Journal of Applied and Computational Mechanics, Vol. 1, No. 1, (2015), 35-43
DOI: 10.22055/jacm.2014.10547

Journal of
Applied and Computational Mechanics

Bearing Fault Detection Based on Maximum Likelihood Estimation and Optimized ANN Using the Bees Algorithm

Behrooz Attaran[1], Afshin Ghanbarzadeh[2]

[1] Master of Science, Department of Mechanical Engineering, Shahid Chamran University
Golestan Street, Ahvaz, 61848-54385, Iran, AttaranBehrooz@yahoo.com
[2] Assistant Professor, Department of Mechanical Engineering, Shahid Chamran University
Golestan Street, Ahvaz, 61357-43337, Iran, Ghanbarzadeh.A@Scu.ac.ir

Received May 5 2014; revised May 18 2014; accepted for publication May 30 2014.
Corresponding author: Behrooz Attaran, AttaranBehrooz@yahoo.com

Abstract

Rotating machinery is the most common machinery in industry. The root of the faults in rotating machinery is often faulty rolling element bearings. This paper presents a technique using optimized artificial neural network by the Bees Algorithm for automated diagnosis of localized faults in rolling element bearings. The inputs of this technique are a number of features (maximum likelihood estimation values), which are derived from the vibration signals of test data. The results show that the performance of the proposed optimized system is better than most previous studies, even though it uses only two features. Effectiveness of the above method is illustrated using obtained bearing vibration data.

Keywords: Fault diagnosis, MLE distributions, RBF neural network, Bees Algorithm.

1. Introduction

Various diagnostic tools exist for diagnosing fault in machineries; the most common is vibration-based tools. Using vibration collected data from defective components, algorithms are developed to detect the time when bearing damage has been occurred. Over the past 25 years, numerous vibration-based algorithms for bearing damage detection have been developed. Unfortunately, up to now, a complete database of existing vibration algorithms along with their capabilities and limitations is not presented. A pertinent review and an application of an ensemble of hybrid intelligent models are presented earlier [1]. The monitoring methods applied to bearings can be achieved in a number of ways [2], with some of the methods being simple to use, and others requiring sophisticated signal processing. Shocks are usually created in the presence of faults and can be analyzed either in the time domain [3] (RMS and max-peak amplitude of vibration level, Crest factor and Kurtosis, detection of shock waves and Julien method [4], statistical parameters applied to the time signal, Cepstrum) or in the frequency domain (spectral analysis around bearing defect frequencies [5], frequency spectrum in the high frequency domain, Spike energy [6], high frequency demodulation [7], acoustic emission [8-9], adaptive filtering [10-11], artificial neural networks [12-13], time-frequency [14], the Bees Algorithm [15] and etc). The application of ANNs has gained importance in the area of automated fault detection and diagnosis of rotating machinery [16]. The neural networks have the advantages of adaptive learning, nonlinear generalization, fault tolerance, resistance to noisy data, and parallel computation abilities.

The proposed method, several domain features (time signal, real Cepstrum, minimum phase reconstruction, discrete cosine transform, discrete Fourier transform, envelope analysis signal and the Hilbert transform) are extracted from the vibration signal and pattern recognition using ANN is used for bearing fault diagnosis. Measured Vibration signals from a single location are used in this method. Some of previous works dealt with signals from multiple locations for fault detection [17-18]. The number of input parameters used in the proposed algorithm is less

than previous works and hence, training speed is high, the performance of the maximum likelihood estimation features are compared with other descriptive statistic and it has been observed that they perform better when used as input features of RBF for the fault diagnosis of bearings. In the final step, the train of RBF network is optimized by means of the Bees Algorithm.

The remainder of this paper is organized as follows. In section 2, the experimental setup is introduced. Then, in section 3, signal processing techniques is developed. Section 4 presents various feature extraction methods. In section 5, the feature selection is invented. Section 6 presents the Bees Algorithm. The results and conclusions of this investigation are summarized in sections 7 and 8.

2. Experimental Setup

The bearings used in this study are deep groove ball bearings generated by SKF. The faults were introduced into the drive end bearing of the motor using the electro discharge machining method. Each bearing was tested under four different loads, (e.g. 0, 1, 2, and 3 hp). The bearing dataset was derived from the experimental system under the four different operating conditions: (1) normal condition; (2) outer race fault; (3) inner race fault; and (4) ball fault (Table 1). The number of training samples is 20 samples and the number of testing samples is 10 [19].

Table 1. Description of the Experimental Dataset

Load (HP)	Motor speed (rpm)	Operating condition	Label of class
0	1797	Normal baseline	C1
1	1772	Normal baseline	C1
2	1750	Normal baseline	C1
3	1730	Normal baseline	C1
0	1797	Inner race fault	C2
1	1772	Inner race fault	C2
2	1750	Inner race fault	C2
3	1730	Inner race fault	C2
0	1797	Ball Fault	C3
1	1772	Ball Fault	C3
2	1750	Ball Fault	C3
3	1730	Ball Fault	C3
0	1797	Outer race fault	C4
1	1772	Outer race fault	C4
2	1750	Outer race fault	C4
3	1730	Outer race fault	C4

3. Signal processing techniques

Features extracted from time-domain vibration signals, real Cepstrum, minimum phase reconstruction, discrete cosine transform, discrete Fourier transform, envelope analysis signal and the Hilbert transform are used as input features for the neural network.

The complex Cepstrum for a sequence x is calculated by finding the complex natural logarithm of the Fourier transform of x, then the inverse Fourier transform of the resulting sequence.

$$\hat{x} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left[X\left(e^{j\omega}\right)\right] e^{j\omega n} d\omega \tag{1}$$

The real Cepstrum is the inverse Fourier transform of the real logarithm of the magnitude of the Fourier transform of a sequence. The real Cepstrum (y) is define as

$$y = real\left(ifft\left(\log\left(abs\left(fft\left(x\right)\right)\right)\right)\right) \tag{2}$$

Appropriate windowing in the Cepstral domain forms the reconstructed minimum phase (ym) signal:

$$w = \begin{bmatrix} 1 \\ 2\times ones\left(n/2-1,1\right) \\ ones\left(1-rem\left(n,2\right),1\right) \\ zeros\left(n/2-1,1\right) \end{bmatrix} \tag{3}$$

$$ym = real\left(ifft\left(\exp\left(fft\left(w.*y\right)\right)\right)\right) \tag{4}$$

The discrete cosine transform (DCT) is closely related to the DFT. The DCT's energy compaction properties are useful for applications like signal coding. Mathematically, the unitary DCT of an input sequence x is:

$$y(k)=w(k)\sum_{n=1}^{N}x(n)\cos\left(\frac{\pi(2n-1)(k-1)}{2N}\right),k=1,...,N \tag{5}$$

where

$$w(k)=\begin{cases}\dfrac{1}{\sqrt{N}},k=1\\ \sqrt{\dfrac{2}{N}},2\le k\le N\end{cases} \tag{6}$$

The faulty features can be detected by Hilbert transform of raw signal in time-domain. The high-frequency vibration amplitude of operating bearings with local faults was modulated by pulse (extra) force. In order to obtain the fault characteristic, the vibration signals of rolling element bearings need to demodulate. Defining a series *x(t)* of raw signal, we can have its Hilbert transform as

$$H\left[x(t)\right]=\frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{x(t)}{t-\tau}d\tau \tag{7}$$

Then, the envelope analysis signal can be given as

$$h[s]=\int_{-\infty}^{+\infty}\sqrt{x^2(t)+H^2\left[x(t)\right]}d\tau \tag{8}$$

4. Feature Extraction

4.1 Descriptive statistics

Data samples can have thousands (even millions) of values. Descriptive statistics can summarize these data into a few numbers that contain most of the relevant information. The Following statistical parameters are used to detect incipient bearing damage:

$$Root\ Mean\ Square,\ \ y_{rms}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i)^2} \tag{9}$$

$$Interquartile\ Range,\ \ y_{iqr}=iqr(X) \tag{10}$$

$$Skewness,\ \ y_{skw}=\frac{E(X-\mu)^3}{\sigma^3} \tag{11}$$

$$Kurtosis,\ \ y_{kur}=\frac{E(X-\mu)^4}{\sigma^4} \tag{12}$$

$$Mean,\ \ y_{mea}=\frac{1}{N}\sum_{i=1}^{N}(x_i) \tag{13}$$

$$Harmonic\ Mean,\ \ y_{har}=\frac{N}{\sum_{i=1}^{N}\frac{1}{X_i}} \tag{14}$$

$$Largest\ Element,\ \ y_{max}=\max(X) \tag{15}$$

$$Smallest\ Element,\ \ y_{min}=\min(X) \tag{16}$$

$$Most\ Frequent\ Value,\ \ y_{mod}=mode(X) \tag{17}$$

$$Std,\qquad y_{std}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i-\bar{X})^2} \tag{18}$$

$$Variance,\ \ y_{var}=variance(X) \tag{19}$$

$$Median,\ \ y_{med} = median(X) \tag{20}$$

$$Sum,\ \ y_{sum} = \sum_{i=1}^{N} X_i \tag{21}$$

$$Trapezpidal\ Numerical\ Integration, y_{trp} = trapz(X) \tag{22}$$

$$Mean\ Absolute\ Deviation,\ \ y_{mad} = mad(X) \tag{23}$$

$$Percentiles,\ \ y_{prc} = prctile(X, p) \tag{24}$$

Where X_i (i=1, …, N) is the amplitude at sampling point i and N is the number of sampling points. μ is the mean of X, σ is the standard deviation of X, and E represents the mathematical expectation.

4.2 Innovative indicators

Two new descriptors called TALAF and THIKAT are presented by S. Sassi [18]. The TALAF is defined as

$$TALAF = \log\left[kur + \frac{rms}{rms_0} \right] \tag{25}$$

Where rms_0 is the root mean square value define for a healthy bearing. The THIKAT is defined as

$$THIKAT = \log\left[\left(kur\right)^{CF} + \left(\frac{rms}{rms_0} \right)^{peak} \right] \tag{26}$$

Where CF is crest factor of X.

4.3 Maximum likelihood estimates

Each function represents a parametric family of distributions. Input arguments are data, presumed to be samples from some member of the selected distribution family. Functions return maximum likelihood estimates (MLEs) of distribution parameters, that is, parameters for the distribution family member with the maximum likelihood of producing the data as a random sample.

MLEs can be heavily biased, especially for small samples. As sample size increases, however, MLEs become unbiased minimum variance estimators with approximate normal distributions. This is used to compute confidence bounds for the estimates.

The various distributions are given in Table 2.

Table 2. Maximum Likelihood Estimation Distributions

Distribution	Describe
Exponential	mean of an exponentially distributed
extreme value (ev1)	location parameter, mu
extreme value (ev2)	scale parameter, sigma
gamma1	gamma distribution parameter1
gamma2	gamma distribution parameter2
generalized extreme value (gev1)	shape parameter, K
generalized extreme value (gev2)	scale parameter, sigma
generalized extreme value (gev3)	location parameter, mu
generalized pareto (gp1)	tail index (shape) parameter, K
generalized pareto (gp2)	scale parameter, sigma
Geometric	Geometric
lognormal1	mean, mu
lognormal2	standard deviation, sigma
normal1	mean, mu
normal2	standard deviation, sigma
Poisson	Poisson distribution, λ
Rayleigh	parameter of the Rayleigh distribution
discrete uniform	max(data)
uniform1	uniform distribution parameter1
uniform2	uniform distribution parameter1
weibull1	Weibull parameter *a*
weibull2	Weibull parameter *b*

5. Feature Selection

Feature selection has a significant impact on the success of pattern recognition. The Euclidean distance between pairs of classes is computed in this section. Given an m-by-n data matrix X, which is treated as *m* (1-by-*n*) row vectors $x_1, x_2, \ldots, x_m$, the Euclidean distances between the vector x_r and x_s is defined as

$$d_{rs}^2 = (x_r - x_s)(x_r - x_s)' \tag{27}$$

$$\begin{aligned}\text{TotalDistance} &= \min(d_{12}, d_{13}, d_{14}) \times \min(d_{21}, d_{23}, d_{24}) \\ &\times \min(d_{31}, d_{32}, d_{34}) \times \min(d_{41}, d_{42}, d_{43})\end{aligned} \tag{28}$$

Where d_{ij} is Euclidean distance between *Ci* and *Cj*. Maximum Total Distances are selected for input features of artificial neural network.

6. The Bees Algorithm

Figure 1 shows the flowchart for the algorithm in the simplest form. For more details, the reader is referred to [15]. The algorithm requires a number of parameters to be set, namely: number of scout bees (*n*), number of sites selected for neighborhood searching (out of *n* visited sites) (*m*), number of top-rated (elite) sites among *m* selected sites (*e*), number of bees recruited for the best *e* sites (*nep*), number of bees recruited for the other (*m-e*) selected sites (*nsp*), the initial size of each patch (*ngh*) (a patch is a region in the search space that includes the visited site and its neighborhood), and the stopping criterion.

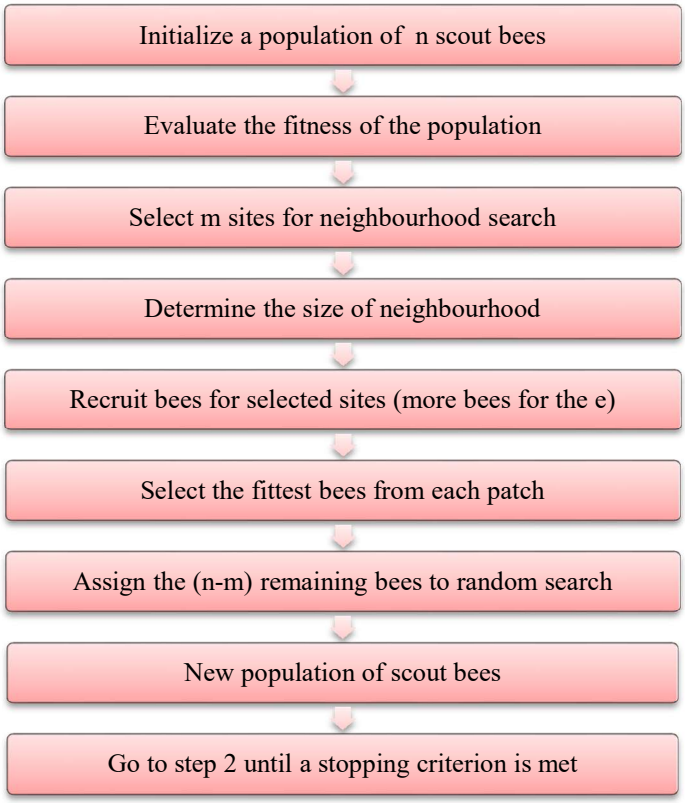

Fig. 1. Flowchart of the Bees Algorithm (BA)

7. Results

7.1 Feature Selection result

There are 287 features in seven domains (time, real Cepstrum, minimum phase reconstruction, discrete cosine transform, discrete Fourier transform, envelope analysis signal and the Hilbert transform). Initially, one feature, two features and three features were selected and then Total Distance for these features was calculated. Finally, features were selected by Total Distance and the first six of them (Table 3).

Table 3. A comparison on Total Distance

Feature 1	Feature 2	Feature 3	Total Distance
Geometric (Envelope)	---	---	0.00342
Geometric (Hilbert)	---	---	0.00328
Interquartile range (Time)	---	---	0.00243
gev2 (Hilbert)	---	---	0.00242
gev2 (Envelope)	---	---	0.00240
weibull1 (Time)	---	---	0.00229
Geometric (Envelope)	ev1 (real cepstrum)	---	0.24890
lognormal1 (DCT)	gev3 (minimum phase)	---	0.24881
Geometric (Hilbert)	ev1 (real cepstrum)	---	0.24811
gev3 (Time)	lognormal1 (DCT)	---	0.24801
lognormal1 (Hilbert)	Exponential (DCT)	---	0.24139
lognormal1 (Hilbert)	Poisson (DCT)	---	0.24139
Geometric (Envelope)	gp1 (FFT)	ev1 (real Cepstrum)	1.04083
Geometric (Hilbert)	gp1 (FFT)	ev1 (real Cepstrum)	1.03884
Exponential (Hilbert)	gp1 (FFT)	ev1 (real Cepstrum)	1.01300
normal1 (Hilbert)	gp1 (FFT)	ev1 (real Cepstrum)	1.01300
Poisson (Hilbert)	gp1 (FFT)	ev1 (real Cepstrum)	1.01300
Mean (Hilbert)	gp1 (FFT)	ev1 (real Cepstrum)	1.01300

7.2 RBF Neural network

These features are used as input features for RBF neural network and calculation of mse of train and mse of test (Table 4).

Table 4. A comparison on mse train and mse test of RBF

Feature 1	Feature 2	Feature 3	mse train	mse test
Geometric (Envelope)	---	---	9.19E-05	8.93E-05
Geometric (Hilbert)	---	---	9.38E-05	8.98E-05
weibull1 (Time)	---	---	2.28E-04	2.13E-04
gev2 (Hilbert)	---	---	2.72E-04	2.62E-04
gev2 (Envelope)	---	---	2.75E-04	2.66E-04
Interquartile range (Time)	---	---	1.80E-03	2.72E-03
lognormal1 (DCT)	gev3 (minimum phase)	---	9.89E-06	2.69E-05
gev3 (Time)	lognormal1 (DCT)	---	9.98E-06	4.55E-05
Geometric (Hilbert)	ev1 (real cepstrum)	---	1.44E-05	4.20E-05
Geometric (Envelope)	ev1 (real cepstrum)	---	1.48E-05	4.43E-05
lognormal1 (Hilbert)	Exponential (DCT)	---	3.88E-05	1.98E-04
lognormal1 (Hilbert)	Poisson (DCT)	---	3.88E-05	1.98E-04
normal1 (Hilbert)	gp1 (FFT)	ev1 (real cepstrum)	3.34E-06	6.13E-05
Poisson (Hilbert)	gp1 (FFT)	ev1 (real cepstrum)	3.34E-06	6.13E-05
Mean (Hilbert)	gp1 (FFT)	ev1 (real cepstrum)	3.34E-06	6.13E-05
Exponential (Hilbert)	gp1 (FFT)	ev1 (real cepstrum)	3.34E-06	6.13E-05
Geometric (Hilbert)	gp1 (FFT)	ev1 (real cepstrum)	2.90E-06	4.77E-04
Geometric (Envelope)	gp1 (FFT)	ev1 (real cepstrum)	3.17E-06	1.47E-04

7.3 Optimized RBF using the Bees Algorithm

The first layer standard radial basis sets weights (w1) to inputs, and the first layer biases (b1) are all set to 0.8326/spread. Spread is 1.

The first layer biases are optimization variables. Variation range is defined between -7 to 7. The optimization fitness for problem is defined as

$$Fitness: \quad minimization[mse\_test] \tag{29}$$

Parameters of the Bees Algorithm are given in Table 5.

Table 5. Bees Algorithm Parameters

Bees Algorithm Parameters	Values
N	20
M	7
E	2
Nep	6
Nsp	2
Ngh	0.01

Optimization answer is -6.5917024. Thus, the first layer biases are -6.5917024. A comparison on train and test error between standard as well as optimum RBF is given in Table 6.

Table 6. Comparison between optimum RBF and standard RBF

State	Input features	mse_train	mse_test
Standard RBF	Lognormal1 (DCT) and gev3 (minimum phase)	9.89E-06	2.69E-05
Optimum RBF	Lognormal1 (DCT) and gev3 (minimum phase)	3.687E-10	1.896E-09

7.4 Comparison of proposed method with other methods in the literature

Table 7 compares the different methods in case of: the recognition accuracy, the used classifier and the used features, and the used number of features.

Table 7. Compares the different methods

Reference no.	Neural network	Feature	Number of features	Recognition accuracy (RA %)
[20]	HMM	Multi-Scale Fractal Dimension (MFD)	6	100
		Mel-frequency Cepstral Coefficient	13	99
[21]	SVM and ENN	Multi-Scale Fractal Dimension (MFD)	4	100
		Mel-frequency Cepstral Coefficient	11	100
[22]	SVM	Combination CoD[1] and InD[2]	3	99.579
		Combination FDF[3] and TDF[4]	3	100
[23]	FFNN	Statistical and Normal negative log-likelihood	10	100
[24]	Improved Fuzzy ARTMAP	Time domain, frequency domain and wavelet	10	99.57
[25]	BP	Statistical time domain	11	99.78
	IL[5]			99.84
	NCL[6]			99.81
	NCCE[7]			99.97
This paper	Proposed method	Lognormal1 (DCT) and gev3 (minimum phase)	2	100

[1] Represents the correlation dimension
[2] Represents the information dimension
[3] Represents the fractal dimension
[4] Represents the time-domain statistical feature
[5] Independent learning
[6] Negative correlation learning
[7] Neural network compact ensemble

8. Conclusion

In this work, we have investigated the problem of automatic bearing fault diagnosis using machine learning methods. These methods consist of feature extraction, feature selection, and classification. We concentrated on time, real cepstrum, minimum phase reconstruction, discrete cosine transform, discrete Fourier transform, envelope analysis signal and the Hilbert transform domains, and verified the pertinence of numerous features from descriptive statistics and maximum likelihood estimation; eventually eighteen features with better performance were chosen as input features to the classifier. We opted for using RBF as the classifier with the accuracy of 100% which is achieved in diagnosing bearing fault type.

Moreover, a method for optimizing radial basis train was presented. The method combines benefits of both radial basis function neural network and the Bees Algorithm to improve classifier accuracy. A comparison between the presented optimum train and the standard radial basis train showed that the presented method is more efficient in classifier accuracy exceptionally.

References

[1] Seera, M., Lim, CH. P., Nahavandi, S., Loo, CH. K., "Condition Monitoring of Induction Motors: A Review and an Application of an Ensemble of Hybrid Intelligent Models", Expert Systems with Applications, Vol. 41, No. 10, pp. 4891-4903, 2014.

[2] Thomas, M., and Fiabilite, 2003, "Maintenance Predictive", *et Vibration de Machines, Publications ETS, Montreal, Qc, Can*, p. 616.

[3] Archambault, J., Archambault, R. and Thomas, M., 2002, "Time domain descriptors for rolling-element bearing fault detection", *Proceedings of the 20th seminar on machinery vibration, CMVA, Québec*, p. 10.

[4] Thomas, M., Archambault, R., and Archambault, J., 2003, "Modified Julien index as a shock detector: its application to detect rolling element bearing defect", *Proceedings of the 21th seminar on machinery vibration, CMVA, Halifax (N.S.)*, pp. 21.1-21.12.

[5] Gluzman, D., 2000, "The use of log scales to analyse bearing failures", *J. Vibrations*, 16 (3), pp. 3-5.

[6] Randall, R. B., Antoni, J., "Rolling Element Bearing Diagnostics _ A Tutorial", Mechanical Systems and Signal Processing, Vol. 25, No. 2, pp. 485-520, 2011.

[7] Yan, R., Gao, R. X., "Multi-Scale Enveloping Spectrogram for Vibration Analysis in Bearing Defect Diagnosis", Tribology International, Vol. 42, No. 2, pp. 293-302, 2009.

[8] Loutas, T. H., Roulias, D., Pauly, E., Kostopoulos, V., "The Combined Use of Vibration, Acoustic Emission and Oil Debris On-line Monitoring Towards a More Effective Condition Monitoring of Rotating Machinery", Mechanical Systems and Signal Processing, Vol. 25, No. 4, pp. 1339-1352, 2011.

[9] Hamel, M., Addali, A., Mba, D., "Investigation of the Influence of Oil Film Thickness on Helical Gear Defect Detection Using Acoustic Emission", Applied Acoustic, Vol. 79, pp. 42-46, 2014.

[10] Safizadeh, M.S., Lakis, A.A., and Thomas, M., 2002, "Time-Frequency distributions and their Application to Machinery Fault Detection", *J. International Journal of Condition Monitoring and Diagnosis Engineering Management*, 5 (2), pp. 41-56.

[11] Khemili, I., Chouchane, M., "Detection of Rolling Element Bearing Defects by Adaptive Filtering", European Journal of Mechanics – A/Solids, Vol. 24, No. 2, pp. 293-303, 2005.

[12] Wang, H., Chen, P., "Intelligent Diagnosis Method for Rolling Element Bearing Faults Using Possibility Theory and Neural Network", Computers & Industrial Engineering, Vol. 60, No. 4, pp. 511-518, 2011.

[13] Castejon, C., Lara, O., Garcia-Prada J. C., "Automated Diagnosis of Rolling Bearings Using MRA and Neural Networks", Mechanical Systems and Signal Processing, Vol. 24, No. 1, pp. 289-299, 2010.

[14] Dong, G., Chen, J., "Noise Resistant Time Frequency Analysis and Application in Fault Diagnosis of Rolling Element Bearings", Mechanical Systems and Signal Processing, Vol. 33, pp. 212-236, 2012.

[15] Pham, D.T., Ghanbarzadeh, A., Koc, E., Otri, S., Rahim, S. and Zaidi, M., 2006, "The Bees Algorithm. A novel tool for complex optimization problems", *In Proceedings of the 2nd International Virtual Conference on Intelligent production machines and systems (IPROMS 2006)*, pp. 454–459 (Elsevier, Oxford), see also URL http://www.iproms.org/

[16] Samanta, B., and Al-Balushi, K.R., 2003, "Artificial neural network based fault diagnostics of rolling element bearings using time-domain features", *J. Mechanical Systems and Signal Processing*, 17 (2), pp. 317-328.

[17] Abbasion, S., Rafsanjani, A., Farshidianfar, A., and Irani, N., 2007, "Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine", *J. Mechanical Systems and Signal Processing*, 21, pp. 2933-2945.

[18] Sassi, S., Badri, B., and Thmoas, M., 2006, ""TALAF" and "THIKAT" as innovative time domain indicators for tracking BALL bearings", *Proceedings of the 24nd Seminar on machinery vibration, Canadian Machinery Vibration Association, éditeur M. Thomas, Montréal, Canada*, pp.404-419.

[19] Loparo, K.A., Bearings vibration data set, 2003, Case Western Reserve University, available from <http://www.eecs.cwru.edu/laboratory/bearing/download.htm>.

[20] Nelwamondo, F.V., Marwala, T., and Mahola, U., 2006, "Early classifications of bearing faults using hidden markov models, Gaussian mixture models, mel-frequency cepstral coefficients and fractals", *J. International Journal of Innovative Computing, Information and Control*, 2 (6), pp. 1281-1299.
[21] Marwala, T., and Vilakazi, C.B., 2007, "Computational intelligence for condition monitoring", *CoRR*.
[22] Yang, J., Zhang, Y., and Zhu, Y., 2007, "Intelligent fault diagnosis of rolling element bearing based on SVMs and fractal dimension", *J. Mechanical Systems and Signal Processing*, 21, pp. 2012-2024.
[23] Sreejith, B., Verma, A.K., and Srividya, A., 2008, "Fault diagnosis of rolling element bearing using time-domain features and neural network", *IEEE region 10 Colloquium and the ICIIS, Kharagpur, India,* No. 409.
[24] Xu, Z., Xuan, J., Shi, T., Wu, B., and Hu, Y., 2009, "A novel fault diagnosis method of bearing based on improved fuzzy ARTMAP and modified distance discriminant technique", *J. Expert Systems with Applications*, 36, pp. 11801-11807.
[25] Wang, Q., Zhang, Y., and Zhu, Y., 2010, "Neural network Compact Ensemble and Its Application", *Chinese Journal of Mechanical Engineering,* 23(2).