Research Article

Multiscale Permutation Entropy Based Rolling Bearing Fault Diagnosis

Jinde Zheng,[1,2] Junsheng Cheng,[1,2] and Yu Yang[1,2]

[1] *State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body, Hunan University, Changsha 410082, China*
[2] *College of Mechanical and Vehicle Engineering, Hunan University, Changsha 410082, China*

Correspondence should be addressed to Jinde Zheng; lqdlzheng@126.com

A new rolling bearing fault diagnosis approach based on multiscale permutation entropy (MPE), Laplacian score (LS), and support vector machines (SVMs) is proposed in this paper. Permutation entropy (PE) was recently proposed and defined to measure the randomicity and detect dynamical changes of time series. However, for the complexity of mechanical systems, the randomicity and dynamic changes of the vibration signal will exist in different scales. Thus, the definition of MPE is introduced and employed to extract the nonlinear fault characteristics from the bearing vibration signal in different scales. Besides, the SVM is utilized to accomplish the fault feature classification to fulfill diagnostic procedure automatically. Meanwhile, in order to avoid a high dimension of features, the Laplacian score (LS) is used to refine the feature vector by ranking the features according to their importance and correlations with the main fault information. Finally, the rolling bearing fault diagnosis method based on MPE, LS, and SVM is proposed and applied to the experimental data. The experimental data analysis results indicate that the proposed method could identify the fault categories effectively.

1. Introduction

The vibration signals of mechanical systems, especially for ones with fault, often show mutation, nonlinearity, and nonstationarity because of the strike, velocity chopping, structure transmutation, loading, and friction. Hence, it is very crucial for mechanical fault diagnosis to extract the fault feature information from the nonlinear and nonstationary signal. A primary method for dealing with the nonlinear and nonstationary signal is time-frequency analysis [1], which has been applied to the mechanical fault diagnosis field widely for its ability to provide local information both in time and frequency domains of vibration signals [2]. However, the time-frequency analysis method, such as wavelet transform or Hilbert-Huang transform [3, 4], which decomposes the vibration signal into several stationary monocomponent signals, cannot reflect the subtle dynamic changes of vibration signal effectively and, therefore, inevitably will have some limitations [5].

With the development of nonlinear dynamic theories, especially in recent years, a number of nonlinear parameters and methods, such as chaos theory, fractal dimension, and information entropy, have been applied to machine condition monitoring and fault diagnosis. For instance, Logan and Mathew elaborated the application of the correlation dimension to vibration fault diagnosis of rolling element bearing [6]; Jiang et al. used the correlation dimension in gearbox condition monitoring [7]. However, reliable estimation of correlation dimension requires very long datasets, which might be difficult or even impossible to be achieved especially in online, real-time monitoring and diagnosis [5]. Lately, approximate entropy (ApEn) was introduced and selected as a tool for rolling bearing health monitoring by Yan and Gao [5]. Unfortunately, the estimation of ApEn depends heavily on the data length, and the estimated value is uniformly lower than the expected one, especially for a short dataset, and lacks relative coherence as well [8, 9]. In order to overcome the shortcomings of ApEn, the sample entropy (SampEn) was proposed by Richman and Moorman [9, 10]. However, ApEn and SampEn both measure the complexity of time series in a single scale. Based on SampEn, multiscale entropy (MSE) was

introduced by Costa as an enhanced approach to evaluate the complexity of complex time series in different scales [11, 12]. MSE has been recently utilized to extract the fault feature information from rolling bearing vibration signal by Zhang et al. [8]. However, the SampEn estimation will be affected by the nonstationarity, outliers, and artifacts of time series, which changes the standard deviation of time series and similarity criterion [13] and hence will cause a bad estimation of MSE. In addition, the computations of MSE are also very time-consuming, especially for a very long time series.

Recently, permutation entropy (PE) was proposed by Bandt and Pompe [14, 15] for measuring the randomicity and detecting the dynamic changes of time series. Compared with the parameters mentioned above, the computation of PE is simple, immune to noise, and suitable for online monitoring. Recently, Yan and Liu [16] viewed PE as a tool for status characterization of rotary machines and their research indicates that PE could effectively detect and amplify the dynamic change of rolling bearing vibration signals. Nicolaou and Georgiou [17] used PE and SVMs to detect the epileptic electroencephalogram. Their findings indicate that the low computational complexity of PE makes it a highly favorable feature to be employed as part of a system for real-time automated seizure detection.

However, like traditional single scale nonlinear dynamic parameters ApEn and SampEn, PE detects the dynamic changes and randomness of time series only in a single scale. Recently, multiscale permutation entropy (MPE) was introduced by Aziz and Arif in the literature [13] to measure the complexity of time series in different scales and is compared with MSE through analyzing the physiological time series and the results show that MPE is more robust than MSE in analyzing the presence of artifacts and white Gaussian noise.

As the vibration signals collected from normal rolling bearing are random and irregular, the randomness and the dynamic behavior of the vibration signal will change abruptly when the rolling bearing of equipment works under a bad condition. Due to the complexity of mechanical system, the vibration signal is much more complex and contains much more important information in different scales. Hence, MPE is employed to detect the dynamic changes and fault features from the rolling bearing vibration signal.

In the paper, firstly the PE values with different scales are served as initial feature parameters to extract fault feature information from the bearing vibration signal. Since the feature vector concludes MPE values in different scales, which will lead to a high dimension and information redundancy, and it is also difficult to find out the features containing the main fault information, in this paper the LS proposed by He et al. [18] is employed to refine the feature vector and rank the feature values according to their importance. Then the several most important features are reconstructed as the new feature vector for the SVM training and testing. Next, naturally, a multifault classifier needs to be constructed to fulfill the diagnostic procedure automatically. As support vector machine (SVM) has the merits of suitability for small sample data classification and fast training, in this paper, SVM is adopted to construct the multifault classifier [19–21].

The rest of the paper is organized as follows. In the second section, the definitions of PE and MPE are introduced, respectively. In the third section, the Laplacian score (LS) is introduced firstly, and then a new rolling bearing fault diagnosis method based on MPE, LS, and SVM is proposed. In the fourth section, the proposed method is applied to rolling bearing experimental data and some comparisons are made. Finally, the fifth section concludes the paper.

2. Algorithms of PE and MPE

2.1. Algorithm of PE. Permutation entropy (PE) was introduced recently to detect dynamic changes of time series by Bandt and Pompe [14, 15], which is based on comparison of neighboring values and therefore has the advantages of simple computation, less calculating amounts and time. Besides, it has been verified that, similar to Lyapunov exponents, PE is particularly useful and robust in the presence of dynamic or observational noise [22], and its algorithm is described as follows.

Consider a time series, $\{x(i), \ i = 1, 2, \ldots, N\}$, with length N, which can be reconstructed as

$$X(1) = \{x(1), x(1+\lambda), \ldots, x(1+(m-1)\lambda)\}$$

$$\vdots$$

$$X(i) = \{x(i), x(i+\lambda), \ldots, x(i+(m-1)\lambda)\}$$

$$\vdots \tag{1}$$

$$X(N-(m-1)\lambda) = \{x(N-(m-1)\lambda),$$
$$x(N-(m-2)\lambda), \ldots, x(N)\},$$

where m is the embedding dimension and λ is the time delay. As described in [16, 22], for a given but arbitrary i, the m real values $\{x(i), x(i+\lambda), \ldots, x(i+(m-1)\lambda)\}$, contained in each $X(i)$, can be rearranged in an increasing order as

$$X(i) = \{x(i+(j_1-1)\lambda) \le x(i+(j_2-1)\lambda)$$
$$\le \cdots \le x(i+(j_m-1)\lambda)\}. \tag{2}$$

If there exist two elements in $X(i)$ that have the same value, for example, $x(i+(j_{i1}-1)\lambda) = x(i+(j_{i2}-1)\lambda)$, then we order the quantities x according to the values of their corresponding j's; namely, if $j_{i1} < j_{i2}$, then $x(i+(j_{i1}-1)\lambda) \le x(i+(j_{i2}-1)\lambda)$ is written. Accordingly, any vector $X(i)$ can be mapped onto a group of symbols as

$$S(g) = [j_1, j_2, \ldots, j_m], \tag{3}$$

where $g = 1, 2, \ldots, k, \ k \le m!$. $m!$ is the largest number of distinct symbols and $S(g)$ is one of the m! permutations of m distinct symbols, which is mapped onto the m number symbols $(j_1, j_2, \ldots, j_m)$ in m-dimensional embedding space. When each such permutation is considered as a symbol, then the reconstructed trajectory in the m-dimensional space is represented by a symbol sequence [22].

Therefore, if we suppose that the probability distribution for the distinct symbols be as $P_1, P_2, \ldots, P_k$, $\sum_{g=1}^{k} P_g = 1$, where $k \leq m!$, then the PE for the time series $\{x(i), i = 1, 2, \ldots, N\}$ can be defined as the Shannon entropy for the k distinct symbols:

$$H_p(m) = -\sum_{g=1}^{k} P_g \ln P_g. \tag{4}$$

It is noticed that $H_p(m)$ attains the maximum value, $\ln(m!)$, when $P_j = 1/m!$. For convenience, $H_p(m)$ can be normalized by $\ln(m!)$ as

$$H_p = \frac{H_p(m)}{\ln(m!)}. \tag{5}$$

Obviously, $0 \leq H_p \leq 1$. A smaller value of H_p indicates that the time series is much more regular and the smallest value of H_p (zero) means that the time series is very regular as the periodic signal. And a larger H_p means a much more random time series and the largest possible value of H_p (one) is realized when all permutations have equal probability, as is in the case of white noise [16]. Therefore, PE is a very suitable tool for describing local order structure and amplifying the dynamic changes of time series.

There are three parameters to be considered in the calculation of PE, namely, the length of time series N, embedding dimension, and time delay λ. Bandt recommended $m = 3 \sim 7$. However, in the following research we will find that $m = 6$ seems to be the most suitable. In order to investigate the effect of N and m on computation of PE, five Gaussian white noise signals, respectively, with lengths 128, 256, 512, 1024, and 2048, are under our consideration. For convenience, their PE values are denoted by PE_1, PE_2, PE_3, PE_4, and PE_5. Figure 1 shows their PE relationships with different m and N when $\lambda = 1$.

As the Gaussian white noise signal is random and it should have an estimated value close to 1, therefore when N is less than 2048, m should be no more than 7 (where the estimated PE is smaller than 0.9). From Figure 1 it can be found that the difference between PE_4 with length 1024 and PE_3 with length 512 is only 0.0659 when $m = 6$. Hence when $m = 6$, $N > 512$ is sufficient for PE calculation.

In addition, the time delay λ has a little effect on the estimation of PE. Take the Gaussian white noise signal with length 512 as an instance. Its PE is shown in Figure 2 with λ ranging from 1 to 6 in different embedding dimensions (m ranging from 2 to 8). And from Figure 2 it can be found that there are very small differences among the PEs between different time delays. Therefore, in this paper, we set $\lambda = 1$.

2.2. Calculation of MPE. Multiscale permutation entropy (MPE) is defined as the PE set of time series in different scales and is calculated as follows [13].
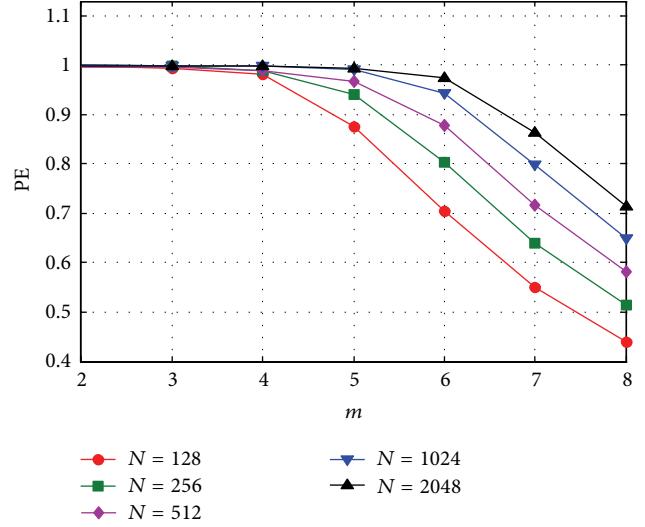

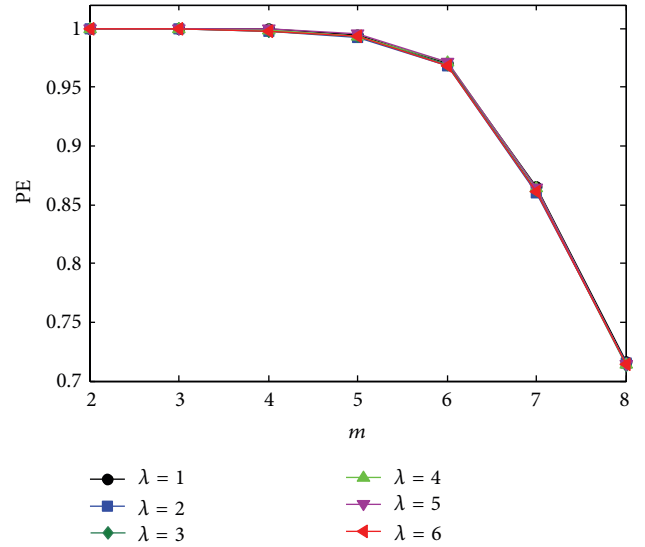FIGURE 1: The PEs of white Gaussian noise signals with different lengths when $\lambda = 1$.

FIGURE 2: The PEs of white Gaussian noise with different time delays.

(1) Consider the time series, $\{x(i), i = 1, 2, \ldots, N\}$, which can be divided into several coarse-grained time series $\{y_j^{(\tau)}\}$ in the form as

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, \quad j = 1, 2, \ldots, \frac{N}{\tau}, \tag{6}$$

where $\tau = 1, 2, \ldots$ is the scale factor. Obviously, when $\tau = 1$, $y_j^{(1)}$ $(j = 1, 2, \ldots, N)$ is the original time series $\{x(i), i = 1, 2, \ldots, N\}$. When $\tau > 1$, $y_j^{(\tau)}$ $(j = 1, 2, \ldots, N)$ is a coarse grained time series with length N/τ.

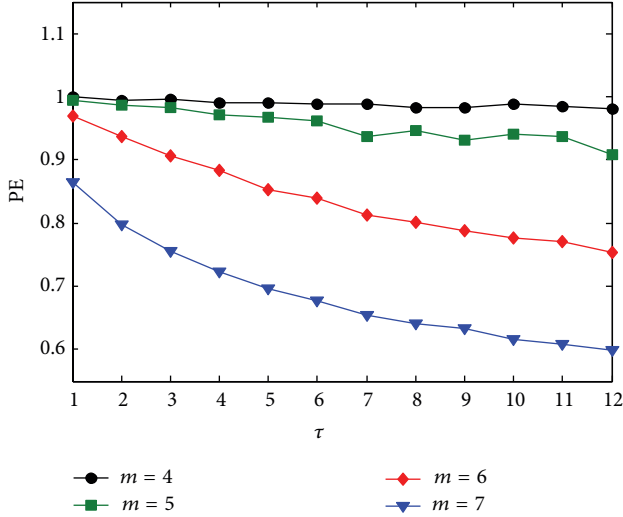(2) Calculate PE of each coarse-grained time series $y_j^{(\tau)}$ $(j = 1, 2, \ldots, N)$ under the same parameters, and

FIGURE 3: The MPE of Gaussian white noise signal with different embedding dimensions. Here $\tau_{\max} = 12$ and $\lambda = 1$.

FIGURE 4: Flow chart of the proposed method.

then plot these PE values as a function of scale factor τ. We call this procedure multiscale permutation entropy analysis.

In order to select the best m for MPE calculation, we take the Gaussian white noise signal with length $N = 2048$ as an example. The MPEs are calculated under embedding dimension $m = 4, 5, 6$, and 7 when the parameters maximal scale factor $\tau_{\max} = 12$ and $\lambda = 1$. Correspondingly, their consuming times are 0.1880 second (s), 0.6710 s, 3.8290 s, and 27.6710 s, when a desktop computer with 2.0 GHz, Pentium Dual-Core CPU, 2.0 GB RAM, and MATLAB (R2011a) platform is utilized. The MPE is plotted as a function of the scale factor and is shown in **Figure 3**. From **Figure 3** it can be concluded that when m is less than 6 ($m = 4$ and 5), with the increase of scale factor τ, the PE values change very slowly with a value close to 1 and could not reflect the dynamic changes sensitively. However, if m is too large (e.g., $m = 7$), the calculation of PE would cost much runtime (27.6710 s for the data with length $N = 2048$) and the PE value is less than the expected one. As when $\tau = 1$, the Gaussian white noise signal should have an expected PE value close to 1, based on these consideration, $m = 6$ may be the most suitable.

3. The Proposed Method

3.1. Laplacian Score (LS) for Feature Selection. Theoretically, the extracted MPE features in 12 scales are able to identify the fault categories. However, the feature vector with a high dimension will be time-consuming and information inefficient for fault diagnosis. Therefore it is necessary to select the most important features which contain the main fault information from the 12 features, which could avoid the dimension disaster and improve the performance and efficiency of rolling bearing automatically fault diagnosis.

Laplacian score (LS) is a popular feature ranking based feature selection method and is mainly founded on Laplacian
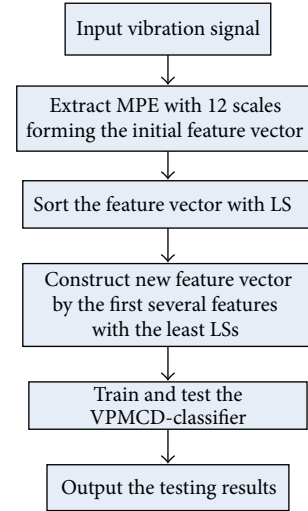
eigenmaps and locality preserving projection. Its basic idea is to estimate the features according their locality preserving power [18]. In LS algorithm those features with the lowest scores are chosen as the most important ones. LS has not been widely used in rolling bearing fault diagnosis for feature selection; in this paper it is employed to decrease the dimension of the initial fault features and select the most important features to represent the main fault information of vibration signal.

3.2. The Proposed Method. Based on the advantages of MPE, LS, and SVM, the proposed rolling bearing fault diagnosis method is described as follows.

(1) Calculate MPE for each rolling bearing vibration signal with parameter selection $m = 6$, $\tau = 1$, $N = 2048$, and the maximum scale factor $\tau_{\max} = 12$.

(2) Then the obtained MPEs in all scales (i.e., 12 PEs) are viewed as the initial feature vector to represent the main fault information of vibration signal.

(3) LS is employed to rank the 12 features from low to high score according to their importance and relationships with fault information.

(4) The first several features with the least scores are selected as the new feature vector.

(5) The new feature vectors are used to train and test the SVM based multifault classifier to fulfill fault diagnosis automatically.

The proposed method can be described briefly as in **Figure 4**.

In step (4) of the proposed method, as too many features will cost much training time and cause information redundancy while too few ones cannot completely reflect the fault information and get a lower accuracy, the novel feature vector in this paper is constructed by the first five features with the lowest LSs to achieve an effective fault diagnosis.
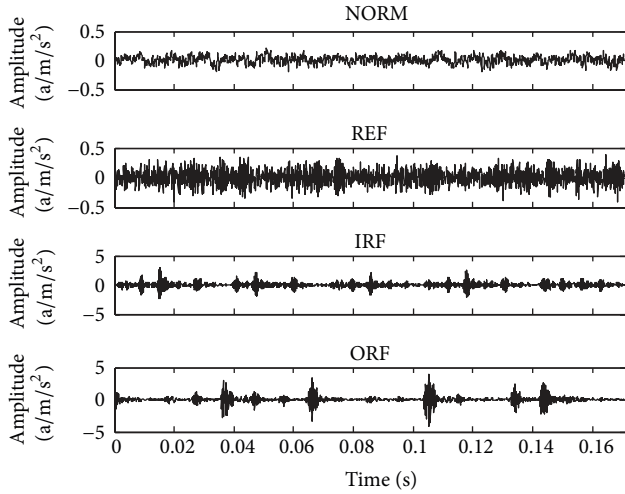
Figure 5: The time domain waveforms of normal and fault bearing vibration signals.

Figure 6: The MPE of normal and faulty bearing vibration signals. The results are the average of ten trials.

4. Analysis of Experimental Data

4.1. Experimental Data Description. The rolling bearing experimental data analyzed in the paper are kindly provided by Bearing Data Center, Case Western Reserve University [8, 23]. The tested bearing is 6205-2RS JEM SKF deep groove ball bearing, with motor load about 2206.50 watts and motor speed 1730 r/min. Here three rolling bearings with outer race fault (ORF), inner race fault (IRF), and rolling element fault (REF) are under our consideration and single point faults with defect sizes 0.5334 mm in diameter and 2.794 mm in depth were set into the tested bearing using electrodischarge machining. The data collection system consists of a high-bandwidth amplifier particularly designed for vibration signals and a data recorder with a sampling frequency of 12000 Hz per channel. A detailed signal collection depiction can be easily found in [8, 23].

4.2. Results and Discussions. The vibration signals of normal (NORM) bearing and bearings with fault (ORF, IRF, REF) are depicted in Figure 5.

It is unobvious to identify the normal and fault rolling bearings from each other especially differentiating NORM from REF and IRF from ORF. Therefore MPE is utilized to analyse above signals and their MPEs are plotted as a function of the scale factor in Figure 6.

From Figure 6 it can be found that the MPE with scale factor $\tau = 1$, namely, the PE of original vibration signal, could detect the dynamic changes of systems when the bearing works under a faulty condition. The PE of original vibration signal of normal rolling bearing is smaller than the PEs of rolling bearings with fault, which is coincident with Yan's conclusions [16]. In the literature [16] Yan and Liu concluded that the PE of normal condition is smaller than PEs of rolling bearings with worn rolling element and broken cage. When the rolling bearing is broken, the dynamic change will occur and can be detected and amplified by PE. Therefore when
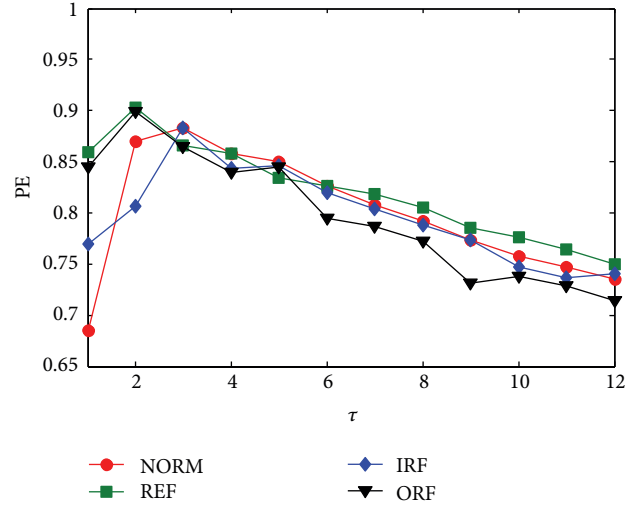
the rolling bearing is broken, the dynamic change will occur and cause a larger PE than that of normal condition. However, the single scale based PE only discriminates the faulty rolling bearing from normal ones (with threshold about 0.73) and cannot clearly identify the fault categories, that is, REF, IRF, or ORF. As the bearing vibration signals contain much more important fault information in other scales, it is essential to deal with the vibration signal using a MPE method.

If the extracted MPEs with 12 scales from the vibration signal are viewed as the feature vector, it will increase computational time and complexity, and the redundant information will decrease the classification accuracy. However, it is difficult for us to find out which feature contains the main fault information. In the literature [8], the statistical features of the MSE are used for reducing the dimension of feature vectors. However, the statistical features ignore the characteristics of the inner relation between the features. Therefore, in this paper the LS is employed to select the most important features to represent the vibration signal.

In this paper, normal and three faults (REF, IRF, and ORF) types of rolling bearing are under our consideration. Each type has 30 samples and there are totally 120 samples. By extracting MPE from each vibration signal, correspondingly, 120 initial feature vectors with 12 PEs can be obtained. For each fault type, 10 samples are randomly chosen for training and the remaining 20 samples are used for testing. Hence, a training dataset (with dimension 10×12) and a testing dataset (with dimension 20×12) are obtained.

Then, the LS is used to rank the 12 features according to their importance and the results are shown as follows:

$$LS_1 < LS_2 < LS_9 < LS_{11} < LS_{10} < LS_{12} < LS_6 < LS_7 < LS_8 < LS_3 < LS_4 < LS_5,$$

where the subscript stands for scale factor number. Therefore, the MPEs with $\tau = 1, 2, 9, 11$, and 10 are adopted to compose the new feature vector.
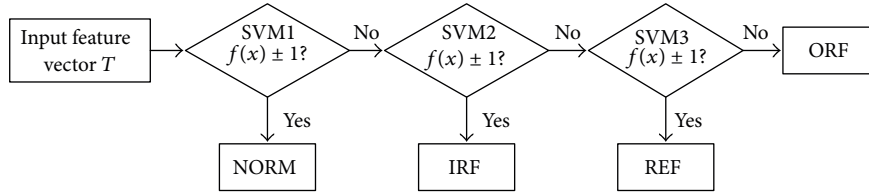
FIGURE 7: Structure diagram of SVM based multifault classifier.

TABLE 1: The SVM-classifier outputs of testing data with feature vector refined by LS.

Testing data	Fault type	SVM1	SVM2	SVM3
T_1–T_{20}	NORM	+1 (20)	—	—
T_{21}–T_{40}	IRF	−1 (20)	+1 (20)	—
T_{41}–T_{60}	REF	−1 (20)	−1 (20)	+1 (20)
T_{61}–T_{80}	ORF	−1 (20)	−1 (20)	−1 (20)

Next, a multi-fault classifier consisting of three SVMs, that is, SVM1, SVM2, and SVM3, is trained, where SVM1 is used to distinguish normal from the fault, SVM2 is used to discriminate IRF from REF and ORF, and SVM3 is used to discriminate REF from ORF. The structure diagram of the multi-fault classifier is depicted as in Figure 7.

After training the SVM-classifier with the 40 training feature vectors, the remaining 80 testing features are used to test the trained SVM-classifier and the outputs of the multiclassifier is shown in Table 1, from which it can be concluded that the classification accuracy of the proposed method on testing data achieves a perfect level (100%) and no samples are misclassified, which indicates that the proposed method can identify the fault categories effectively.

For comparison, a multiclassifier based on back propagation (BP) neural network [24–26] consisting of two layers in which the node numbers of input layer and output layer are 8 and 4, respectively, is used to fulfill the same classification problems. For convenience, NORM is labeled 1, IRF is labeled 2 (the output of BP-classifier plus 1), REF is labeled 3 (the output plus 2), and ORF is labeled 4 (the output plus 3). The training and testing samples are the same as the above SVM-classifier. The classification results of BP-classifier are given in Figure 8. The results indicate that the BP-classifier also recognizes the fault categories with accuracy of 100%. However, the training time of BP-classifier is much longer than the SVM-classifier's. Moreover, the accuracy cannot be high enough due to the limitations of "overfitting," slow convergence velocity, and relapsing into local extremum easily [27].

In addition, in order to verify the essentiality of multiscale analysis using MPE, the PE value of the original vibration signal, namely, the MPE with scale factor $\tau = 1$ is taken as the feature vector. Then train the SVM-classifier and BP-classifier with the same 40 training samples, respectively. And the outputs of SVM-classifier are given in Table 2, and the outputs of BP-classifier are depicted in Figure 9, respectively. It can be found that there are six testing data misclassified
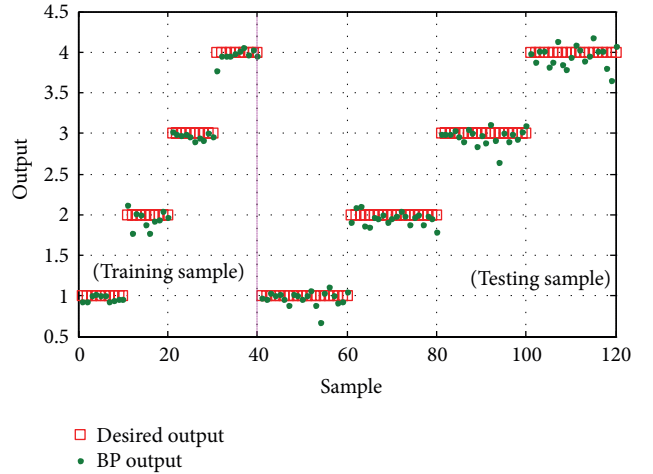

FIGURE 8: The BP classifier outputs of all samples with the same feature vector as SVM-classifier. The first 40 outputs are training data and the remaining 80 outputs are testing data.

TABLE 2: The outputs of SVM-classifier with feature vector consisting of one PE.

Samples	Fault type	SVM1	SVM2	SVM3
T_1–T_{20}	NORM	+1 (20)	—	—
T_{21}–T_{40}	IRF	−1 (20)	+1 (20)	—
T_{41}–T_{60}	REF	−1 (20)	−1 (20)	+1 (18), −1 (2)
T_{61}–T_{80}	ORF	−1 (20)	−1 (20)	−1 (16), +1 (4)

by both SVM- and BP-classifier with recognition rate of accuracy 92.5%. Therefore, the analysis results of Table 2 and Figure 9 indicate that the single scale based PE of original signal cannot reflect the nature of fault information, and it is necessary to handle the vibration signal using MPE method for getting much more fault information.

To verify that it is necessary and superior to refine the feature vectors using LS, without loss of generality, the MPE with scales 1, 2, 3, 4, and 5 are taken as the feature vector to train and test the SVM-classifier. After training the classifier, the outputs of testing data are given in Table 3. It is easy to find from the Table 3 that two testing samples are misclassified and the identification rate is 97.5%, which is lower than the proposed method (100%). Therefore, the analysis result indicates that it is essential to optimize features using LS.
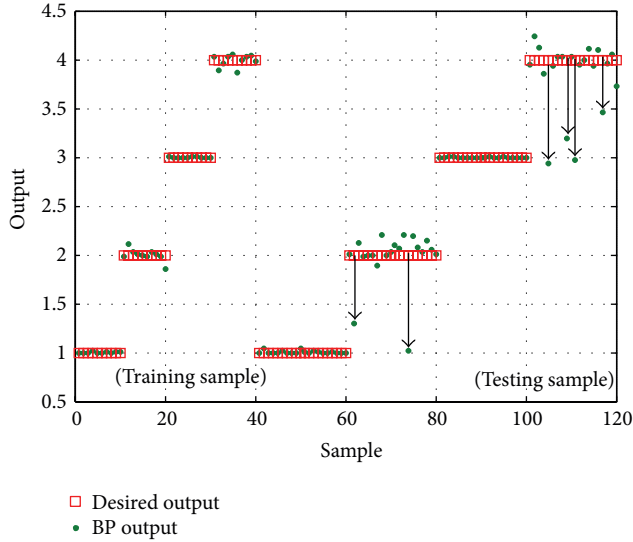
FIGURE 9: The outputs of BP-classifier with feature vector consisting of the PE when $\tau = 1$.

TABLE 3: The outputs of SVM-classifier with feature vector consisting of the first five PEs.

Samples	Fault type	SVM1	SVM2	SVM3
T_1–T_{20}	NORM	+1 (20)	—	—
T_{21}–T_{40}	IRF	−1 (20)	+1 (20)	—
T_{41}–T_{60}	REF	−1 (20)	−1 (20)	+1 (20)
T_{61}–T_{80}	ORF	−1 (20)	−1 (20)	−1 (18), +1 (2)

5. Conclusion

In consideration of the nonlinearity and nonstationarity of rolling bearing vibration signal, a novel rolling bearing fault diagnosis method based on MPE, LS, and SVM is proposed. Permutation entropy (PE) is defined to detect the dynamic changes of time series. For the complexity of the mechanical system, the vibration signal always contains much more important failure information in different scales. Therefore, in this paper MPE is adopted to extract the nonlinear fault characteristics from vibration signal. Besides, in order to achieve the fault diagnosis automatically, the SVM is utilized to construct the multifault classifier. Meanwhile, to refine the feature vector and select the most important features, Laplacian score (LS) is employed for feature selection. Finally, the proposed method is applied to rolling bearing experimental data. Also, the SVM-classifier is compared with BP-classifier and the single scale based PE is compared with MPE by analyzing the experimental data, and the comparison result indicates that the proposed method could get much higher identifying accuracy and has verified the necessities of analyzing the vibration signal with MPE and selecting feature by LS as well. Finally, the proposed method is aiming to fault diagnosis of rolling bearing and has been verified as an effective way by experiment data. However, the proposed method also has some problems, such as the number selection of feature vector refined by LS, the construction of

multiclassifier, and its generalization to other bearings or gear fault diagnosis, and they will be discussed and solved in the future work.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.
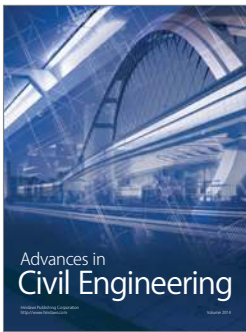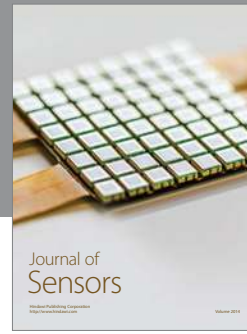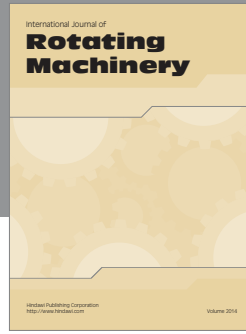
Acknowledgments

References

[1] D. Yu, J. Cheng, and Y. Yang, "Application of EMD method and Hilbert spectrum to the fault diagnosis of roller bearings," *Mechanical Systems and Signal Processing*, vol. 19, no. 2, pp. 259–270, 2005.

[2] E. Sejdić, I. Djurović, and J. Jiang, "Time-frequency feature representation using energy concentration: an overview of recent advances," *Digital Signal Processing*, vol. 19, no. 1, pp. 153–183, 2009.

[3] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hubert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society A*, vol. 454, no. 1971, pp. 903–995, 1998.

[4] Z. Wu and N. E. Huang, "A study of the characteristics of white noise using the empirical mode decomposition method," *Proceedings of the Royal Society A*, vol. 460, no. 2046, pp. 1597–1611, 2004.

[5] R. Yan and R. X. Gao, "Approximate entropy as a diagnostic tool for machine health monitoring," *Mechanical Systems and Signal Processing*, vol. 21, no. 2, pp. 824–839, 2007.

[6] D. Logan and J. Mathew, "Using the correlation dimension for vibration fault diagnosis of rolling element bearings—I. Basic concepts," *Mechanical Systems and Signal Processing*, vol. 10, no. 3, pp. 241–250, 1996.

[7] J. D. Jiang, J. Chen, and L. S. Qu, "The application of correlation dimension in gearbox condition monitoring," *Journal of Sound and Vibration*, vol. 223, no. 4, pp. 529–541, 1999.

[8] L. Zhang, G. Xiong, H. Liu, H. Zou, and W. Guo, "Bearing fault diagnosis using multi-scale entropy and adaptive neuro-fuzzy inference," *Expert Systems with Applications*, vol. 37, no. 8, pp. 6077–6085, 2010.

[9] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate and sample entropy," *American Journal of Physiology—Heart and Circulatory Physiology*, vol. 278, no. 6, pp. H2039–H2049, 2000.

[10] Y.-H. Pan, Y.-H. Wang, S.-F. Liang, and K.-T. Lee, "Fast computation of sample entropy and approximate entropy in biomedicine," *Computer Methods and Programs in Biomedicine*, vol. 104, no. 3, pp. 382–396, 2011.

[11] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of complex physiologic time series," *Physical Review Letters*, vol. 89, no. 6, Article ID 068102, 4 pages, 2002.

[12] M. Costa, A. L. Goldberger, and C. K. Peng, "Multiscale entropy analysis of biological signals," *Physical Review E*, vol. 71, Article ID 021906, 2005.

[13] W. Aziz and M. Arif, "Multiscale permutation entropy of physiological time series," in *Proceedings of the 9th International Multitopic Conference (INMIC '05)*, December 2005.

[14] C. Bandt and B. Pompe, "Permutation entropy: a natural complexity measure for time series," *Physical Review Letters*, vol. 88, no. 17, Article ID 174102, 4 pages, 2002.

[15] C. Bandt, G. Keller, and B. Pompe, "Entropy of interval maps via permutations," *Nonlinearity*, vol. 15, no. 5, pp. 1595–1602, 2002.

[16] R. Yan, Y. Liu, and R. X. Gao, "Permutation entropy: a non-linear statistical measure for status characterization of rotary machines," *Mechanical Systems and Signal Processing*, vol. 29, pp. 474–484, 2012.

[17] N. Nicolaou and J. Georgiou, "Detection of epileptic electroencephalogram based on Permutation Entropy and Support Vector Machines," *Expert Systems with Applications*, vol. 39, no. 1, pp. 202–209, 2012.

[18] X. He, D. Cai, and P. Niyogi, "Laplacian score for feature selection," in *Advances in Neural Information Processing System*, MIT Press, Cambridge, Mass, USA, 2005.

[19] Y. Yang, D. Yu, and J. Cheng, "A fault diagnosis approach for roller bearing based on IMF envelope spectrum and SVM," *Measurement*, vol. 40, no. 9-10, pp. 943–950, 2007.

[20] P. Konar and P. Chattopadhyay, "Bearing fault detection of induction motor using wavelet and Support Vector Machines (SVMs)," *Applied Soft Computing Journal*, vol. 11, no. 6, pp. 4203–4211, 2011.

[21] J. Cheng, D. Yu, J. Tang, and Y. Yang, "Application of SVM and SVD technique based on EMD to the fault diagnosis of the rotating machinery," *Shock and Vibration*, vol. 16, no. 1, pp. 89–98, 2009.

[22] Y. Cao, W.-W. Tung, J. B. Gao, V. A. Protopopescu, and L. M. Hively, "Detecting dynamical changes in time series using the permutation entropy," *Physical Review E*, vol. 70, no. 4, Article ID 046217, 7 pages, 2004.

[23] Y. Lei, Z. He, and Y. Zi, "A new approach to intelligent fault diagnosis of rotating machinery," *Expert Systems with Applications*, vol. 35, no. 4, pp. 1593–1600, 2008.

[24] P. K. Kankar, S. C. Sharma, and S. P. Harsha, "Fault diagnosis of ball bearings using machine learning methods," *Expert Systems with Applications*, vol. 38, no. 3, pp. 1876–1886, 2011.

[25] P. Qin, Y. Shen, J. Zhu, and H. Xu, "Dynamic analysis of hydro-dynamic bearing-rotor system based on neural network," *International Journal of Engineering Science*, vol. 43, no. 5-6, pp. 520–531, 2005.

[26] Y. Yang, D. Yu, and J. Cheng, "A roller bearing fault diagnosis method based on EMD energy entropy and ANN," *Journal of Sound and Vibration*, vol. 294, no. 1-2, pp. 269–277, 2006.

[27] S.-W. Fei and X.-B. Zhang, "Fault diagnosis of power transformer based on support vector machine with genetic algorithm," *Expert Systems with Applications*, vol. 36, no. 8, pp. 11352–11357, 2009.