

Research Article

A Wavelet Based Multiscale Weighted Permutation Entropy Method for Sensor Fault Feature Extraction and Identification

Qiaoning Yang and Jianlin Wang

College of Information Science & Technology, Beijing University of Chemical Technology, Beijing City Chaoyang District North Third Ring Road 15, Beijing 100029, China

Correspondence should be addressed to Jianlin Wang; wangjl@mail.buct.edu.cn

Received 20 February 2016; Revised 5 May 2016; Accepted 8 May 2016

Academic Editor: Fanli Meng

Copyright © 2016 Q. Yang and J. Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Sensor is the core module in signal perception and measurement applications. Due to the harsh external environment, aging, and so forth, sensor easily causes failure and unreliability. In this paper, three kinds of common faults of single sensor, bias, drift, and stuck-at, are investigated. And a fault diagnosis method based on wavelet permutation entropy is proposed. It takes advantage of the multiresolution ability of wavelet and the internal structure complexity measure of permutation entropy to extract fault feature. Multicluseter feature selection (MCFS) is used to reduce the dimension of feature vector, and a three-layer back-propagation neural network classifier is designed for fault recognition. The experimental results show that the proposed method can effectively identify the different sensor faults and has good classification and recognition performance.

1. Introduction

At present, the sensor is widely used in various processes to obtain a variety of physical quantity of data. In practical applications, due to the harsh external environments, battery depletion, aging, and other reasons, the sensor is prone to failure or even damage [1, 2]. The data obtained from the fault sensor has low reliability, and the subsequent judgment, recognition, decision, and control based on these low quality data will lose the meaning. The reliability of sensor data and the identification of sensor fault are important research subjects. Sensor fault identification mainly consists of two aspects: fault feature extraction and fault pattern classification [3, 4].

Wavelet transform is a widely used time-frequency analysis technology. Using wavelet transform, signals are decomposed into multilevel time-frequency components. Suitable wavelet basis for wavelet decomposition is important for fault information representation [5]. The selection method of wavelet base includes many kinds, such as the minimum joint entropy standard, the minimum conditional entropy standard, the maximum mutual information criterion, the minimum relative entropy standard, and maximum energy-Shannon entropy criterion [6, 7]. Maximum energy-Shannon

entropy criterion takes energy intensity and energy distribution into consideration and is capable of extracting the sensor fault variations effectively. The wavelet basis with maximum entropy Shannon energy ratio is the most appropriate wavelet basis.

After wavelet decomposition, the main problem is how to extract fault information from the coefficients in decomposed subbands. The traditional Shannon entropy only considers the probability distribution of the signal value and does not consider the order structure of the signal value. Paper [8] combines the concept of Shannon entropy with the theory of symbol to propose a new complexity measure, which is the permutation entropy (PE). Permutation entropy is a time series complexity measure based on comparison of neighborhood values and the numerical mapping into symbol sequence pattern. It can describe the local structure features of time series signal and enlarge the subtle changes in the signal with low complexity and antinoise ability. PE is a kind of effective method for classification of different signal state, identification of the breakpoint in time sequence, prediction of the future trend of the time series, determination of causal relationship [9], and so forth. In order to overcome the shortcomings of PE's single scale, Aziz and Arif [10] proposed a multiscale permutation entropy (MPE) to estimate the

complexity in different scales of time series. MPE can describe the structural characteristics and complexity of time series in multiple scales and is widely used in heart sound signal analysis [10] and bearing fault diagnosis [11].

On the other hand, the disadvantage of PE is the lack of amplitude information about the signal except sequential pattern [12]. Paper [13] puts forward the weighted permutation entropy (WPE). It extracts the sequential pattern of time series and retains the amplitude information of time series. Although the amplitude information of time series is used by WPE, it can only reflect the complexity of time series in one single scale. Multiscale analysis and weighted permutation entropy are combined, and multiscale weighted permutation entropy (MWPE) emerges. MWPE represents the complexity measure of the signal on multiscale and reflects the microlocal structure complexity and the amplitude information of the signal. It is widely used in a variety of signal analysis, such as analysis for bubbly oil-in-water [14], flow bearing fault diagnosis [15–17], biomedical signal analysis [18–21], and stock information analysis [22].

From the viewpoint of structure feature presentation, PE can extract the local microstructure feature and wavelet transform can extract the global macrostructure feature. So the combination of wavelet transform with MWPE can comprehensively represent the feature of the sensor fault. A wavelet based multiscale weighted permutation entropy (WMWPE) is proposed in this paper. WMWPEs of different subbands are used to extract signal features. Because the dimensions of WMWPE features are relatively high, it may cause low identification accuracy and time consuming. So the selection of the most important features in WMWPEs is needed [23, 24]. In this paper, the multicluster feature selection (MCFS) [23] is used as feature selection method, which takes into account both the importance of each feature itself and the correlation between all features. By sorting the score of MCFS, the first K features with larger MCFS score are selected as the K important feature. Through the MCFS feature selection algorithm, the recognition accuracy is guaranteed, the feature vector dimension is reduced, and the computational efficiency is improved.

Naturally, after feature selecting using MCFS, the multifault classifier is needed to conduct the fault diagnosis. A three-layer BP neural network is adopted as classifier to identify fault. The K features selected by MCFS are fed into the classifier to identify sensor fault.

The remainder of this paper is organized as follows. Section 2 introduces permutation entropy and multiscale permutation entropy. In Section 3, the wavelet based multiscale weighted permutation entropy and corresponding fault identification method are presented in detail. Using practical data, the performance of the proposed method is investigated in Section 4. Section 5 makes some concluding remarks.

2. Permutation Entropy and Multiscale Permutation Entropy

2.1. Permutation Entropy. The permutation entropy (PE) is defined as follows [8]. Given a time series $x(n)$ with the length

N , then the m dimensional embedding vector at i time is defined as

$$X_i^m = \{x(i), x(i + \tau), \dots, x(i + (m - 1)\tau)\}, \quad (1)$$

$$i = 1, 2, \dots, N - (m - 1)\tau,$$

where m is the embedding dimension and τ is time delay. X_i^m has a permutation $\pi_{r_0 r_1 \dots r_{m-1}}$, if it satisfies that

$$x(i + r_0 \tau) \leq x(i + r_1 \tau) \leq \dots \leq x(i + r_{m-1} \tau), \quad (2)$$

where $0 \leq r_i \leq m - 1$ and $r_i \neq r_j$. There are $m!$ possible permutations for an m -tuple vector. For each permutation, we determine the relative frequency by

$$p(\pi_k) = \frac{\#\{i \mid 1 \leq i \leq N - (m - 1)\tau, X_i^m \text{ has type } \pi_k\}}{N - (m - 1)\tau}, \quad (3)$$

where $\sum_{k=1}^K p(\pi_k) = 1$, $1 \leq K \leq m!$, and $\#\{\}$ represents the number of one set.

The definition of PE with m dimension is defined as

$$H_{\text{PE}}(m) = -\sum_{k=1}^K p(\pi_k) \ln p(\pi_k). \quad (4)$$

The maximum value of $H_{\text{PE}}(m)$ is $\ln(m!)$, when all possible permutations appear with the same probability. Therefore, the normalized permutation entropy (NPE) can be obtained as

$$H_{\text{NPE}}(m) = \frac{H_{\text{PE}}(m)}{\ln(m!)}. \quad (5)$$

For any time series, $0 \leq H_{\text{NPE}}(m) \leq 1$ is satisfied. The value of $H_{\text{PE}}(m)$ depends on the selection of the embedding dimension m and delay τ . If m is too small, the scheme will not work well since there are too few distinct states. However, it is often inappropriate to choose m as a large value for detecting the dynamic change of a time series.

2.2. Weighted Permutation Entropy. Weighted permutation entropy (WPE) incorporates significant amplitude information from the time series when retrieving the sequential patterns. The main motivation aims at saving useful amplitude information carried by the signal. WPE is defined as follows [13].

Given a time series $x(n)$ and m dimensional embedding vector X_i^m as (1) shows, the relative frequency of each motif in (3) is modified to include the weighted information w_i for each X_i^m . Weight values w_i are calculated in (6) based on the variance or energy of each subsequence X_i^m :

$$w_i = \frac{1}{m} \sum_{l=1}^m (x_{i+(l-1)\tau} - \overline{X_i^m})^2, \quad (6)$$

where $\overline{X_i^m}$ is the mean of X_i^m :

$$\overline{X_i^m} = \frac{1}{m} \sum_{l=1}^m x_{i+(l-1)\tau}. \quad (7)$$

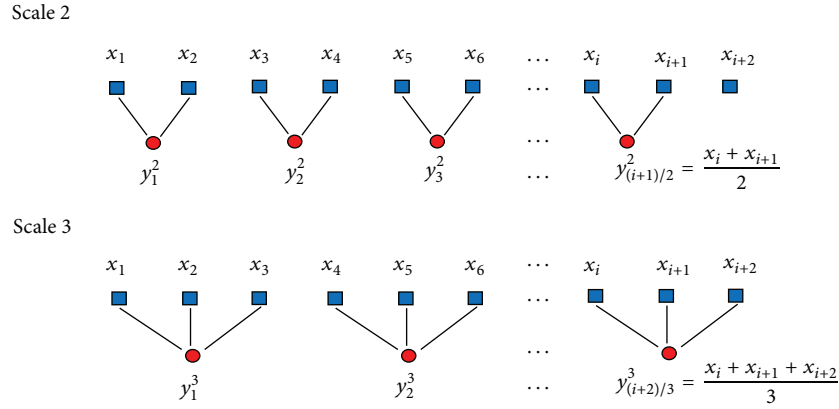


FIGURE 1: Coarse-grained procedure.

Thus, each pair of weight value w_i and motif type π_k can represent full feature for each vector X_i^m . By using weight value, WPE extends the concept of PE with the addition of amplitude information prior to the computing of probability occurrence of each motif defined in (6). Weighted relative frequency is defined as

$$p_w(\pi_k) = \frac{\sum \{w_i \mid 1 \leq i \leq N - (m-1)\tau, X_i^m \text{ has type } \pi_k\}}{\sum w_i} \quad (8)$$

Then WPE of time series $x(n)$ is

$$H_{\text{WPE}}(m) = -\sum_{k=1}^K p_w(\pi_k) \ln p_w(\pi_k). \quad (9)$$

2.3. Multiscale Weighted Permutation Entropy. Multiscale analysis [22] is to obtain new time series of the original time series by a coarse-grained size process, which can estimate the complexity of time signals at different scales. Multiscale weighted permutation entropy (MWPE) is the combination of weighted permutation entropy and multiscale analysis. MWPE can totally describe the microstructure complexity and amplitude information of the time series on multiscale.

The MWPE procedure is summarized in the following steps. Firstly, the original time series $x(n) = \{x_1, x_2, \dots, x_N\}$ is divided into nonoverlapping windows of length s . Secondly, the data points inside each window are averaged by (10), and the coarse-grained time series y_j^s is got. The schematic illustration of the coarse-grained procedure is shown in Figure 1. Consider the following:

$$y_j^s = \frac{1}{s} \sum_{i=(j-1)s+1}^{js} x_i, \quad 1 \leq j \leq \frac{N}{s}. \quad (10)$$

Weighted permutation entropy of y_j^s is MWPE, as shown in

$$\text{MWPE}(x, s, m, \tau) = \text{WPE}(y_j^s, m, \tau). \quad (11)$$

MWPE is a function of scale factor s and represents the complexity of time series on each scale. When $s = 1$, MWPE

is the same as WPE; that is to say, WPE is a special form of MWPE. For most of signals, the MWPE on one single scale is not enough to describe complexity of the signal, and the multiscale weighted permutation entropy is more suitable for the analysis of all kinds of actual signals.

2.4. Parameters Selection. Before computing multiscale weighted permutation entropy, four important parameters including the length N of time series data, embedding dimension m , time delay τ , and scale factor s are needed to set.

The value of PE mainly depends on the embedding dimension m and time delay τ . Embedding dimension m determines the number of sequential pattern, and the maximum sequential pattern number is $m!$, so m plays an important role in calculating the probability of the sequential pattern. If embedding dimension is too small (such as less than 3), it is hard to differentiate the sequential patterns. If the embedding dimension m is too large (such as more than 8), calculation of PE is time consuming, and it is not easy to observe the small changes in the signal [16].

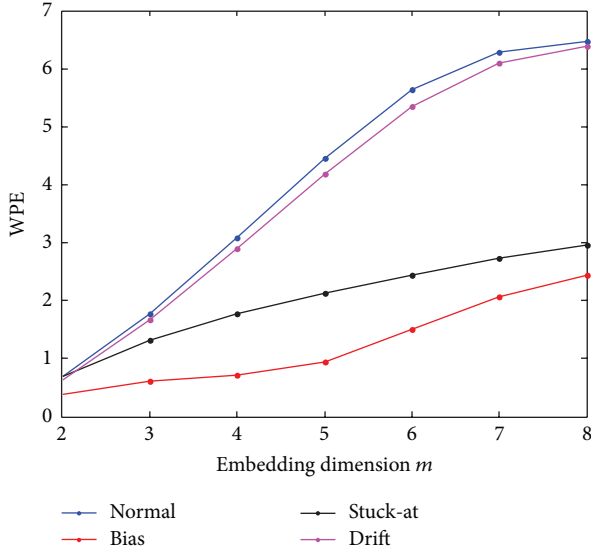
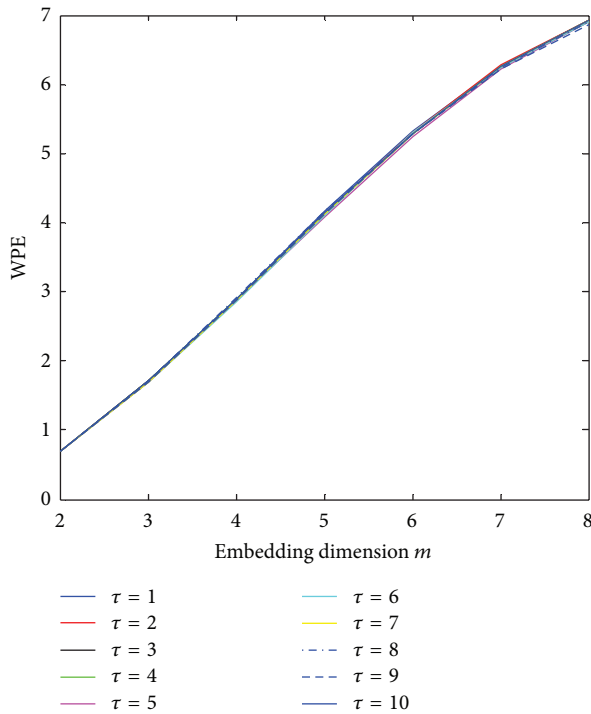
The time delay τ is related to the sampling rate of the signal. As [8] suggests, the time delay τ is set to 1 in this paper.

The length of time series N has a great influence on the calculation of PE, a larger N will cause low computational efficiency, and a smaller N cannot completely describe the complexity of the time series. $N \geq 5m!$ is recommended by [9].

To illustrate the rationality of parameters selection, some experiments are conducted. The experiment data used in this paper is 1-minute gas sensor data. Sampling rate is 100 Hz and the data has 6000 sampling points. Given that the scale factor s is a fixed value, Figure 2 shows the weighted permutation entropy (WPE) of four kinds of sensor data under different embedding dimensions, $m = 2-8$, and the time delay $\tau = 1$. As shown in Figure 2, when $m = 5$, WPE can separate the four kinds of signals. So $m = 5$ is feasible and is used in this paper.

In Figure 3, it is easy to find that different τ almost has no effect on WPE, so it is reasonable for the fact that τ is set to 1.

In this paper, the scale factor s is set to 20. MWPEs on 20 scales are used as the features to identify and classify faults.

FIGURE 2: The relationship of m and WPE.FIGURE 3: The relationship of τ and WPE.

3. Wavelet Based Multiscale Weighted Permutation Entropy (WMWPE) and Fault Identification Method

Although the MWPE takes advantage of the local microstructure information and amplitude information of signal, the macrostructure information is not explored. Wavelet transform is a powerful method to explore it and can be used to extract the global macrostructure information of the signal.

So the combination of wavelet transform with MWPE can comprehensively describe the multiple features of the signal.

3.1. WMWPE. Given the analyzed signal $x(n)$ and wavelet decomposed level J , wavelet decomposed coefficients are expressed as $\{D_1(n), D_2(n), \dots, D_J(n), A_J(n)\}$, corresponding to the coefficients $D_1(n)$ of the first high frequency subband (HF₁), the coefficients $D_2(n)$ of the second high frequency subband (HF₂), ..., the coefficients $D_J(n)$ of the J high frequency subband (HF _{J}), and the coefficients $A_J(n)$ of the J low frequency subband (LF _{J}). $A_J(n)$ belongs to the low frequency subband, which represents the macrostructure of the analyzed signal. So $A_J(n)$ is the preferable subband to calculate WMWPE:

$$\text{WMWPE}(A_J(n), s, m, \tau) = \text{WPE}(y_j^s, m, \tau), \quad (12)$$

where $y_j^s = (1/s) \sum_{i=(j-1)s+1}^{js} A_J(n)$, $1 \leq j \leq N_J/s$, N_J is the number of coefficients in LF _{J} , and s is the scale factor.

3.2. WMWPE Based Fault Identification Method. The WMWPE based fault identification method takes advantage of wavelet transform, WMWPE, MCFS, and BP-NN. It provides a full working flow of feature selection and fault identification as shown in Figure 4. Detailed procedures are described as follows.

Step 1. Use maximum energy-to-Shannon entropy ratio criterion to choose a proper wavelet base.

Step 2. The sensor signals are decomposed by the selected wavelet base. A series of wavelet subband signals are obtained, and the appropriate subband signal is selected to extract feature.

Step 3. Calculate multiscale weighted permutation entropy of the selected wavelet subbands, and WMWPEs are got.

Step 4. After feature extraction, calculate MCFS score of the WMWPE features. According to the MCFS score ranked from high to low, select the features corresponding to the top K MCFS score as the best subset of features.

Step 5. The selected K features by MCFS are fed into BP-NN classifier for sensor fault recognition.

4. Experiments and Result Analysis

In this paper, the experimental data set is the measurement recordings collected from an array of 72 metal-oxide gas sensor array-based chemical detection platform [25]. The sampling rate of data is 100 Hz. One-minute sampling data (6000 sampling points) of each sensor is used as original data. Three kinds of fault (bias, stuck-at, and drift) are injected into original data. Each kind has 120 groups of data. For bias fault, the bias constant is 2% of the average value of original data. In each stuck-at fault, there are two segments of data with constant value (98% of the average value of the original). In drift fault groups, drift rate is 0.1% of the average value of

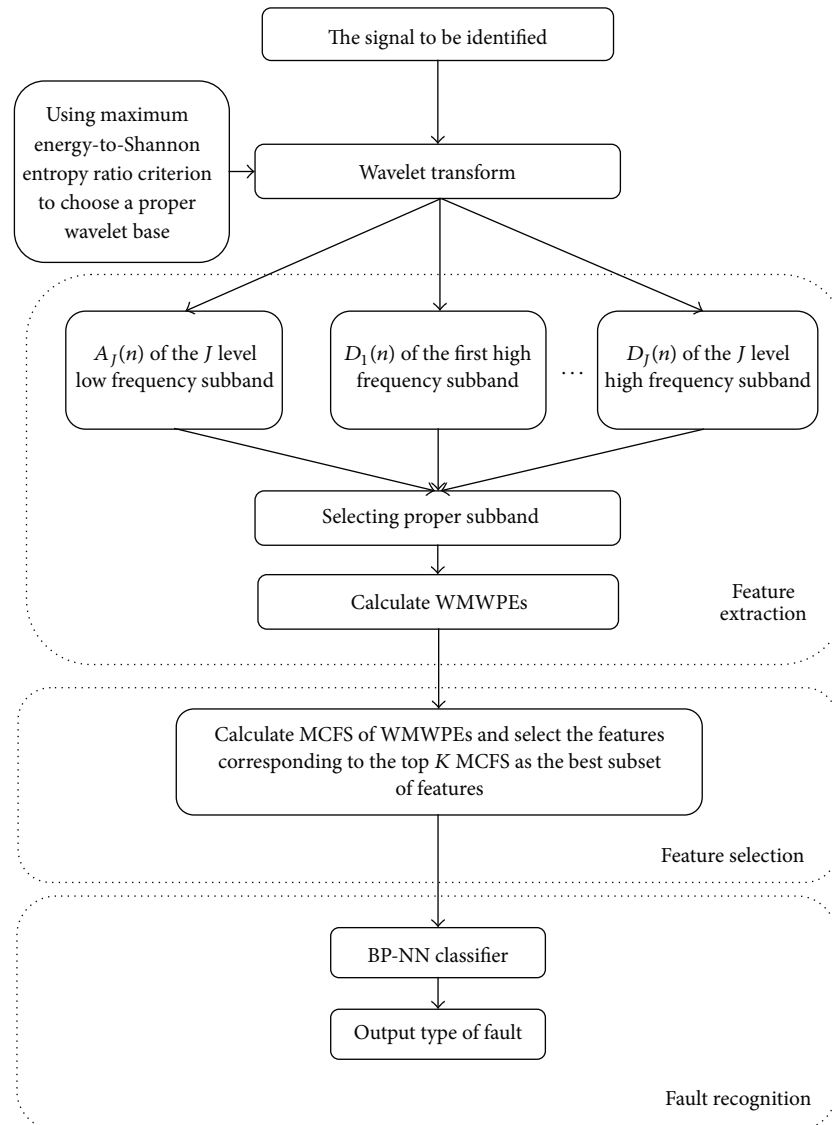


FIGURE 4: Flowchart of sensor fault feature extraction and recognition method.

original data. Normal sensor data and fault sensor data are shown in Figure 5, respectively.

(1) *Selection of Wavelet Subband.* Before selection of wavelet subband, the maximum energy-to-Shannon entropy ratio criterion is used for wavelet base selection. More detailed information can be found in [26]. Figure 6 is the average WMWPE of 120 sensor signals after 3-level wavelet decomposition. It is easy to find that WMWPE of the low frequency subband can easily differentiate the four kinds of sensor signals (normal, bias, stuck-at, and drift) on different scale factor s , and WMWPE of three high frequency subbands cannot distinguish the four kinds of sensor signals. The experiment results show that selecting the low frequency subband for feature extraction is feasible.

(2) *Feature Extraction and Fault Identification.* After 3-level wavelet decomposition, feature extraction on the third-level

low frequency subband is the further job. The WMWPEs of the selected low frequency subband signal on 20 scales are calculated. These WMWPEs will be fed into BP-NN classifier to identify fault.

A three-layer BP-NN neural network is used as classifier in the experiments. The hidden layer of the network has 10 neural nodes for learning. The network is trained by scaled conjugate gradient back-propagation method. Mean squared error is used as performance function. To obtain the generalized identification performance, 10-fold cross-validation [27, 28] is used in this paper. In 10-fold cross-validation, the proposed feature vectors are randomly partitioned into 10 equal sized subsets. Of the 10 subsets, a single subset is retained as the validation data for testing, and the remaining 9 subsets are used as training data. The cross-validation process is then repeated 10 times, with each of the 10 subsets used exactly once as the validation data. The 10 results from the folds can then be averaged to produce a single estimation.

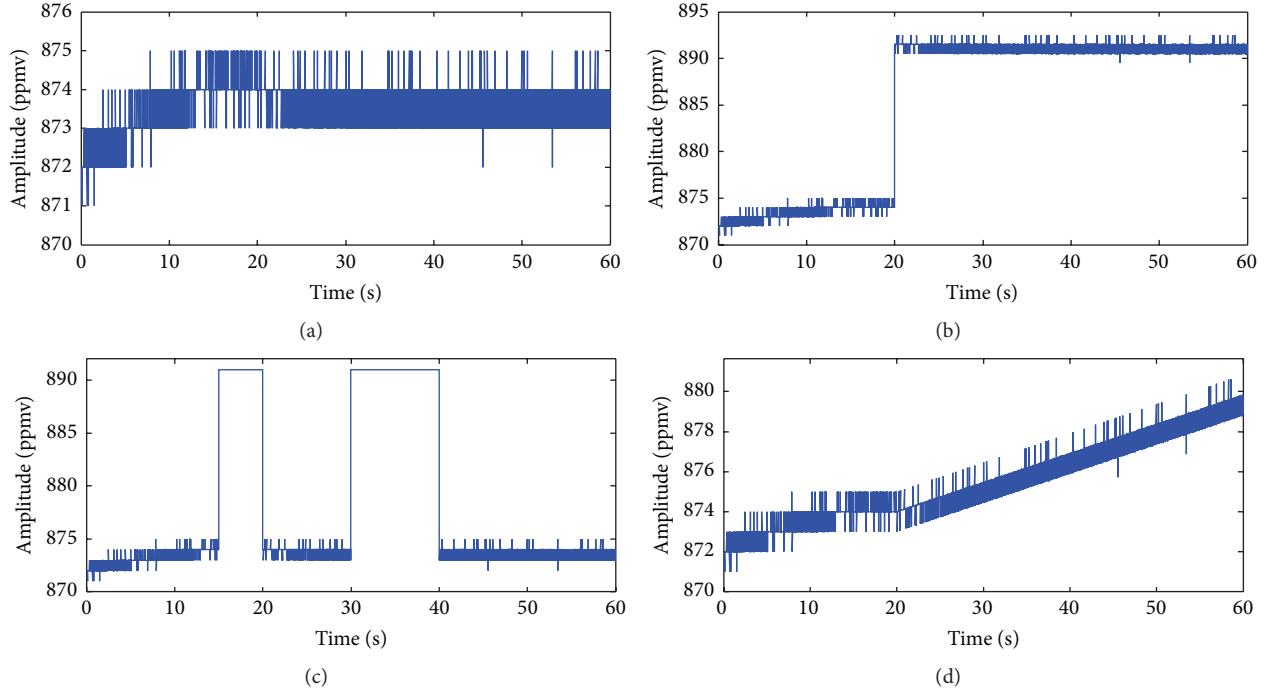


FIGURE 5: Normal sensor data and fault sensor data. (a) Normal data, (b) bias, (c) stuck-at, and (d) drift.

TABLE 1: Average identification accuracy and standard deviation of 20 WMWPEs of different subband.

Subband	Average identification accuracy \pm standard deviation
LF ₃	99.50% \pm 0.08
HF ₁	60.94% \pm 1.42
HF ₂	72.99% \pm 0.64
HF ₃	76.11% \pm 0.65

TABLE 2: Identification results of WMWPE feature of LF₃ subband.

Actual type	Identification results			
	Normal	Bias	Stuck-at	Drift
Normal	99.43% \pm 0.24	0%	0%	0.57% \pm 0.24
Bias	0%	100%	0%	0%
Stuck-at	0%	0%	100%	0%
Drift	1.18% \pm 0.17	0.28% \pm 0.02	0%	98.54% \pm 0.16

To illustrate the identification performance of WMWPE on different subbands (HF₁, HF₂, HF₃, and LF₃), all WMWPEs of 20 scales are used to recognize faults. Table 1 shows the average identification accuracy of different subband. WMWPE of LF₃ has much better identification performance than that of other subbands. The average identification accuracy of LF₃ is 99.5% and about 20% higher than others. At the same time, LF₃ has the smallest identification standard deviation than others. It explains the rationality of selecting WMWPE of LF₃ subband as features.

Table 2 is the detailed identification results of WMWPE from LF₃ subband. For bias and stuck-at, the identification

accuracy is 100%. There is false judgment in the identification of normal and drift signal. The reason is that the local structure pattern and the amplitude of bias and stuck-at are significantly different from those of normal signal. So it is easy to use WMWPE to identify them. Although the drift causes the change of signal amplitude, small change of the amplitude may bring little change to the sequential pattern. And some false negative results exists.

(3) *Feature Selection*. There exists large consuming time and information redundancy, if selecting all 20 scales of WMWPE as features and implementing classification. In order to improve the efficiency of the proposed method, MCFS is used to reduce the dimension of feature vector.

After calculating MCFS of 20 WMWPE features, the top K features with larger MCFS score are selected as the inputs of BP-NN. The feature selection result and the corresponding identification accuracy are shown in Table 3.

Table 3 shows that the average identification accuracy and standard deviation obtained by 5 new selected features are almost the same as those of all 20 features. The experiment results show that, using feature selection algorithm MCFS, the dimension of the feature vector is reduced from 20 to 5, but the recognition accuracy is maintained at about 99.5% with standard deviation of 0.08. That is because MCFS not only considers the classification ability of single feature, but also considers the relationship between the features.

(4) *Performance Comparison of Different Features*. In order to further illustrate the superiority of the proposed method, the identification performance of the proposed method WMWPE is compared with that of MWPE, WWPE, WPE,

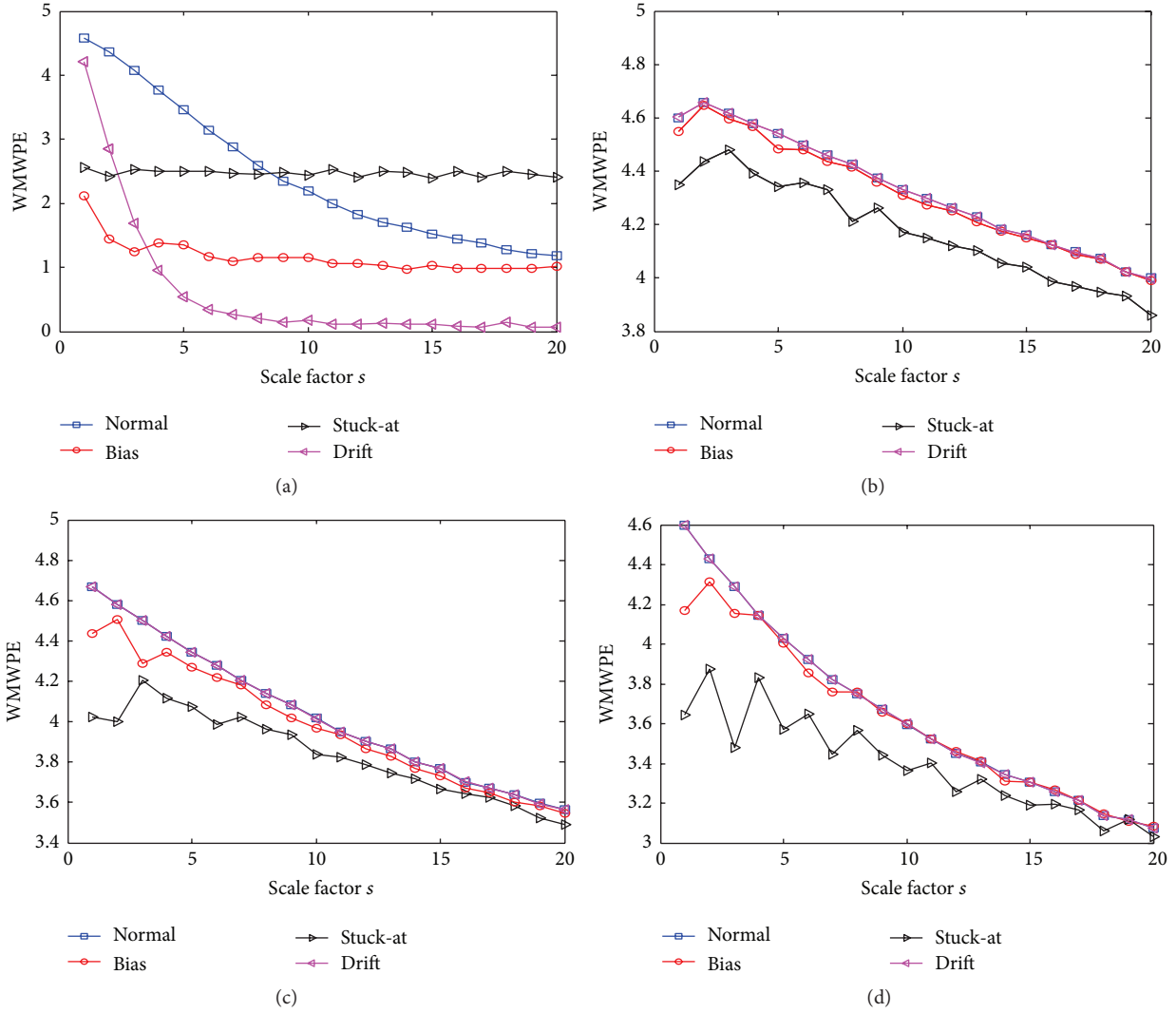


FIGURE 6: The average WMWPE of 120 sensor signals after wavelet decomposition. (a) Low frequency-level 3. (b) High frequency-level 1. (c) High frequency-level 2. (d) High frequency-level 3.

TABLE 3: Average identification accuracy of the selected feature vectors.

Feature dimension K	The scale index of selected features	Average identification accuracy \pm standard deviation
1	[1]	83.46% \pm 0.53
2	[19, 20]	87.19% \pm 0.72
3	[18, 13, 17]	97.42% \pm 0.27
4	[18, 17, 3, 4]	98.74% \pm 0.23
5	[18, 4, 6, 5, 7]	99.46% \pm 0.09
6	[18, 4, 5, 2, 1, 6]	99.37% \pm 0.15
7	[4, 18, 5, 2, 1, 6, 17]	99.30% \pm 0.08
8	[4, 18, 5, 2, 1, 6, 3, 7]	99.42% \pm 0.13
9	[4, 5, 18, 2, 1, 6, 3, 7, 8]	99.33% \pm 0.07
10	[4, 5, 18, 2, 1, 3, 6, 7, 8, 12]	99.43% \pm 0.10
All 20 features	[1, ..., 20]	99.50% \pm 0.08

TABLE 4: Average identification accuracy of using different features.

Features	Average identification accuracy \pm standard deviation
WMWPE	99.50% \pm 0.08
MWPE	98.06% \pm 0.20
WWPE	83.34% \pm 0.62
WPE	45.15% \pm 0.49
MPE	73.42% \pm 0.40
PE	35.27% \pm 0.20

MPE, and PE, wherein WWPE is WPE of the low frequency subband signal under wavelet decomposition. Table 4 shows the average recognition accuracy and standard deviation for the four kinds of signals using different features over 10 times of 10-fold cross-validation. The more detailed recognition results are shown in Table 5.

TABLE 5: Identification results of different features.

Actual type	Identification results			
	Normal	Bias	Stuck-at	Drift
		MWPE		
Normal	97.93% ± 0.56	0%	0%	2.07% ± 0.51
Bias	0%	98.13% ± 0.48	1.87% ± 0.46	0%
Stuck-at	0%	1.43% ± 0.23	98.57% ± 0.22	0%
Drift	1.43% ± 0.2	0.8% ± 0.3	0%	97.77% ± 0.19
		WWPE		
Normal	90.87% ± 0.52	0%	0%	9.13% ± 0.46
Bias	0%	74.5% ± 1.62	25.50% ± 1.51	0%
Stuck-at	0%	3.67% ± 0.21	96.33% ± 0.26	0%
Drift	26.83% ± 0.96	0%	1.47% ± 0.41	71.7% ± 1.34
		WPE		
Normal	40.87% ± 6.97	13.03% ± 1.43	6.97% ± 0.82	39.13% ± 7.42
Bias	15.70% ± 2.05	35.30% ± 4.52	27% ± 3.63	22% ± 2.85
Stuck-at	6.2% ± 1.31	20.47% ± 2.95	64.13% ± 2.71	9.20% ± 1.47
Drift	42.80% ± 6.35	10.70% ± 1.60	6.20% ± 0.84	40.30% ± 6.76
		MPE		
Normal	57.76% ± 1.40	37.91% ± 1.38	4.33% ± 0.17	0%
Bias	42.06% ± 0.82	53.74% ± 0.71	4.20% ± 0.32	0%
Stuck-at	3.20% ± 0.44	10.13% ± 1.42	86.67% ± 1.42	0%
Drift	4.50% ± 0.08	0%	0%	95.50% ± 0.24

As shown in Tables 2, 4, and 5, the identification accuracy of WMWPE is higher than that of MWPE. The same conclusion can be got from the performance comparison of WWPE with WPE. That is because wavelet transform brings macrostructure information into the feature. Combined with macro- and microstructure information, WMWPE and WWPE can get better performance.

Comparing the results of WMWPE to WWPE, MWPE to WPE, and MPE to PE, the multiscales analysis can improve the identification precision over 16%. The main reason is that the multiscale feature can explore more local structure information of the signals than the single scale feature.

Comparing the results of MWPE to MPE and WPE to PE, the amplitude information can bring about 25% and 9% increase of average identification accuracy, respectively.

So the macro- and microstructure information and amplitude information are all explored by WMWPE. The experiment results validate that the proposed method based on WMWPE can achieve high identification accuracy for sensor fault.

5. Conclusion

How to find an effective feature extraction method for sensor fault analysis and identification is always an important issue. Taking full advantage of macrostructure information, microstructure information, and amplitude information of the typical sensor faults, this paper proposed a new sensor fault feature extraction and identification method based on wavelet transform and multiscale weighted permutation entropy. The appropriate based wavelet selection, feature

extraction, multicluster feature selection, and BP classifier are investigated. Actual chemical gas concentration data is used to evaluate the performance of the proposed method. Experiment results show that the proposed WMWPE extracts more comprehensive feature information and can achieve higher fault recognition accuracy than other kinds of features.

Competing Interests

The authors declare no competing interests.

References

- [1] G. Heredia and A. Ollero, "Sensor fault detection in small autonomous helicopters using observer/Kalman filter identification," in *Proceedings of the IEEE International Conference on Mechatronics (ICM '09)*, pp. 1–6, April 2009.
- [2] J. C. D. Silva, A. Saxena, E. Balaban, and K. Goebel, "A knowledge-based system approach for sensor fault modeling, detection and mitigation," *Expert Systems with Applications*, vol. 39, no. 12, pp. 10977–10989, 2012.
- [3] M. Zhang, J. Wu, and Y. Wang, "Sensor soft fault detection method of autonomous underwater vehicle," in *Proceedings of the IEEE International Conference on Mechatronics and Automation (ICMA '09)*, pp. 4839–4844, Changchun, China, August 2009.
- [4] H. Liu, M. Huang, I. Janghorban, P. Ghorbannezhad, and C. Yoo, "Faulty sensor detection, identification and reconstruction of indoor air quality measurements in a subway station," in *Proceedings of the 11th International Conference on Control, Automation and Systems (ICCAS '11)*, pp. 323–328, October 2011.

- [5] W. He, Q. Miao, M. Azarian, and M. Pecht, "Health monitoring of cooling fan bearings based on wavelet filter," *Mechanical Systems and Signal Processing*, vol. 64–65, pp. 149–161, 2015.
- [6] R. Yan, *Base wavelet selection criteria for non-stationary vibration analysis in bearing health diagnosis [Ph.D. thesis]*, University of Massachusetts Amherst, Amherst, Mass, USA, 2007.
- [7] P. K. Kankar, S. C. Sharma, and S. P. Harsha, "Fault diagnosis of ball bearings using continuous wavelet transform," *Applied Soft Computing Journal*, vol. 11, no. 2, pp. 2300–2312, 2011.
- [8] C. Bandt and B. Pompe, "Permutation entropy: a natural complexity measure for time series," *Physical Review Letters*, vol. 88, no. 17, Article ID 174102, 2002.
- [9] M. Zanin, L. Zunino, O. A. Rosso, and D. Papo, "Permutation entropy and its main biomedical and econophysics applications: a review," *Entropy*, vol. 14, no. 8, pp. 1553–1577, 2012.
- [10] W. Aziz and M. Arif, "Multiscale permutation entropy of physiological time series," in *Proceedings of the Pakistan Section Multitopic Conference*, pp. 1–6, Karachi, Pakistan, December 2005.
- [11] S.-D. Wu, P.-H. Wu, C.-W. Wu, J.-J. Ding, and C.-C. Wang, "Bearing fault diagnosis based on multiscale permutation entropy and support vector machine," *Entropy*, vol. 14, no. 8, pp. 1343–1356, 2012.
- [12] P. L. Vuong, A. S. Malik, and J. Bornot, "Weighted-permutation entropy as complexity measure for electroencephalographic time series of different physiological states," in *Proceedings of the 3rd IEEE Conference on Biomedical Engineering and Sciences (IECBES '14)*, pp. 979–984, Kuala Lumpur, Malaysia, December 2014.
- [13] B. Fadlallah, B. Chen, A. Keil, and J. Principe, "Weighted-permutation entropy: a complexity measure for time series incorporating amplitude information," *Anesthesiology*, vol. 87, no. 2, pp. 1647–1650, 2013.
- [14] X. Chen, N.-D. Jin, A. Zhao, Z.-K. Gao, L.-S. Zhai, and B. Sun, "The experimental signals analysis for bubbly oil-in-water flow using multi-scale weighted-permutation entropy," *Physica A: Statistical Mechanics & its Applications*, vol. 417, pp. 230–244, 2015.
- [15] J. Zhou, J. Xiao, H. Xiao, W. Zhang, W. Zhu, and C. Li, "Multi-fault diagnosis for rolling element bearings based on intrinsic mode permutation entropy and ensemble optimal extreme learning machine," *Advances in Mechanical Engineering*, vol. 6, Article ID 803919, 2014.
- [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] X. Zhang, Y. Liang, J. Zhou, and Y. Zang, "A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM," *Measurement*, vol. 69, pp. 164–179, 2015.
- [18] C. Bian, C. Qin, Q. D. Y. Ma, and Q. Shen, "Modified permutation-entropy analysis of heartbeat dynamics," *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, vol. 85, no. 2, Article ID 021906, pp. 439–441, 2012.
- [19] U. Parlitz, S. Berg, S. Luther, A. Schirdewan, J. Kurths, and N. Wessel, "Classifying cardiac biosignals using ordinal pattern statistics and symbolic dynamics," *Computers in Biology and Medicine*, vol. 42, no. 3, pp. 319–327, 2012.
- [20] 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.
- [21] A. Ravelo-García, J. Navarro-Mesa, U. Casanova-Blancas et al., "Application of the permutation entropy over the heart rate variability for the improvement of electrocardiogram-based sleep breathing pause detection," *Entropy*, vol. 17, no. 3, pp. 914–927, 2015.
- [22] Y. Yin and P. Shang, "Weighted multiscale permutation entropy of financial time series," *Nonlinear Dynamics*, vol. 78, no. 4, pp. 2921–2939, 2014.
- [23] D. Cai, C. Zhang, and X. He, "Unsupervised feature selection for multi-cluster data," in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '10)*, pp. 333–342, ACM, Washington, DC, USA, July 2010.
- [24] M. Kang, M. R. Islam, J. Kim, J.-M. Kim, and M. Pecht, "A hybrid feature selection scheme for reducing diagnostic performance deterioration caused by outliers in data-driven diagnostics," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 5, pp. 3299–3310, 2016.
- [25] "Gas sensor arrays in open sampling settings Data set," 2016, <https://archive.ics.uci.edu/ml/datasets/Gas+sensor+arrays+in+open+sampling+settings>.
- [26] Q. Yang and J. Wang, "Multi-level wavelet shannon entropy-based method for single-sensor fault location," *Entropy*, vol. 17, no. 10, pp. 7101–7117, 2015.
- [27] J. Rodriguez, A. Perez, and J. Lozano, "Sensitivity analysis of k-fold cross validation in prediction error estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 3, pp. 569–575, 2010.
- [28] M. Kang, J. Kim, J.-M. Kim, A. C. C. Tan, E. Y. Kim, and B.-K. Choi, "Reliable fault diagnosis for low-speed bearings using individually trained support vector machines with kernel discriminative feature analysis," *IEEE Transactions on Power Electronics*, vol. 30, no. 5, pp. 2786–2797, 2015.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

