

1255. Health assessment and fault diagnosis for centrifugal pumps using Softmax regression

Jian Ma¹, Chen Lu², Wenjin Zhang³, Youning Tang⁴

^{1,2}Science and Technology on Reliability and Environmental Engineering Laboratory
Beijing, 100191, China

^{1,2,3}School of Reliability and Systems Engineering, Beihang University, Beijing, 100191, China

⁴Aeronautics Computing Technique Research Institute, Xi'an, 710000, China

³Corresponding author

E-mail: ¹majian3128@126.com, ²luchen@buaa.edu.cn, ³zwjok@buaa.edu.cn, ⁴yntang@aero-software.com

(Received 7 February 2014; received in revised form 2 April 2014; accepted 7 April 2014)

Abstract. Real-time health monitoring of industrial components and systems that can detect, classify, and predict impending faults is critical to reduce operating and maintenance costs. This paper presents a softmax regression-based prognostic method for on-line health assessment and fault diagnosis. System conditions are evaluated by processing the information gathered from access controllers or sensors mounted at different points in the system, and maintenance is performed only when the failure or malfunction prognosis is indicated. Wavelet packet decomposition and fast Fourier transform techniques are used to extract features from non-stationary vibration signals. Wavelet packet energies and fundamental frequency amplitude are used as features, and principal component analysis is used for feature reduction. Reduced features are input into softmax regression models to assess machine health and identify possible failure modes. The gradient descent method is used to determine the parameters of softmax regression models. The effectiveness and feasibility of the proposed method are illustrated by applying to a real application.

Keywords: health assessment, fault diagnosis, wavelet packet decomposition, PCA, softmax regression.

1. Introduction

Considerable efforts have been made to develop methods and tools for failure diagnosis. However, limited results have been provided on prognostics that can detect, analyze, and correct equipment problems before failures manifest, as well as provide system operators with a sufficient time window to schedule maintenance without disrupting the operations [1-4]. This paper presents a prognostic method for the health assessment and fault diagnosis of on-line centrifugal pumps using multiple softmax regression.

Realizing effective centrifugal pumps health assessment and fault diagnosis, one challenge is selecting a proper feature space that can reflect comprehensive performance [5]. Traditional time-domain and frequency-domain analysis are based on the assumption that the processing signals are stationary and linear. However, the vibration signal of worn centrifugal pumps is a nonlinear and non-stationary signal [6-8]. Wavelet packet decomposition has particular advantages for collecting abundant information with arbitrary time-frequency resolution. It allows extraction of features that combine nonlinear and non-stationary characteristics [9]. Thus, Wavelet packet decomposition is employed for feature extraction in this method.

The paper is organized as follows. Section 2 provides a state-of-the-art prognostic methodology and its related mathematics. Section 3 illustrates the test bed setup of a centrifugal pump and assessment/classification results obtained from an application of the proposed schemes on real data. Section 4 provides the summary and future research directions.

2. Methodology The prognostic scheme is based on monitored data that contain centrifugal pump incipient failure signatures and on intelligent mathematical techniques that can be

incorporated to detect and evaluate the risk of failure over a protracted period and classify which particular type of failure may occur.

2.1. Procedures of the methodology

The methodology has two major steps. First is the extraction and selection of features for health assessment and fault diagnosis. The search space is also reduced for fast computation. Second is the use of softmax regression for health assessment and root cause classification.

The process of the methodology can be summarized as Fig. 1.

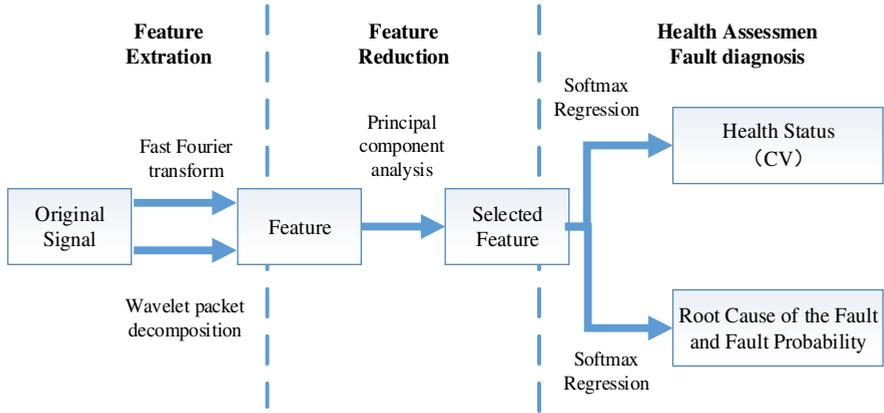


Fig. 1. The process of the methodology

2.2. Feature extraction using wavelet packet decomposition

The fault signals of a centrifugal pump are usually distributed in both high- and low-frequency bands where wavelet packet decomposition can reach a delicate degree. Wavelet packet analysis is a sophisticated method of orthogonal decomposition based on multi-resolution analysis, which can divide the full frequency band of signals in multi-level, so that each band signal contains elaborate information about the original signal. Therefore, wavelet packet decomposition is suitable for extracting both low- and high-frequency features [10-13]. The energy index of all bands reflecting signal characteristics can be constructed by statistically analyzing all bands of signals decomposed by wavelet packet. Determining the scale of wavelet packet decomposition is important. Very low wavelet packet decomposition scale reduces the efficiency of fault feature extraction, whereas very high wavelet packet decomposition scale increases the dimension of feature vector, consequently affecting the calculation rate [14]. Therefore, in centrifugal pump health assessment and fault classification, according to the vibration signal characteristics, the original signal $f(t)$ can be constructed by the sum of 8 components as Eq. (1), and eight frequency band energy indices $E_{3,j}$ can be calculated by three-layer decomposition:

$$f(t) = \sum_{l=1}^8 f_8^l(t), \tag{1}$$

where $f_8^j(t)$ is the wavelet packet component signal that can be expressed by a linear combination of wavelet packet functions as:

$$f_8^j(t) = \sum_{k=-\infty}^{\infty} S_{8,k}^j \psi_{8,k}^j(t), \tag{2}$$

where j and k are integers and defined as the modulation, scale and translation parameter, respectively; $S_{8,k}^j$ is the wavelet packet coefficients, $\psi_{8,k}^j(t)$ is wavelet packet function:

$$E_{3,j} = \int |S_{3,j}(t)|^2 dt = \sum_{k=1}^n |x_{j,k}|^2, \tag{3}$$

where $x_{j,k}$ ($j = 0, 1, \dots, 7; k = 1, 2, \dots, n$) stands for the amplitude of reconstruction signal $S_{3,j}$.

When a centrifugal pump has a fault, the energy of each band signal will have a great impact. Thus, the energy should be normalized into a feature vector T :

$$T = [E_{3,0}, E_{3,1}, E_{3,2}, E_{3,3}, E_{3,4}, E_{3,5}, E_{3,6}, E_{3,7}] \cdot \left(\sum_{j=0}^7 |E_{3,j}|^2 \right)^{-1/2}. \tag{4}$$

2.3. PCA-based feature reduction

The PCA procedure employed in this paper will be briefly presented for a more comprehensive data. At a given centrifugal pump operation stage S (in this paper, only the normal machine behavior and machine operation with a worn bearing/impeller are considered), the signal features X are characterized by the multivariate Gaussian distribution with mean $\bar{\mu}_s$ and the covariance matrix K_s . The symmetric matrix K_s can now be represented as:

$$K_s = \sum_{i=1}^r \lambda_i \bar{v}_i \bar{v}_i^T = V \Lambda V^T, \tag{5}$$

where r is the rank of the covariance matrix K_s , $\lambda_i, i = 1, 2, \dots, r$ are the non-zero eigenvalues of K_s , \bar{v}_i are the corresponding unit norm eigenvectors, and:

$$V = [\bar{v}_1 \quad \bar{v}_2 \quad \dots \quad \bar{v}_r], \quad \Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_r \end{bmatrix}. \tag{6}$$

All eigenvalues of K_s are real and greater than or equal to zero because of its positive semi-definiteness. Each eigenvalue $\lambda_i, i = 1, 2, \dots, r$ depicts the amount of the covariance matrix energy projected in the direction of the corresponding eigenvector \bar{v}_i [15]. Only a few of the eigenvalues in Λ account for most of the energy in the covariance matrix K_s when a high degree of correlation exists among the components of X . Thus, assuming that eigenvalues $\lambda_i, i = 1, 2, \dots, r$ are arranged in descending order, Eq. (5) can be represented as:

$$K_s = \sum_{i=1}^p \lambda_i \bar{v}_i \bar{v}_i^T = V_p \Lambda_p V_p^T, \tag{7}$$

where:

$$V = [\bar{v}_1 \quad \bar{v}_2 \quad \dots \quad \bar{v}_p], \quad \Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_p \end{bmatrix}, \tag{8}$$

where p is the number of the principal components of K_s , λ_i , $i = 1, 2, \dots, r$ are the largest p eigenvalues of K_s , and \bar{v}_i are the corresponding unit norm eigenvectors.

A query item \tilde{X} can now be transformed into a p component random variable \tilde{Y} given as:

$$\tilde{Y} = T(\tilde{X} - \bar{\mu}_s), \tag{9}$$

$$T = \Lambda_p^{-1/2} \Lambda_p. \tag{10}$$

If \tilde{X} belongs to the class of signals from the centrifugal pump state S , then \tilde{Y} should be normally distributed with zero mean and variance I_p , where I_p is the unity matrix of order p . Thus, for each query item \tilde{X} , its adherence to the class S can be assessed through the Euclidean norm of the vector \tilde{Y} , which in turn corresponds to assessment and classification based on the softmax regression of the query item from the training classes [16].

2.4. Logistic regression method

The machine condition description from daily maintenance records/logs is a dichotomous problem (either normal or failed) that can be represented using an LR function [1]. The goal of logistic regression (LR) is to find the best fitting model to describe the relationship between the categorical characteristics of dependent variables (the probability of an event, constrained between 0 and 1) and independent variables. The logistic function is:

$$prob(event) = p(x) = (1 + e^{-g(x)})^{-1}. \tag{11}$$

The logistic or logit model is:

$$Logit = g(x) = \log(p(x)(1 - p(x))^{-1}) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k, \tag{12}$$

where $g(x)$ is a linear combination of the independent variables x_1, x_2, \dots, x_k .

The precondition for figuring out $P(x)$ is determining parameters α and β_1, \dots, β_k in advance. Dichotomous-dependent variables make estimation using ordinary least squares inappropriate [17]. Hence, estimation in LR chooses the parameters α and β_1, \dots, β_k using the maximum likelihood method rather than those that minimize the sum of squared errors [1]. Then, the probability of failure for each input vector x can be calculated according to Eq. (11).

2.5. Softmax regression method

In the softmax regression setting, multi-class classification (as opposed to only binary classification) is interested, and so the label y can take on k different values, rather than only two. Thus, in training set $\{(x^{(1)}, y^{(1)}), \dots, (x^{(m)}, y^{(m)})\}$, we now have that $y^{(i)} \in \{1, 2, \dots, k\}$.

Given a test input x , the hypothesis to estimate the probability that $p(y = j|x)$ for each value of $j = 1, 2, \dots, k$ is wanted. I.e., the probability of the class label taking on each of the k different possible values is estimated. Thus, the hypothesis will output a k dimensional vector (whose elements sum to 1) giving k estimated probabilities [18].

Concretely, our hypothesis $h_\theta(x)$ takes the form:

$$h_\theta(x^{(i)}) = \begin{bmatrix} p(y^{(i)} = 1|x^{(i)}; \theta) \\ p(y^{(i)} = 2|x^{(i)}; \theta) \\ \vdots \\ p(y^{(i)} = k|x^{(i)}; \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^k e^{\theta_j^T x^{(i)}}} \begin{bmatrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ \vdots \\ e^{\theta_k^T x^{(i)}} \end{bmatrix}. \tag{13}$$

Here $\theta_1, \theta_2, \dots, \theta_k \in \mathfrak{R}^{n+1}$ are the parameters of the softmax regression model. Notice that the term $1/\sum_{j=1}^k e^{\theta_j^T x^{(i)}}$ normalizes the distribution, so that it sums to one.

In the special case where $k = 2$, softmax regression reduces to logistic regression. This shows that softmax regression is a generalization of logistic regression. The cost function used for softmax regression is:

$$J(\theta) = -\frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^k 1\{y^{(i)} = j\} \log \frac{e^{\theta_j^T x^{(i)}}}{\sum_{l=1}^k e^{\theta_l^T x^{(i)}}} \right], \quad (14)$$

where $1\{\cdot\}$ is the indicator function, so that $1\{\text{a true statement}\} = 1$, and $1\{\text{a false statement}\} = 0$.

There is no known closed-form way to solve for the minimum of $J(\theta)$, and thus as usual well resort to an iterative optimization algorithm such as gradient descent or L-BFGS. Taking derivatives, the gradient is:

$$\nabla_{\theta_j} J(\theta) = -\frac{1}{m} \sum_{i=1}^m \left[x^{(i)} \left(1\{y^{(i)} = j\} - p(y^{(i)} = j | x^{(i)}; \theta) \right) \right]. \quad (15)$$

Softmax regression has an unusual property that it has a “redundant” set of parameters. Adding weight decay to the cost function is needed. This will take care of the numerical problems associated with softmax regression's overparameterized representation.

The modified cost function is:

$$J(\theta) = -\frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^k 1\{y^{(i)} = j\} \log \frac{e^{\theta_j^T x^{(i)}}}{\sum_{l=1}^k e^{\theta_l^T x^{(i)}}} \right] + \frac{\lambda}{2} \sum_{i=1}^k \sum_{j=0}^n \theta_{ij}^2. \quad (16)$$

The derivative of the modified cost function is:

$$\nabla_{\theta_j} J(\theta) = -\frac{1}{m} \sum_{i=1}^m \left[x^{(i)} \left(1\{y^{(i)} = j\} - p(y^{(i)} = j | x^{(i)}; \theta) \right) \right] + \lambda \theta_j. \quad (17)$$

By minimizing $J(\theta)$ with respect to θ , a working implementation of softmax regression is given.

3. Experimental result

The methodology was implemented in a centrifugal pump (Fig. 2) to evaluate dynamic health condition. In addition, fault mode analysis was performed to identify the possible root cause. The centrifugal pump is driven by a motor with stabilized speed of 2900 r/min. Four commonly occurring faults of the centrifugal pump were set, namely, bearing roller wearing, bearing inner race wearing, bearing outer race wearing and centrifugal pump impeller wearing.

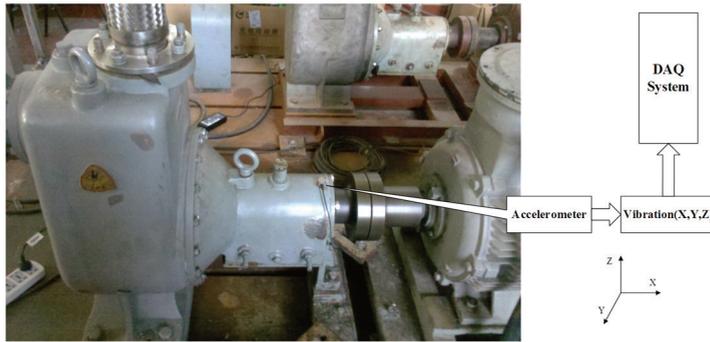


Fig. 2. Centrifugal pump data acquisition system

3.1. Data acquisition system description

Three vibration signals were acquired from an installed accelerometer, with a sampling rate of 10.24k Hz.

3.2. Feature extraction and reduction

FFT was used for each vibration signal to obtain the fundamental frequency amplitude. Three-level wavelet packet decomposition using Daubechies wavelet (DB10) was adopted for each vibration signal, and fundamental frequency amplitude and packet energies were used as features. A subset of feature components was determined using PCA. In this case, a four-dimensional feature vector was finally selected as a feature vector for health assessment and fault mode classification after feature reduction.

3.3. Softmax regression models training

- Softmax regression model trained for health assessment.

When Softmax regression model was trained for health assessment, softmax regression reduced to logistic regression. A total of 120 sets of data were used as training data, including 40 sets of data sampled under normal conditions [$P(x) = 0$] versus 80 sets of fault data [$P(x) = 1$]. The parameters α and β_1, \dots, β_4 were estimated using the maximum likelihood method to eventually obtain the model for performance assessment as softmax regression model 1.

- Softmax regression model trained for fault diagnosis.

Four fault modes of the centrifugal pump which include bearing roller wearing, bearing inner race wearing, bearing outer race wearing and centrifugal pump impeller wearing, were considered. Four sets of centrifugal pump vibration data (one set for each fault mode of the centrifugal pump) were used for training centrifugal pump fault diagnosis model based on softmax regression.

Set 1 (bearing roller wearing): 40 subsets of bearing roller wearing data ($p(y = 1|x) = 1$) versus 40 subsets of non-wearing data ($p(y = 1|x) = 0$);

Set 2 (bearing inner race wearing): 40 subsets of bearing inner race wearing data ($p(y = 2|x) = 1$) versus 40 subsets of non-wearing data ($p(y = 2|x) = 0$);

Set 3 (bearing outer race wearing): 40 subsets of bearing outer race wearing data ($p(y = 3|x) = 1$) versus 40 subsets of non-wearing data ($p(y = 3|x) = 0$);

Set 4 (centrifugal pump impeller wearing): 40 subsets of centrifugal pump impeller wearing data ($p(y = 4|x) = 1$) versus 40 subsets of non-wearing data ($p(y = 4|x) = 0$).

The parameters $\theta_1, \theta_2, \dots, \theta_k$ were estimated using gradient descent method to eventually obtain the model for fault diagnosis as softmax regression model 2.

3.4. Validation

Four sets of centrifugal pump vibration data (one set for each fault mode) were used for validation of centrifugal pump health assessment and fault diagnosis models based on softmax regression.

- Set 1 (bearing roller wearing): 10 data under normal condition versus 40 bearing roller wearing data;
- Set 2 (bearing inner race wearing): 10 data under normal condition versus 32 bearing inner race wearing data;
- Set 3 (bearing outer race wearing): 10 data under normal condition versus 40 bearing outer race wearing data;
- Set 4 (centrifugal pump impeller wearing): 10 data under normal condition versus 30 impeller wearing data.

FFT and three-level wavelet packet decomposition using Daubechies wavelet (DB10) were adopted for extracting features from four sets of validation data. The wavelet package energies (E0 to E7) extracted from data set 1 to 4 is respectively show in Figs. 3-6. The fundamental frequency amplitude (or ffa, for short) extracted from data set 1 to 4 (ffa 1, ffa 2, ffa 3 and ffa 4) is show in Fig. 7.

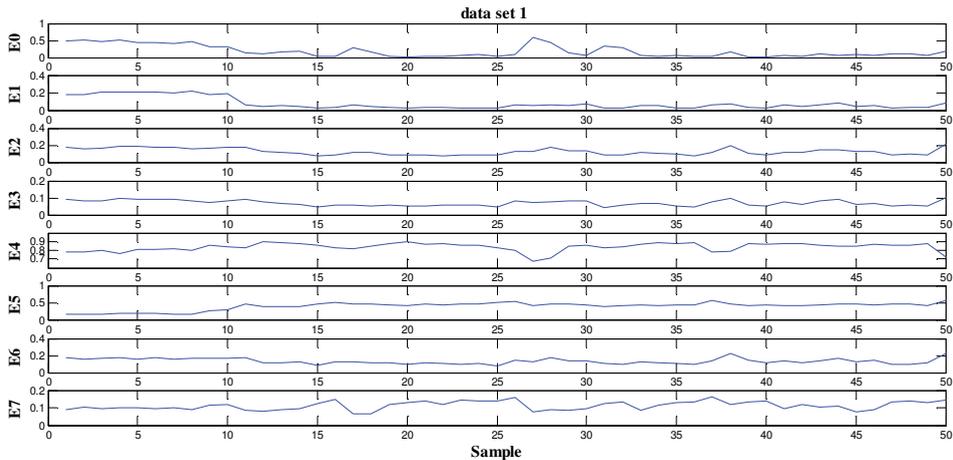


Fig. 3. The wavelet package energies extracted from data set 1

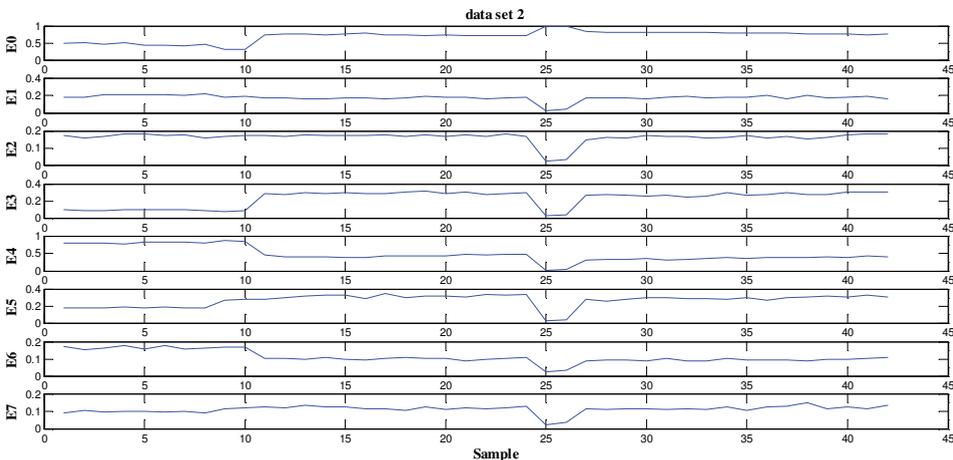


Fig. 4. The wavelet package energies extracted from data set 2

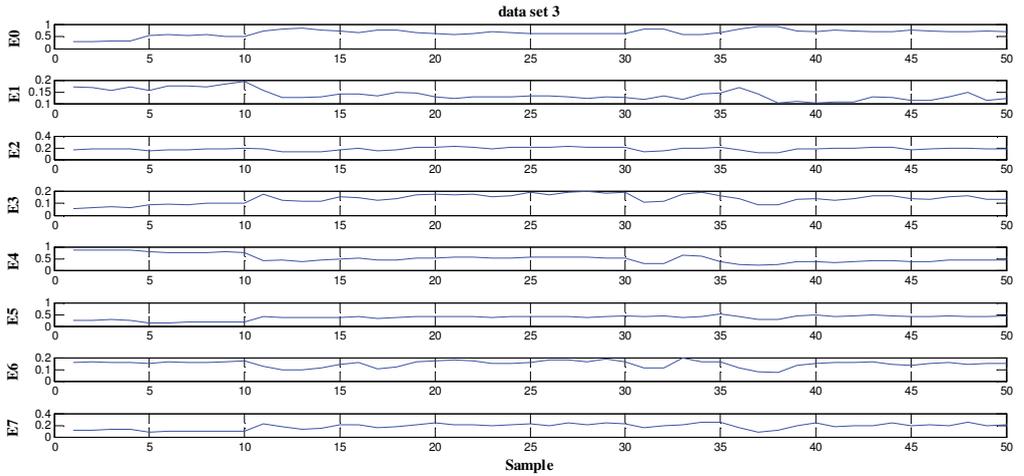


Fig. 5. The wavelet package energies extracted from data set 3

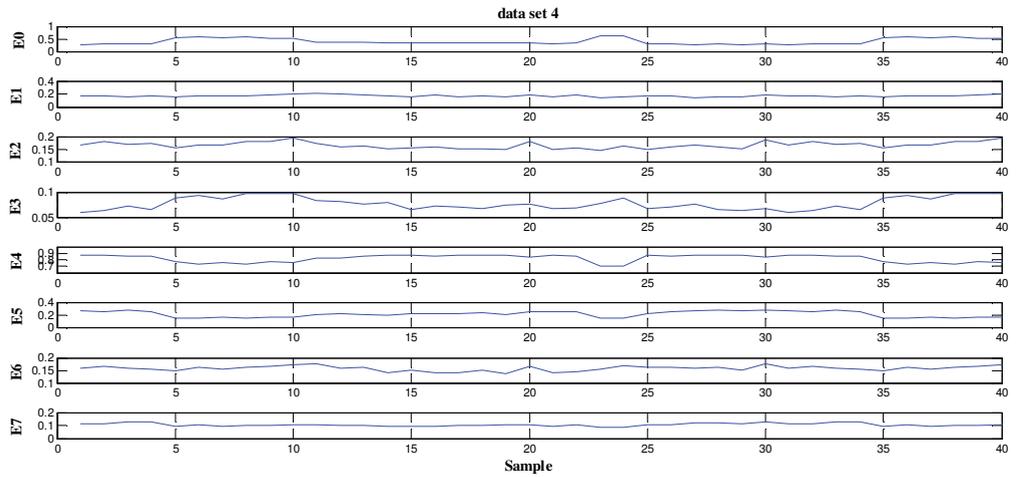


Fig. 6. The wavelet package energies extracted from data set 4

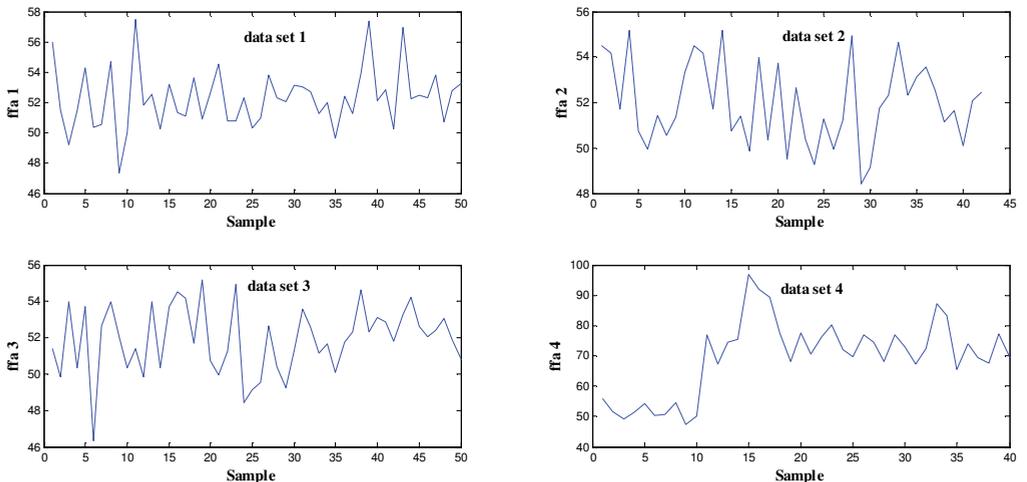


Fig. 7. The fundamental frequency amplitude extracted from data set 1 to 4

Reduced features are input into softmax regression models to assess the centrifugal pump health condition and identify possible failure modes. The confidence value (CV) was calculated based on the probability of failure. Define $CV = 1 - P(x)$. When the centrifugal pump operates normally, CV is close to 1; if the centrifugal pump is going to fail, CV is approaching 0 correspondingly; If the confidence value is less than a predetermined threshold (e.g., 0.8), the root fault diagnosis module will be triggered, and features are input into fault diagnosis models to calculate the probability of each fault.

Fig. 8(a)-(d) show the overall health assessment of the four sets of centrifugal pump vibration data using model softmax regression model 1. The probability of different fault modes conducted from softmax regression model 2 is shown in Fig. 9(a)-(d).

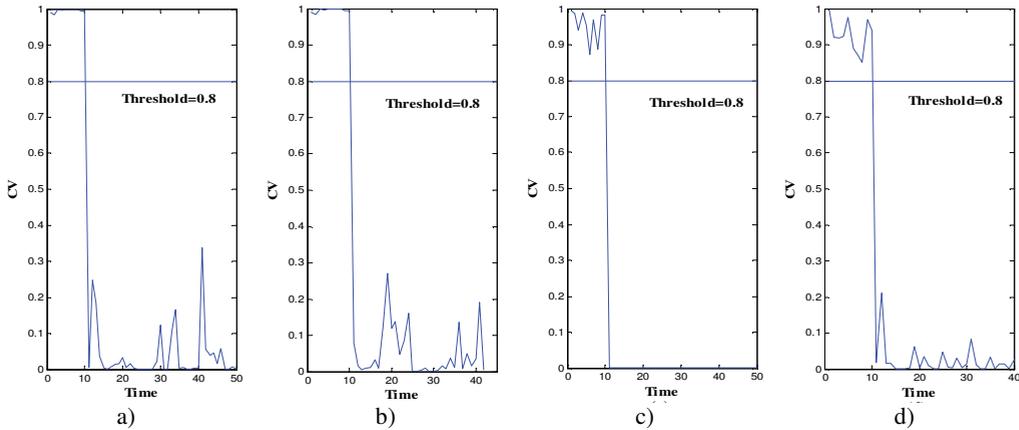


Fig. 8. Health assessment result of four fault modes

In Fig. 8, both bearing and impeller problems can be detected from the CV drops. However, the difference among the four drops and their cause are difficult to clarify. In this methodology, the fault diagnosis module is triggered as long as the confidence value is below a predetermined threshold (0.8) by inputting the corresponding features into the trained models (the softmax regression model 2) to calculate the probability of fault modes. From time 10, the probability of fault mode 1 (bearing roller wearing, $p1$), fault mode 2 (bearing inner race wearing, $p2$), fault mode 3 (bearing outer race wearing, $p3$), and fault mode 4 (impeller wearing, $p4$) is very high [solid line in Fig. 9(a)-(d), respectively]. Consequently, the minor probability of the failure of these points is indicated in Fig. 9.

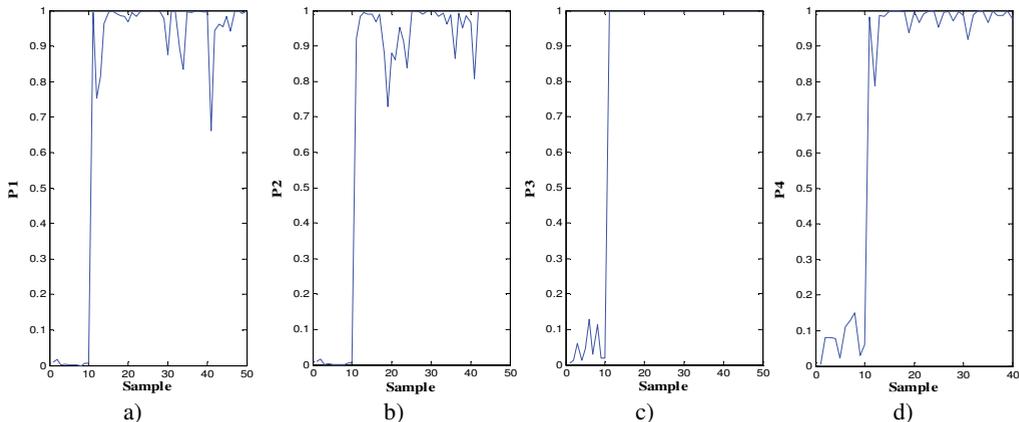


Fig. 9. Probability of fault modes 1, 2, 3, and 4

In conclusion, a softmax classifier is more suitable for centrifugal pump fault diagnosis than 4 separate binary classifiers using logistic regression. The four fault modes considered in this paper are mutually exclusive, so softmax regression classifier would be appropriate. Besides that, softmax regression is more simple, useful and easy to implement.

4. Conclusions

A softmax regression-based approach for centrifugal pump health assessment and root cause classification is presented in this paper. Softmax regression combined with the gradient descent method is an effective and efficient tool for dynamic health assessment and root cause classification. WPT combined with PCA is a suitable feature extraction step where appropriate features can be obtained from non-stationary signals. The method is generic and shows promising results for analyzing both stationary and non-stationary signals. Thus, it could be applied to other centrifugal pumps.

However, four types of centrifugal pump fault modes are considered in this paper. More fault modes should be taken into consideration for better health assessment and fault diagnosis of centrifugal pumps. When the process is not time shifted, the coefficients of WPT can be directly used as features instead of using packet energy, which will be investigated further in future research and applications.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant Nos. 61074083, 50705005 and 51105019), the Technology Foundation Program of National Defense (Grant No. Z132013B002), as well as the Innovation Foundation of BUAA for PhD Graduates.

References

- [1] **Yan J., Lee J.** Machine degradation assessment and root cause classification using logistic regression method. *ASME Journal of Manufacturing Science and Engineering*, Vol. 127, 2005, p. 912-914.
- [2] **Yan J., Lee J., Koc M.** Predictive algorithm for machine degradation detection using logistic regression. *Fifth International Conference on Managing Innovations in Manufacturing*, Milwaukee, 2002, p. 172-178.
- [3] **Kacprzynski G. J., Roemer M. J.** Health management strategies for 21st century condition-based maintenance systems. *International Comadem Congress*, Houston, TX, 2000.
- [4] **Liao L. X., Lee J. J.** Design of a reconfigurable prognostics platform for machine tools. *Expert Systems with Applications*, Vol. 37, 2010, p. 240-252.
- [5] **Pan Y., Chen J., Li X.** Bearing performance degradation assessment based on lifting wavelet packet decomposition and fuzzy c-means. *Mech. Syst. Sign. Proc.*, Vol. 24, 2010, p. 559-566.
- [6] **Soylemezoglu A., Jagannathan S., Saygin C.** Mahalanobis-Taguchi system as a multi-sensor based decision making prognostics tool for centrifugal pump failures. *Reliability, IEEE Transactions on* 2011, Vol. 60, 2005, p. 864-878.
- [7] **Ahonen T., Tiainen R., Viholainen J., Ahola J., Kestila J.** Pump operation monitoring applying frequency converter. *Power Electronics, Electrical Drives, Automation and Motion*, 2008, p. 184-189.
- [8] **Wu X. W.** Vibration faults diagnosis for centrifugal ventilator based on DDAGSVM. *Instrumentation & Measurement, Sensor Network and Automation*, Vol. 1, 2012, p. 318-321.
- [9] **Li Z. X., Yan X. P., Yuan C. Q., Peng Z. X., Li L.** Virtual prototype and experimental research on gear multi-fault diagnosis using wavelet-autoregressive model and principal component analysis method. *Mech. Syst. Sign. Proc.*, Vol. 25, 2011, p. 2589-2607.
- [10] **Jafar Z., Javad P.** Bearing fault detection using wavelet packet transform of induction motor stator current. *Tribology International*, Vol. 40, 2007, p. 763-769.
- [11] **Li J., Jiang P. F., Xiang Y. Y., Ti J. W.** Experimental investigation for fault diagnosis based on a hybrid approach using wavelet packet and support vector classification. *Scientific World Journal*, Vol. 10, 2014, p. 1155-1160.

- [12] **Wang X., Liu. C. W., Bi F. R., Bi X. Y., Shao K.** Fault diagnosis of diesel engine based on adaptive wavelet packets and EEMD-fractal dimension. *Echanical Systems and Signal Processing*, Vol. 41, 2013, p. 581-597.
- [13] **Zhang Z. Y., Wang Y., Wang K. S.** Fault diagnosis and prognosis using wavelet packet decomposition, Fourier transform and artificial neural network. *Journal of Intelligent Manufacturing*, Vol. 24, 2013, p. 1213-1227.
- [14] **Yen G. G., Lin K. C.** Wavelet packet feature extraction for vibration monitoring. *IEEE Trans. Ind. Electron.*, Vol. 47, 2000, p. 650-667.
- [15] **Pirra M., Gandino E., Torri A., Garibaldi L., Machorro-López J. M.** PCA algorithm for detection, localisation and evolution of damages in gearbox bearings. *Journal of Physics: Conference Series*, 2011, p. 305.
- [16] **Upadhyaya S.** Baker-Demaray. Comparison of NN and LR classifiers in the context of screening native American elders with diabetes. *Expert Systems with Applications*, Vol. 40, 2013, p. 5830-5838.
- [17] **Djurdjanovic D., Ni J., Lee J.** Time-frequency based sensor fusion in the assessment and monitoring of machine performance degradation. *ASME International Mechanical Engineering Congress and Exposition*, 2002.
- [18] **D'Ambrosio R., Iannello G., Soda P.** Solving biomedical classification tasks by softmax reconstruction in ECOC framework. *Computer-Based Medical Systems*, 2013, p. 433-436.