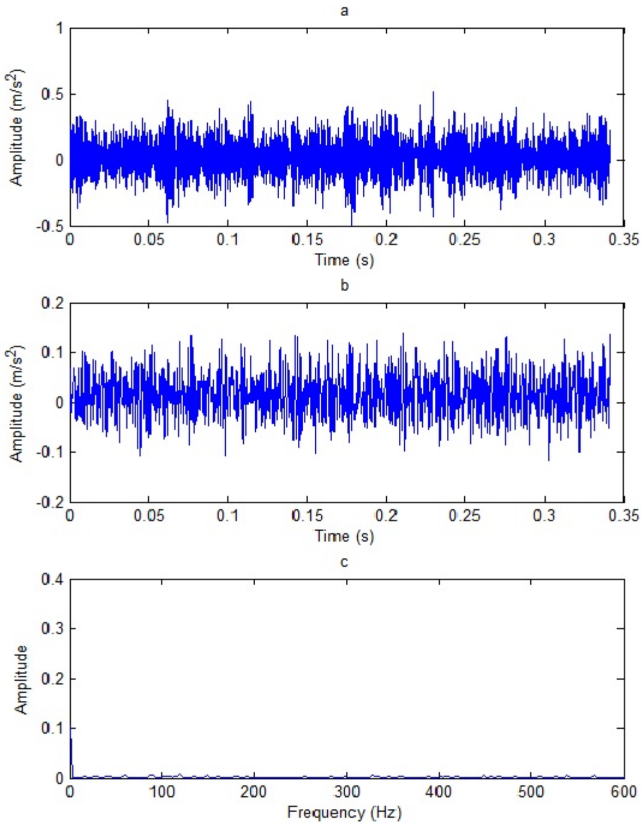

Fig. 2. **a** Measured signal, **b** Reconstructed signal and **c** Its envelope spectrum. Bearing with ball defect.

- **Step 1** Filtering the vibration signals using discrete wavelet transform.
- **Step 2** Extracting the envelope using Hilbert transform.
- **Step 3** Applied the Fourier transform to calculate the envelope spectrum.

After the obtaining of the frequency spectrum, it will be diveded into eight bands.
then we apply the energy at each band as an input variable for SVM classification. The Measured signals, their reconstructed signals and their envelope spectrums are presented in Fig.1, Fig.2, Fig.3 and Fig.4.

B. Features in the time domain

To describe the vibration signal in the time domain, 4 statistical quantities were used: mean value, root mean square value, skewness and kurtosis.

1) The Mean:

$$M = mean(x) \tag{1}$$

2) The standard deviation :

$$SD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x(i)-\overline{x})^2} \tag{2}$$

3) The impulse factor :

$$IMF = \frac{PV}{\frac{1}{N}\sum_{i=1}^{N} x(i)} \tag{3}$$

4) The kurtosis:

$$x_{kur} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x(i)-\bar{x})^4}{\left[\frac{1}{N}\sum_{i=1}^{N}(x(i)-\bar{x})^2\right]^2} \tag{4}$$

III. Support Vector Machine

The basic idea of SVM is to transform the input to a higher dimensional feature space.Then the SVM solves binary problem in which an hyperplane separate data.The hyperplane is defined through the use of support vectors [23].
Support vector machine (SVM) based on statical learning theory is proposed according to optimal hyperplane in the case of linear separable [1].
If all samples are correctly separated by an hyperplane, it must satisfy the following condition [4]:

$$y_k(\langle w; x\rangle - \lambda_0) \geq +1, \forall k \in \{1, ..., n\} \tag{5}$$

The following function should be minimized In order to find the optimal hyperplane [4].

$$\varphi(w) = \frac{1}{2}\left\|w\right\|^2 \tag{6}$$

Saddles of Lagrange function give solution of the optimal problem as below:

$$L(w, \lambda_0, \alpha) = \frac{\left\|w\right\|^2}{2} - \sum_{k=1}^{n}\alpha_k\left[y_k(\langle w; x\rangle - \lambda_0) - 1\right] \tag{7}$$

ISSN 1737-930X

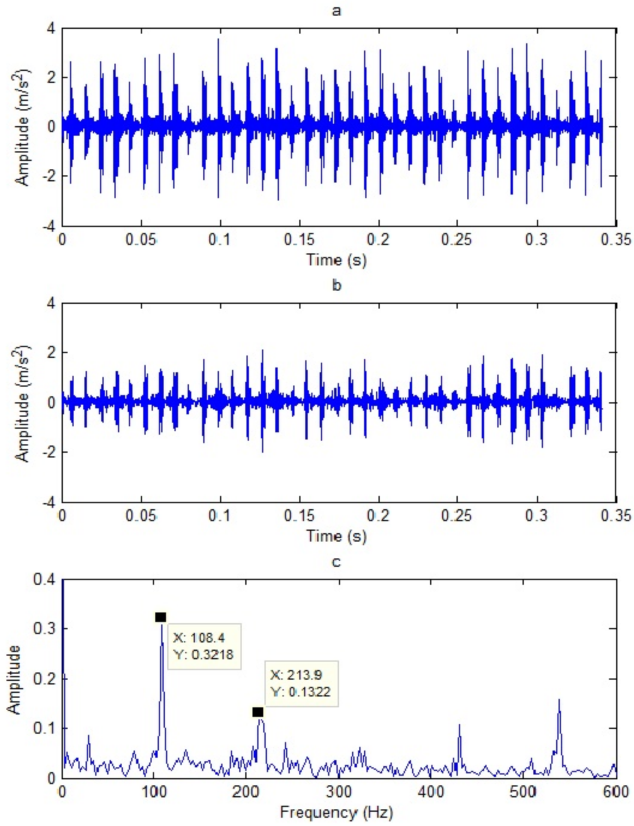

Fig. 3. **a** Measured signal, **b** Reconstructed signal and **c** Its envelope spectrum. Bearing with outer race defect.

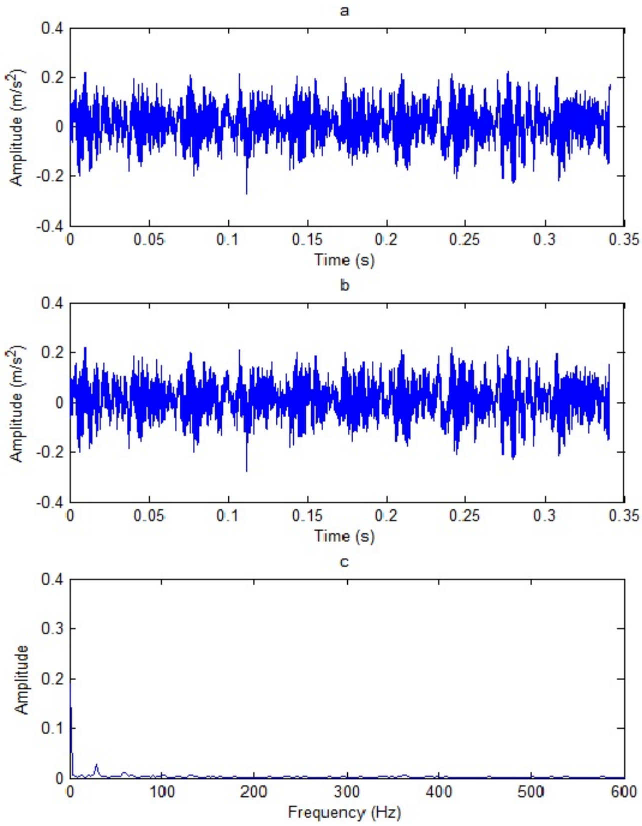

Fig. 4. **a** Measured signal, **b** Reconstructed signal and **c** Its envelope spectrum. Normal Bearing .

where $\alpha = (\alpha_1 ... \alpha_n)$ is the Lagrange coefficient;$\alpha_i \geq 0, \forall i$
The dual problem can be obtained as below :

$$\left\{ W(\alpha) = \sum_{k=1}^{n} \alpha_k - \frac{1}{2} \sum_{k,k'=1}^{n} \alpha_k \alpha_{k'} y_k y_{k'} \left\langle x_k ; x_{k'} \right\rangle , \right\} \quad (8)$$

subject to:

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

If α^* is the optimal solution, then

$$\langle w^* ; x \rangle = \sum_{k=1}^{n} \alpha_k^* y_k \langle x_k ; x \rangle \quad (9)$$

According to the $Kuhn-Tucker$ condition, the solution must satisfy

$$\alpha_k^* \left[y_k (\langle w^* ; x_k \rangle - \lambda_0^*) - 1 \right] = 0. \quad (10)$$

were λ_0^* is given by

$$\lambda_0^* = \frac{1}{N_{sv}} \sum_{s=1}^{N_{sv}} (y_s - x_s^T w^*), s = 1 : N_{sv} \quad (11)$$

The decision function is given by:

$$D(x) = sgn \left[\sum_{k \in S} \alpha_k^* y_k x_i^T x - \lambda_0^* \right] \quad (12)$$

soft-margin SVM solved The nonseparable problem [15], [17], [23].

If we used the inner $\kappa(x_k, x)$ to perform the transformation of original feature space into higher dimensional feature space, the dual problem 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:
$\sum_{k=1}^{n} \alpha_k y_k = 0$ et $0 \leq \alpha_k \leq c$
The decision function is written as below:

$$D(x) = sgn \left[\sum_{k \in S} \alpha_k^* y_k \kappa(x_i^T x) - \lambda_0^* \right] \quad (13)$$

here, $\kappa(x_k, x)$ is called kernel function.
Some kernel functions can be expressed as below [12], [23]:
Polynomial:

$$\kappa(x_1, x_2) = (1 + \langle x_1 ; x_2 \rangle)^q \quad (14)$$

where parameter q is the degree of the polynomial.
Radial basis function (RBF):

$$\kappa(x_1, x_2) = exp(- \|x_1 - x_2\|^2 / 2\sigma^2) \quad (15)$$

where parameter σ^2 is the variance of the Gaussian function.
Sigmoid:

$$\kappa(x_1, x_2) = tanh(\alpha_0 \langle x_1 ; x_2 \rangle + \beta_0) \quad (16)$$

where α_0 and β_0 are the parameters of kernel function.
Support Vector Machines are binary classifiers. However Most of cases in practical, usually more than two classes.

ISSN 1737-930X

In order to overcome this particularity of SVMs, different multiclass strategies have been proposed [12], [24], [25]. Three techniques affect The classification performance of SVM : the selecting of the kernel, the choosing of the kernel parameters, and the choosing of the regularization parameter c [4].
As presented in different works, 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 [26], [27], c and σ exhibit a (strong) interaction. As a consequence, they should be optimized simultaneously, rather than separately.

IV. The Particle Swarm Optimization

The particle swarm optimization (PSO) consists of a swarm of particles flying through the search space. Each particle is treated as a point in a D-dimensional space. The i-th particle is represented $Z_i = (Z_{i1}, Z_{i2}, ..., Z_{id}, ..., Z_{iD})$. The best previous position of any particle is recorded and represented as $P_i = (P_{i1}, P_{i2}, ..., P_{id}, ..., P_{iD})$. The index of the best particle among all the particles in the population is represented by the symbol G. The rate of the position change (velocity) for the i-th particle is represented as $V_i = (V_{i1}, V_{i2}, ..., V_{id}, ..., V_{iD})$. The updated velocity and position of the i-th particle at the k-th iteration are [18]:

$$V_{id}^k = \omega.V_{id}^{k-1} + c_1.r_1.(P_{id} - Z_{id}^{k-1}) + c_2.r_2.(P_{Gd} - Z_{id}^{k-1}) \quad (17)$$

$$Z_{id}^k = Z_{id}^{k-1} + V_{id}^k \quad (18)$$

Where c_1 and c_2 are constants knows as the cognitive and social acceleration coefficients, respectively, ω is the inertia weight, r_1 and r_2 are random numbers between 0 and 1.
The first part of (17) represents the previous velocity, which provides the necessary momentum for particles to fly across the search space. The second part is the ”cognition” part, which represents the private thinking of the particle itself.. The third part is known as the ”social” component, which represents the collaboration among the particles. In addition, the implementation of PSO also requires placing a limit on the particle velocity, and the limit, i.e. the maximum allowed velocity V_{max}, determines the searching granularity of space. The inertia weight ω plays the role of balancing the global search and local search, and it can be a positive constant or even a positive linear or nonlinear function of time.
The classification performance of SVM are affected by two techniques, the choosing of the kernel parameters, and the choosing of the regularization parameter c [5].
The proposed approaches for SVM parameter optimization with PSO, is as follows:
Step 1. *Particle initialization and PSO parameters setting*: Set the PSO parameters including: c_1, c_2, position of each particle, velocity of each particle, number of particles, number of iterations and velocity limitation .
Step 2. *Fitness evaluation*: Perform SVM on each particle in population and compute the prediction accuracy.
Step 3. Update the global and personal best (P_i and P_G) according to the fitness evaluation results.
Step 4. *Particle manipulations*: Each particle moves to its next position using formula (17) and (18).

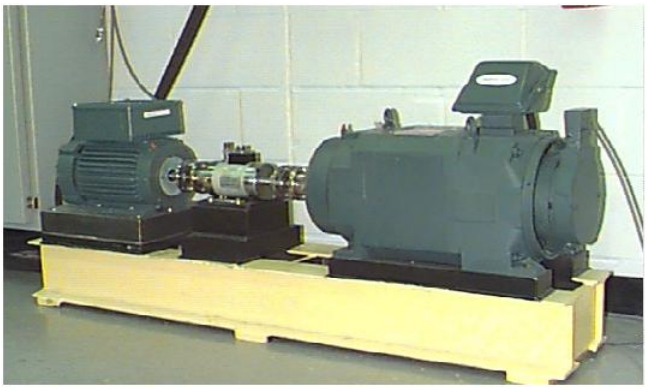

Fig. 5. The test apparatus of bearing

Step 5. *Stop condition checking*: If stopping criteria (maximum iterations predefined) are not met, go to **step 2**, otherwise, go to the next step.
Step 6. End the training and testing procedure and save optimal c , σ for SVM.

In order to select the optimal values of the PSO parameters, a series of experiments had been carried out by varying the values of these parameters.

The swarm size is set to 20 particles. The searching ranges for c and σ are as follows: $c \in [0, 10]$, and $\sigma \in [0, 10]$. Preliminary experiments also let this study set the personal and social learning factors (c1, c2) = (1.3, 1.3) that achieves better classification accuracy.
The inertia weight is set to the following equation:

$$\omega(k) = \omega_{max} - \frac{(\omega_{max} - \omega_{min})}{k_{max}}.k \quad (19)$$

where ω_{max} is the initial weight, ω_{min} is the final weight, k_{max} is the maximum number of iterations or generation, and k is the current iteration number. The predefined maximum iteration is 10. When the maximum iteration is reached, the accuracy of test set is calculated by the predicted output of the trained SVM classifier.

V. Experimental Setup and Vibration Data

The vibration data used in this paper have been obtained from the ball bearing test data set of the Western Reserve University Bearing Data Center Website [28]. As shown in Fig. 5, the test stand consists of a 2hp Reliance Electric motor (left), a torque transducer/encoder (center), a dynamometer (right), and control electronics (notshown) [29]. The data is sampled at a rate of 12 kHz, Each signal is 4096 samples long. Data was gathered for four different conditions: (i) inner race fault (IF) ; (ii) ball fault (BF); (iii) outer race fault (OF); (iv) normal (H) . Faults were introduced into the drive end bearing by using electro-discharge machining.
The bearing monitored is a deep groove ball bearing. it is a $6205 - 2RSJEM$ bearing with a FI, a BF and a FO frequency equal to 5.4152, 4.7135 and 3.5848 times the shaft frequency, respectively.
Theoretical estimations of the FI, BF and FO frequencies are presented at Table I.

ISSN 1737-930X

TABLE I
THEORETICAL ESTIMATIONS OF CHARACTERISTIC FAULT FREQUENCIES

FI	BF	FO
162.1852	141.1693	107.3647

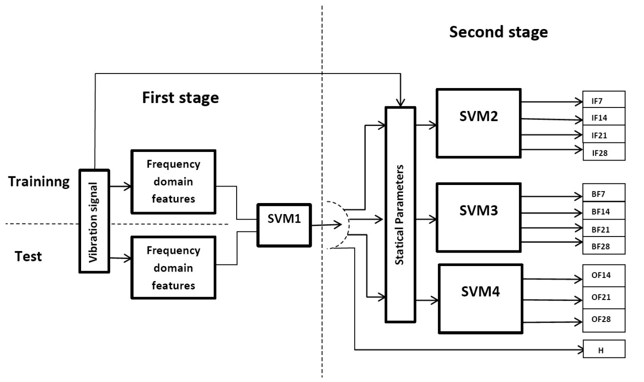

Fig. 6. Proposed diagnosis methodology scheme including feature Extraction and classification.

VI. Classification Strategy

The complete methodology is represented in Fig 6. A a two stage Support Vector Machine (SVM) approach is adapted. First, five Support Vector Machines are trained.The first SVM(SVM1) is trained in order to detect the type of fault (Normal (H), outer race (OF), inner race (OF) or ball fault (BF)). For this reason, the energie of the frequency spectum features are used. When the test signal represents a normal bearing condition, the classification process is over. Otherwise, the fault diagnosisis is transferred to the second stage, where SVM2, SVM3, SVM4 are used to identify the quality of fault one for each fault scenario.

The bearing fault conditions of data are shown in table II

TABLE II
The database

Bearing condition	Fault specifications		class
	Diameter (inches)	Depth (inches)	
Healthy	-	-	H
Inner race fault	0.007	0.0011	IF7
	0.014	0.0011	IF14
	0.021	0.0011	IF21
	0.028	0.0011	IF28
Outer race fault	0.014	0.0011	OF14
	0.021	0.0011	OF21
	0.028	0.0011	OF28
Ball fault	0.007	0.0011	BF7
	0.014	0.0011	BF14
	0.021	0.0011	BF21
	0.028	0.0011	BF28

The leave-one-out ”LOO” validation technique is adapted to guarantee valid results for making predictions regarding new data.

VII. Analysis of Experimentation Results

To identify the mechanical failure of roller bearing, the feature extraction procedure was performed using the proposed technique.
the the one-against-the-rest SVM classification method is used. Once the data dimension was reduced, The leaving-one-out method is applied by excluding a certain sample from the set of known patterns, and the classifier is designed without it; afterward, the classifier is tested precisely on this pattern. The procedure is repeated for all patterns from both classes.
Two remarks can be extracted:

- The input vector varies depending on the input signal. for healthy bearing kurtosis is around 3 this parameter is increased in the case of defect rolling.
- When the state information changes, the energy of spectral bands varies also.

Table III gives the classification result for this bearing fault classification problem.

TABLE III
Classification result of bearing fault in validation and test

Classifier	Optimal ”c”	Optimal ”σ”	Validation Rate	Test Rate
SVM1	33.97	0.004	100	100
SVM2	1.056	3.838	100	100
SVM3	5.007	3.642	100	100
SVM4	1.524	5.075	100	100

VIII. Conclusion

The present paper introduced a new automated diagnosis method for different types of rolling bearing faults. It should be noted that, this scheme is applied here not only to distinguish between ball, inner and outer faults, but also for faults severity detection.
The proposed strategy applied two stage Support Vector Machine(SVM) approach for the automatic diagnosis of defective rolling element bearings condition. For feature extraction, Frequency domain features are used in the first stage, whereas, statistical parameters are utilized in the second stage. Particle Swarm Optimisation (PSO) is employed to simultaneously optimize the SVM kernel function parameter and the penalty parameter.
PSO optimization requires only simple mathematical operators. This algorithm is simple to implement and effective, and is inexpensive in terms of memory and time required.
the good results obtained highlight the effectiveness of the proposed method for bearing fault diagnosis.

References

[1] T.Thelaidjia, A. Moussaoui and S.Chenikher, “Feature extraction and optimized support vector machine for severity fault diagnosis in ball bearing,” *Engineering Solid Mechanics*, vol. 4, no. 4, pp. 167–176, 2016.
[2] Z. Chen, S. Deng, X. Chen, C. Li, R.V. Sanchez and H. Qin , “Deep neural networks-based rolling bearing fault diagnosis,” *Microelectronics Reliability*, vol. 75, no. 14, pp. 327–333, 2017.
[3] Z. Tong, W. Li, B. Zhang and M. Zhang , “Bearing fault diagnosis based on domain adaptation using transferable features under different working conditions,” *Shock and Vibration*, 2018.
[4] D. Dyer and M. Stewart, “Detection of rolling element bearing damage by statistical vibration analysis,” pp. 229–235, 1978.
[5] G. Xian, “Mechanical failure classification for spherical roller bearing of hydraulic injection moulding machine using dwt-svm,” *ELSEVIER*, August 2010.
[6] D. Dyer and M. Stewart, “Detection of rolling element bearing damage by statistical vibration analysis,” pp. 229–235, 1978.
[7] A. Boulenger and C. Pachaud, “Diagnostic vibratoire en maintenance prventive,” *Dunod, Paris*, p. 1998.

ISSN 1737-930X

[8] R. S. C. Pachaud and C. Fray, “Crest factor and kurtosis contributions to identify defects inducing periodical impulsive forces,” *Mech Syst Signal Process*, vol. 11(6), pp. 903–916, 1997.

[9] A. B. Patil , J. A. Gaikwad , J. V. Kulkarni, “Bearing fault diagnosis using discrete wavelet transform and artificial neural network,” *IEEE International Conference on Applied and Theoretical Computing and Communication Technology*, 2016.

[10] C. Pachaud, R. Salvetas, C. Fray, “Crest factor and kurtosis contributions to identify defects inducing periodical impulsive forces,” *Mech Syst Signal Process*, vol. 11(6), pp. 903–916, 1997.

[11] L. Shuang and L. Meng, “Bearing fault diagnosis based on pca and svm,” *IEEE International Conference on Mechatronics and Automation*, August 2007.

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

[13] ——, *Statistical Learning Theory*. Springer, 1998.

[14] 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.

[15] S. Mallat, “A theory for multiresolution signal decomposition: The wavelet representation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11(7), pp. 674–693, August 1989.

[16] P. K. Kankar, C. S. Satish and S. P. Harsha, “Fault diagnosis of ball bearings using machine learning methods,” *ELSEVIER*, 2011.

[17] 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.

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

[19] S. ZHANG, F. KUANG, Y. WANG and L. WANG , “A novel svm model with pso on power transformer fault diagnosis,” *Journal of Computational Information Systems*, vol. 8, no. 14, 2012.

[20] J. Kennedy and R. C. Eberhart, “Particle swarm optimization,” *Proceedings of International Conference on Neural Networks, IEEE*, pp. 1942–1948, 1995.

[21] J Kennedy, R C Eberhart and Y.Shi, “Swarm intelligence,” *Morgan Kaufmann Publishers Inc*, 2001.

[22] J. LI, “A combination of pso and svm for road icing forecast,” *Journal op computers*, vol. 5, no. 9, p. 2010, September.

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

[24] 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.

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

[26] 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.

[27] 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.

[28] K. A. Loparo, “Bearings vibration data set.case western reserve university,” */http://www.eecs.case.edu/laboratory/bearing/download.htm*, consulted on september 2012.

[29] http://csegroups.case.edu/bearingdatacenter/pages/apparatus-procedures .

ISSN 1737-930X