

Article

A Novel Arc Fault Detection Method Integrated Random Forest, Improved Multi-scale Permutation Entropy and Wavelet Packet Transform

Zhendong Yin, Li Wang *, Yaojia Zhang and Yang Gao

Department of Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China; zd_yin@nuaa.edu.cn (Z.Y.); zhangyaojia@nuaa.edu.cn (Y.Z.); gaoyang0313@nuaa.edu.cn (Y.G.)

* Correspondence: liwang@nuaa.edu.cn; Tel.: +86-025-84893497

Received: 13 March 2019; Accepted: 29 March 2019; Published: 2 April 2019



Abstract: Arc faults are one of the important causes of electric fires. In order to solve the problem of randomness, diversity, the concealment of series arc faults and to improve the detection accuracy, a novel arc fault detection method integrated random forest (RF), improved multi-scale permutation entropy (IMPE) and wavelet packet transform (WPT) are designed. Firstly, singular value decomposition (SVD) was applied to filter the current signal and then the high-dimensional fault features were constructed by extracting IMPE, the wavelet packet energy and the wavelet packet energy-entropy. Afterward, the high-dimensional fault features were employed to train the RF to realize the arc fault detection of different load types and the experimental results verify the effectiveness of the arc fault detection method designed in this paper. Finally, the comparative experiments demonstrates that the RF shows better performance in arc fault detection compared to the back-propagation neural network (BPNN) and least squares support vector machines (LSSVM), and that the experiments of transient events indicate that RF is able to effectively avoid incorrectly detecting different load types during the start operations and stop operations.

Keywords: arc fault; multi-scale permutation entropy; random forest; wavelet packet transform; singular value decomposition

1. Introduction

The level of residential electrification has been greatly raised along with the social economic level growth as well as scientific and technological progress. The increase of household appliances not only increases the load of electricity, but also increases the potential safety hazards of electrical systems.

The United States fire service [1] reports that there have been about 380,200 residential building fires in the United States during 2013 and 2015. These fires caused a total of 2695 deaths, about 12,000 people were injured, the economic losses amounted to \$700 million and it caused great harm to society. Electrical fires account for 30.4% of the total number of residential building fires and 46% of extraordinarily serious residential building fires [2], which are one of the most frequent types of residential building fires. Electrical fires are caused by the chemical reaction between the ignition source (hot spot, arc, spark), oxidizer (oxygen) and combustible (close to the fire source in the line). Through the research of residential building fires data, it was found that the causes of electrical fire accidents [1] are usually arcs, electric leakages, over-currents and overheating of electrical equipment. Circuit breakers in the circuit are able to realize the protection of over-currents, electric leakages and other electrical faults, but lack of the ability to effectively realize the detection of arc faults. The fires caused by arc faults account for 82% [3] of electric fires, posing a great threat to the safety of the people's lives and property.

Arcs are luminous electrical discharge between electrodes, which are able to generate a huge amount of heat. The temperature of the arc column is capable of reaching 20,000 K [4] and nearby combustible material can be easily ignited. There are three types of arc faults [5]: ground arc faults, series arc faults and parallel arc faults. The effective value of current is higher than that under normal operations when parallel arc faults and grounding arc faults occur in the circuit. These cases are equivalent to short circuit faults and the breaker is competent in protecting the circuit in time. However, the effective current value in the line is less than that during normal operations when arc faults occur, which are hard to identify [6,7]. Therefore, how to effectively identify series arc faults in the line is an urgent problem needing to be solved, which is of vital practical significance.

There has been much research on arc fault detection methods at present. Some researchers have studied the mathematical modeling of arcs by collecting experimental data [8–10] and thus obtained the parameters of the arc model. However, the research on arc fault models has merely stayed in the simulation stage. There will be strong light and a high temperature, electromagnetic radiation, noise and other physical phenomena when an arc fault occurs, therefore, researchers are able to identify arc faults by detecting the occurrence of these physical phenomena [11–13]. Based on the features extracted from the electromagnetic radiation signal, least squares [14] and extreme learning machine [15] are used to identify arc faults by collecting the voltage signals in the circuit. However, the arc faults are very random and their locations are uncertain. It is difficult to apply the detection methods which are based on sound, light, heat, voltage and electromagnetic signals in the actual environment because the sensors are placed in fixed positions.

In order to solve the problems of fault detection in terms of location and isolation on the transmission lines, a lot of research has been conducted. A fault location method that uses the faulted negative-sequence voltage and locates the faulted sections [16] was proposed by applying the relationship between the fault distance and the clustered measurement groups. For a load change that would not lead to a 180° phase angle change, the current-only method [17] was able to detect the overcurrent fault without requiring the measurement of voltage. The EnKF-based approach [18] is able to accurately locate the short-circuit faults on transmission lines by ignoring the foreknowledge of either the fault type and an approximate guess of the fault location. Arc faults are very random and intermittent when compared to short-circuit faults and high-impedance faults, therefore, it is unsuitable to apply the methods for fault detection and the detection of their location in the transmission line mentioned above in the detection of arc faults.

Nowadays, the fault detection method of series arcs based on the line current signal is getting a lot of attention. Lu [19] took the root mean square of the current signal as the fault feature. This method is simple and shows good performance in real-time, but determining the threshold of different loads is difficult. The harmonic power and the peak-to-peak value of the current are extracted under the conditions of linear loads and nonlinear loads, and the Mahalanobis distance method was employed as a classifier to realize the discrimination of AC arc faults [20]. Based on the multi-resolution feature of wavelet decomposition, Zhang [21] analyzed the components of the current signal at different frequency bands by wavelet transform (WT) and the wavelet energy was extracted as the fault feature. Chirp zeta transform (CZT) [22] and WT [23] were utilized to analyze the spectrum of the current signal and extract the time-frequency domain characteristics for arc fault detection. However, the microwave oven, computer and other nonlinear loads will produce a wealth of harmonic components under normal operation and this method is prone to misjudgment if only a time-frequency component is used as the fault characteristic. By obtaining the wavelet coefficient sequence, the characteristic matrix based on singular value decomposition is defined and used as the basis of series arcing fault detection [24]. The different load types are of different characteristics, therefore, the threshold is hard to determine. The image processing method is another way to realize the detection of arc faults and the gray level-gradient co-occurrence matrix was proposed to extract the features [25]. However, the process of constructing the gray image is able to be undertaken with the loss of detailed information about the current. Sparse representation is able to directly reduce the dimension of the

original signal to construct the high-dimensional feature [26] and a neural network is applied to realize the arc fault detection. However, the overlapping points of the features in the feature space make the features unable to effectively represent the characteristics of the arc faults under different load types. The two-dimensional identification features are constructed by extracting the flat shoulder phenomenon and the information dimension of the current signal [27] but non-linear loads also have flat shoulder phenomena, which are likely to cause misjudgments. Based on the fault features extracted from the current signals, some researchers have applied machine learning methods such as support vector machine (SVM) [28], LSSVM [29], BPNN [30] and the Kalman filter (KF) [31] to detect arc faults. However, SVM, LSSVM and KF are binary-classifiers and several classifiers need to be trained at the same time when handling multiple classification problems. The BPNN is prone to over-fitting and falling into the local extremum.

Permutation entropy (PE), a method to detect the random fluctuation and dynamic change of time series [32], has been successfully applied in areas such as mechanical fault diagnosis [33,34], biomedicine [35,36] and so on. PE is able to measure the complexity of the signal effectively, has excellent anti-noise ability, and is capable of detecting non-linear and non-stationary signals as well. The randomness and complexity of the current signals are enhanced when series arc faults occur in the circuit so that the permutation entropy can be utilized to measure the current change when a series arc fault occurs. The calculation of permutation entropy is based on the single scale time series similar to the traditional parameter extraction method. Costa [37] proposed the method of multi-scale analysis to measure the complexity and randomness of the time series at different time scales so that the important information of the current signal at different time scales can be extracted by using the multi-scale analysis method to calculate the entropy of permutation. However, the multi-scale analysis is the procedure of averaging the original time series within a τ -length window and then downsampling it by a scale factor of τ , which will result in an unstable measurement of the permutation entropy. Azami [38] proposes an improved multi-scale permutation entropy (IMPE) to avoid this problem. Due to the excellent performance of IMPE in extracting the intrinsic features of the signal, this paper utilizes IMPE to be a part of the features for the detection of serial AC arc faults.

WPT [39] is an improved form of wavelet transform (WT) which is able to decompose the signals into a high-frequency part and low-frequency part at the same time and it is a more detailed decomposition method compared to WT. When the signal state changes, the proportion of the signals at each wavelet packet layer in terms of the total energy will change [40]. Compared with normal operations, the energy distribution of signals in different frequency bands will change when series arc faults occur. The energy-entropy of the wavelet packet is suitable for measuring the homogeneity of the energy distribution of multi-layer current signals decomposed by wavelet packet transforms. It is applicable to measure the state change of the current signal state by the energy-entropy of the wavelet packet when series arc faults occur. Therefore, in this paper, wavelet packet energy-entropy is used as one of the fault characteristics to detect serial arc faults.

Random forest (RF) [41] is an effective machine learning method, which can be applied to solve the problem of classification and regression. Random forest is an ensemble learning model based on the decision tree model, which combines Bagging and random subspace theory. RF is robust and is adept in processing high-dimensional features, and it has been widely used in fields such as machinery [42] power electronics [43], image processing [44] and the biological field [45]. There are no reports on the application of RF in series AC arc fault detection. In this paper, the method based on RF for series AC arc fault detection is proposed.

In order to improve the detection efficiency of the series arc fault detection method for different working states of different load types, this paper proposes a novel arc fault detection method by taking advantage of IMPE, WPT, singular value decomposition (SVD) and RF. The experimental results indicate that the working state of different loads of the test set can be accurately detected based on the method proposed by this paper and that the comparison experiment results demonstrate that the

method proposed by this paper has a better performance than the prior methods. This method is also able to avoid incorrect detection in transient event experiments.

The rest of this paper is organized as follows. In Section 2, we introduce the experimental platform and collect the experimental data. Section 3 utilizes the IMPE, WPT and SVD to acquire high-dimensional fault features. In Section 4, the designed RF model is applied to identify the work states of the different load types, the comparison with prior methods is given, and the reliability of RF is further verified by the transient event experiments. Section 5 presents the conclusions.

2. Experimental Platform and Data Acquisition

2.1. Experimental Platform

In order to carry out research on arc fault detection methods, our research group established the arc fault experimental platform by referring to the UL1699 standard, as shown in Figure 1. In this study, the arc generation is of the rod-type. The anode rod made of carbon is held stationary while the cathode rod made of copper is moved apart by a step motor. The two electrodes of the arc generator are closed at the beginning of the experiment. At the same time, the arc generator, power supply and load constitute a series circuit. The parameters of AC power supply are 220V/50 Hz and the data-acquisition board card collects the current data from the circuit through a current sensor at a sampling frequency of 100 KHz. Then the rods of the arc generator are slowly separated at a certain speed by computer software (Labview). There is an arc phenomenon when the two electrodes are separated to a certain extent.

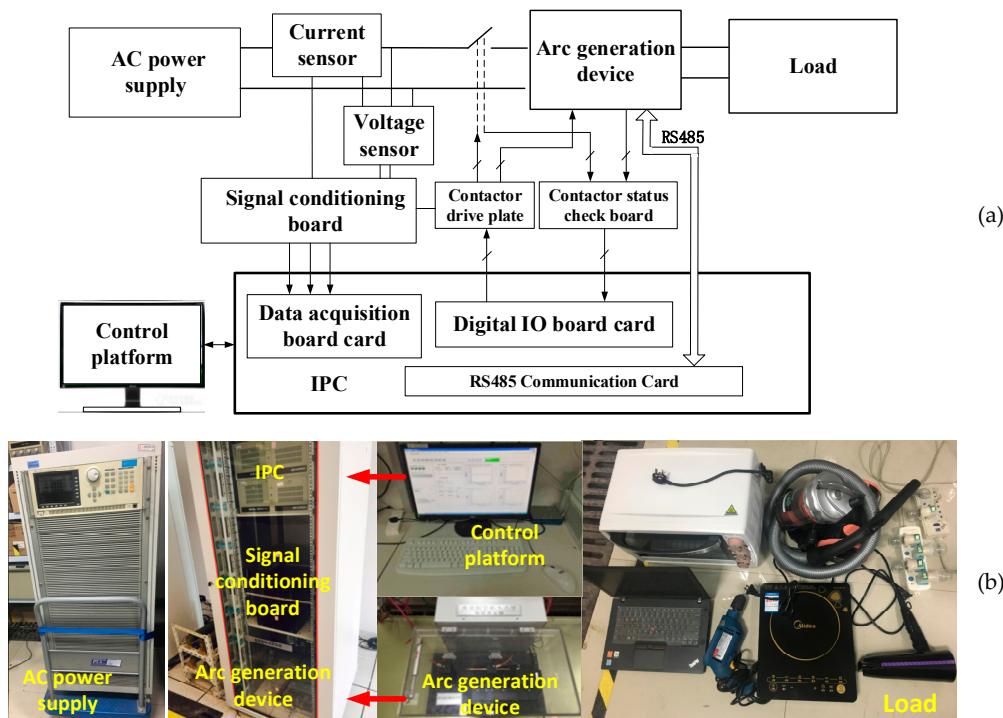


Figure 1. The arc fault experimental platform: (a) schematic; (b) actual platform.

2.2. Experimental Data

Seven load types are used in this paper and the operation conditions of different load types are shown in Table 1. Incandescent lamps and electric ovens are resistive loads. Hairdryers, electric drills and vacuum cleaners are inductive loads. Induction cookers and notebook computers are non-linear loads. The current waveform of the seven load types during normal operations and serial arc fault conditions are shown in Figure 2.

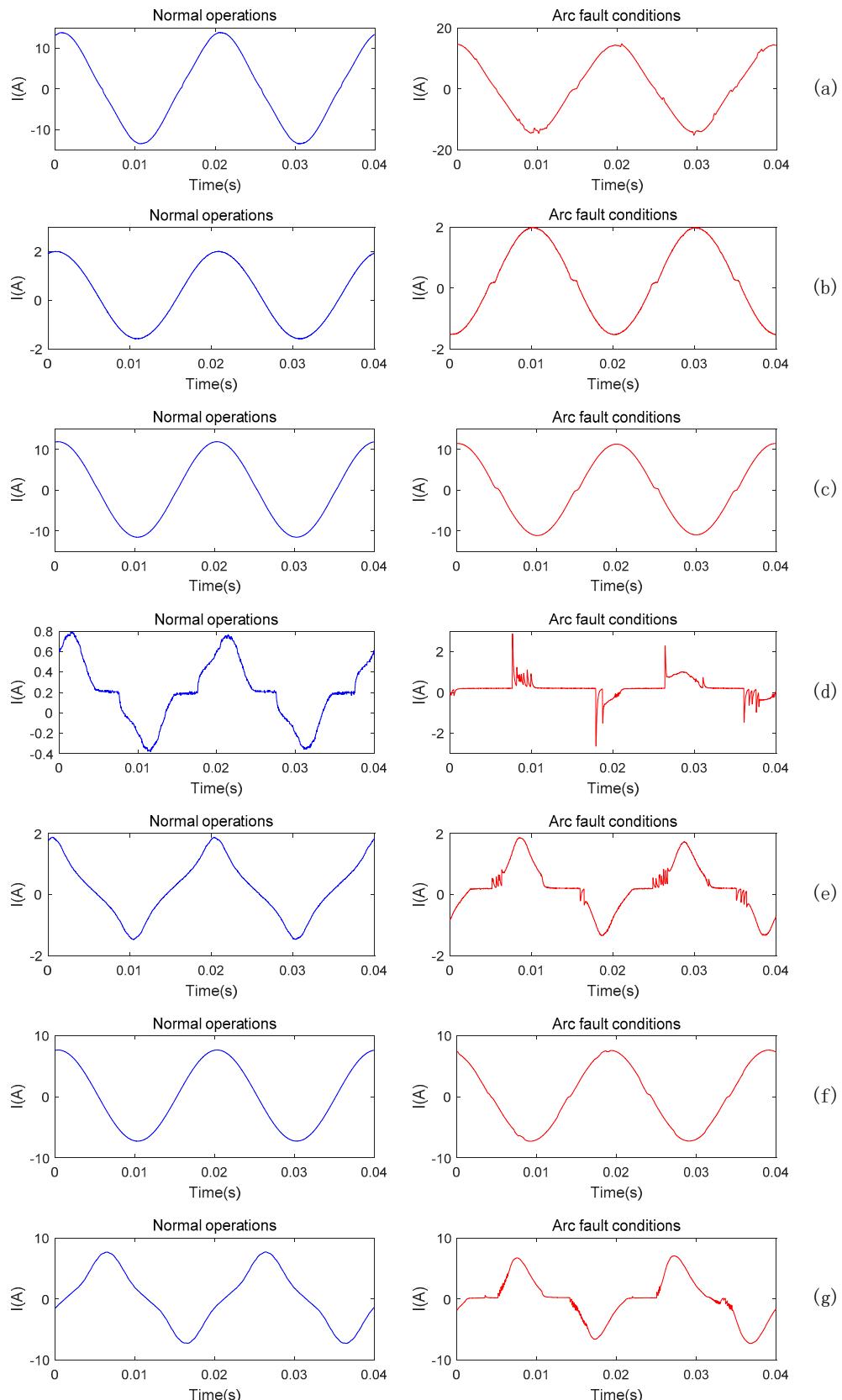


Figure 2. The waveform of different load types during normal operations and serial arc fault conditions. (a) Induction cooker. (b) Incandescent lamp. (c) Hairdryer. (d) Notebook computer. (e) Electric drill. (f) Electric oven. (g) Vacuum cleaner.

Table 1. The operation conditions of the different loads.

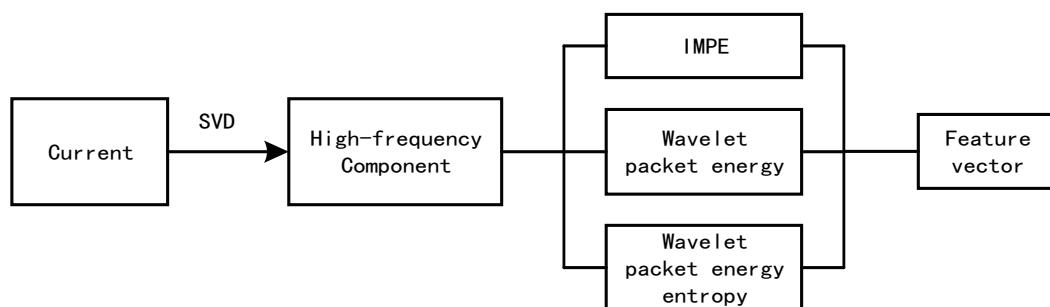
Load Type	Operating Condition	
	Operating Voltage/V	Power/W
Induction cooker	220	2100
Incandescent lamp	220	200
Hairdryer	220	2000
Notebook computer	220	148.3
Electric drill	220	600
Electric oven	220	1200
Vacuum cleaner	220	1200

As seen from Figure 2, the current waveforms of incandescent lamps and electric ovens are sine waves during normal operations and flat shoulder phenomenon occurs during serial arc fault conditions. When the induction cooker operates during normal operations, the current fluctuates slightly near zero during serial arc fault conditions and there are no only flat shoulder phenomenon nor catastrophe points in the current wave shapes. The current wave shapes of the electric drill and vacuum cleaner during normal operations is non-sinusoidal and the flat shoulder phenomenon is more obvious than that of the resistive load during serial arc fault conditions. Besides, there are harmonic components behind the flat shoulder, the current wave shapes of the notebook computer (nonlinear load) are obviously flat shoulder phenomena during normal operations, and the harmonic components of the current are significantly enhanced during the serial arc fault conditions compared to the other load types. The waveform distortion is more serious as well.

The samples are recorded during the stable arc period. Each sample contains 1000 points and the record length is 0.01 s. This paper collects 1400 samples from 7 load types. A total of 700 samples are used as the training set to train the classifiers and another 700 samples are used as the test set to test the performance of the trained classifiers. A total of 200 samples are recorded under each load type, half of the samples are collected during normal operations, and the others are collected during serial arc fault conditions.

3. Feature Extraction

The arc gap will generate a large amount of ionized air when the serial arc faults occur in the circuit and a mass of harmonic components will be introduced into the current. Therefore, the effect of the serial arc fault on the circuit is more reflected in the harmonic components of the current. In this paper, first of all, SVD is utilized to filter the current signal and the harmonic components of the signal are acquired. Then, IMPE, the wavelet packet energy and the wavelet packet energy-entropy are extracted to form the 9-dimensional feature vectors and the feature vectors are the input of the RF. The extraction process is shown in Figure 3.

**Figure 3.** The process of feature extraction.

3.1. SVD

SVD is an effective signal filtering method [46] which has been successfully applied in image processing, medicine, machinery and other fields. The principle of SVD is as follow:

For any matrix H_m whose size is $p \times q$, there are orthogonal matrices U and V

$$H_m = USV^T \quad (1)$$

where, the size of U and V^T are $p \times p$ and $q \times q$, and S is the diagonal matrix with a size of $p \times q$. The main diagonal element of S is λ_i ($i = 1, 2, \dots, k$), $k = \min(p, q)$, λ_i is the singular value of the matrix H_m , and $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k \geq 0$. The following is known from the singular value decomposition theory and the matrix optimal approximation theorem in the Frobenius norm sense: the low-frequency components of the signal are reflected by the bigger singular values (generally, the sum of the first three singular values can account for more than 90% of the sum of the total singular value).

In SVD, the smaller singular values reflect the harmonic components of the signal so that the high pass filtering can be realized by removing the first n large singular values and by performing inverse singular value decomposition. For example, when the load type is of the induction cooker, Figure 4 shows the proportion of different singular values in the sum of the total singular values during normal operations. The first singular value accounts for 93% of the total singular value and the second singular value accounts for 6% of the total singular value. However, the remaining singular value only accounts for 1%. The first 2 singular values are much larger than the last 62 singular values. It is indicated that the first 2 singular values contain more low-frequency information and the last 62 singular values contain more high-frequency information. This paper sets n to 2 in order to realize the high pass filtering of the current. The reason that setting n to 2 will be explained latter.

The proportion of different singular values in the sum of the total singular values

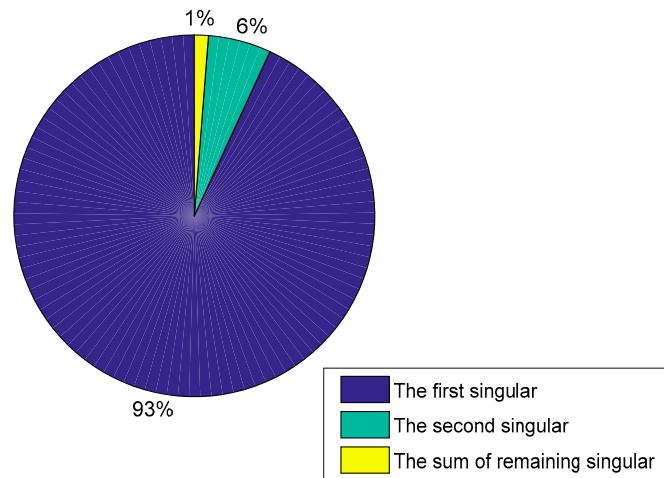


Figure 4. The proportion of different singular values in the sum of the total singular value.

Firstly, the matrix construction method proposed by Reference [47] is applied to convert the current signal into the matrix before filtering the current signal by SVD. This matrix construction method is simple and effective, which is without an adjustable parameter. In References [47,48], the fault features were extracted based on this matrix construction method and the fault detection was successfully achieved. The validity of this matrix construction method is proved as well. The matrix construction process is shown in Figure 5. To obtain an $M \times M$ size matrix, a signal with a length M^2 needs to be obtained from the raw time series. Let $E(i)$, $i = 1, \dots, M^2$, which represents the value obtained from the raw time series. $A(j, k)$, $j = 1, \dots, M$, $k = 1, \dots, M$, which represents the constructed matrix A . E would be driven into M segments. The first segments constitute the first row of A and the

second set of segments constitute the second row of A , and so on. The structure of time series E and matrix A is as follows

$$E = \{a_1, a_2, a_3, \dots, a_{MM}\}, A = \begin{bmatrix} a_1 & a_2 & \dots & a_M \\ a_{M+1} & a_{M+2} & \dots & a_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M(M-1)} & a_{M(M-1)+1} & \dots & a_{MM} \end{bmatrix} \quad (2)$$

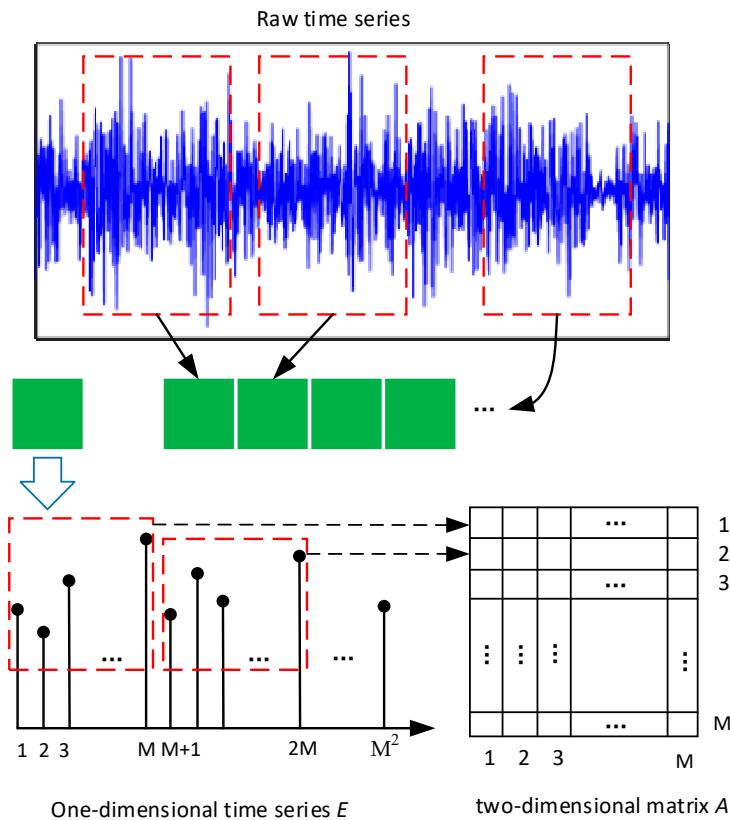


Figure 5. The matrix construction process.

For a current signal, the 64×64 matrix A is constructed by using Equation (2). So, the 64 singular values were acquired by performing SVD on matrix A . After removing the first 2 singular values, the remaining 62 singular values were used to reconstruct the signal and the high pass filtering process is completed. Figure 6 shows the waveform of the filtered signal during normal operations and serial arc faults conditions, and the reconstructed signal mainly contains the harmonic components of the signal. The comparison between Figures 2 and 6 indicates that the low-frequency components are removed and high pass filtering is realized when n is set to 2. The features during normal operations are marked differently from those marked during serial arc fault conditions, which also demonstrate that setting n to 2 is appropriate. During normal operations, the fluctuation amplitude of the current waveform is smaller and the random fluctuation is stronger. During serial arc fault conditions, the local tendency of fluctuation is strengthened (more data segments are monotonous), the waveform has a large fluctuation amplitude at the zero rest position and the distortion of the waveform is stronger than that during normal operations.

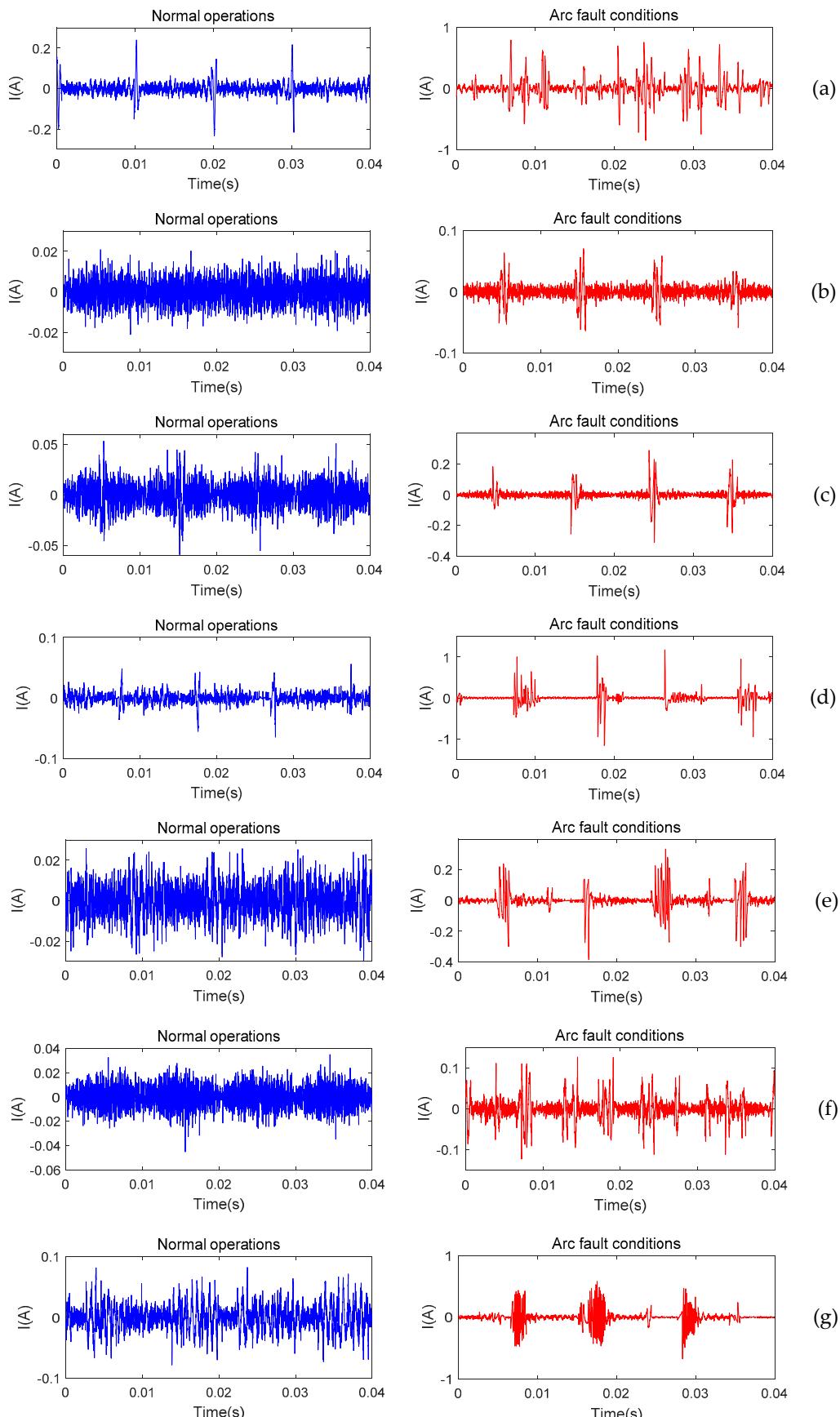


Figure 6. The waveform of different load types during normal operations and serial arc fault conditions after filtering. (a) Induction cooker. (b) Incandescent lamp. (c) Hairdryer. (d) Notebook computer. (e) Electric drill. (f) Electric oven. (g) Vacuum cleaner.

3.2. IMPE

3.2.1. PE

The principle of PE is based on the comparison of the adjacent data points. According to the embedding theorem, the phase space reconstruction of a time series $\{x(i), I = 1, 2, \dots, N\}$ is carried out as follows

$$\begin{aligned} X(1) &= \{x(1), x(1 + \lambda), \dots, x(1 + (m - 1)\lambda)\} \\ &\vdots \\ X(i) &= \{x(i), x(i + \lambda), \dots, x(i + (m - 1)\lambda)\} \\ &\vdots \\ X(N) &= \{x(N), x(N + \lambda), \dots, x(N + (m - 1)\lambda)\} \end{aligned} \quad (3)$$

where, m is the embedded dimension and λ is the time delay.

Rearrange the m data points of $X(i)$ in ascending order

$$\begin{aligned} X(i) &= \{x(i + (j_1 - 1)\lambda) \leq \\ &x(i - (j_2 - 1)\lambda) \leq \dots \leq x(i + (j_m - 1)\lambda)\} \end{aligned} \quad (4)$$

If $x(i - (j_{i1} - 1)\lambda) = x(i - (j_{i2} - 1)\lambda)$, $x(i - (j_{i1} - 1)\lambda) \leq x(i - (j_{i2} - 1)\lambda)$ can be written in the case of $j_{i1} < j_{i2}$. Therefore, for any $X(i)$, a symbol sequence $S(g)$ can be obtained.

$$S(g) = \{j_1, j_2, \dots, j_m\} \quad (5)$$

where, $g = 1, 2, \dots, k$, $k \leq m!$, $S(g)$ is composed of m elements having an $m!$ arrangement and there are $m!$ symbol sequences. We calculate the probability of each symbol sequence P_g ($g = 1, 2, \dots, k$), $\sum_{g=1}^k P_g = 1$, the permutation entropy of time series $X(i)$ can be defined as the Shannon entropy.

$$H'_p(m) = -\sum_{g=1}^k P_g \ln P_g \quad (6)$$

$H'_p(m)$ reaches the maximum $\ln(m!)$ when $P_g = 1/m!$, and $H'_p(m)$ can be standardized by $\ln(m!)$

$$H_p(m) = H'_p(m) / \ln(m!) \quad (7)$$

where, the range of H_p is $[0,1]$, and the size of H_p indicates the complexity and randomness of the time series. The larger H_p is, the more uniform the probability distribution of the permutation pattern of the symbol sequence is and the stronger the uncertainty of the signal.

3.2.2. IMPE

PE is only able to detect the randomness and dynamic mutation of the time series on a single scale. The current signal not only contains important information on a single scale, but also contains important information on other scales because of the high complexity of the current signal under different load types. Multi-scale entropy (MPE) is able to measure the complexity and randomness of the time series at different scales. Therefore, multi-scale analysis of the current signal can excavate deep information on the signal from different time scales.

However, the multi-scale process operation did not make full use of the sample points. For example, Figure 7a shows the coarse-grained procedure of MPE. When the time scale is 3, the first data point of the coarse-grained sequence is able to be acquired by taking the average of x_1, x_2 and x_3 , but the sample point x_2, x_3 and x_4 is not applied to calculate the second data point of the coarse-grained sequence. The second data point of the coarse-grained sequence is calculated directly by using x_4, x_5

and x_6 . The number of data points contained in the coarse-grained sequence will decay exponentially with the increase of the time scales in the time sequence, which makes the calculation of multi-scale entropy unstable. To overcome these problems, Azami [38] proposed the IMPE; the principle is as follows:

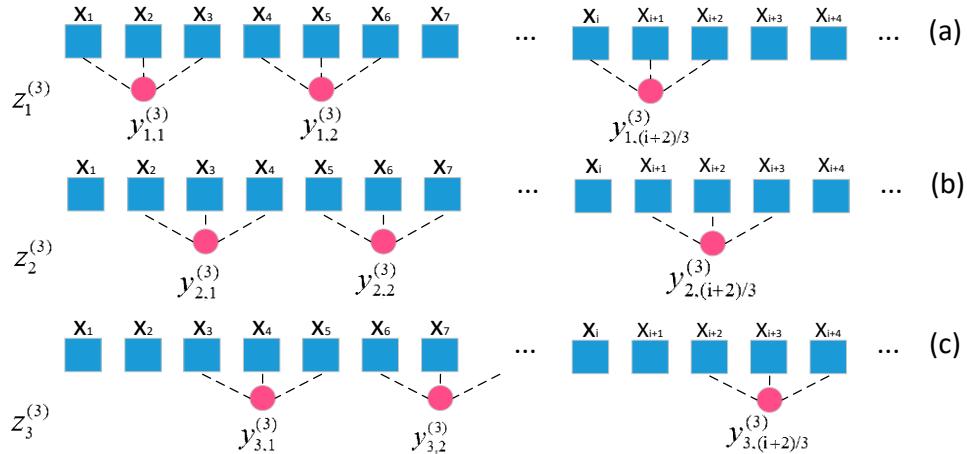


Figure 7. The diagram of the coarse granulation process when $\tau = 3$. (a) The computational process of $z_1^{(3)}$. (b) The computational process of $z_2^{(3)}$. (c) The computational process of $z_3^{(3)}$.

Considering the time series $\{x(i), i = 1, 2, \dots, N\}$, the coarse granulation process of $x(i)$ was carried out to obtain the coarse granulation sequence $z_i^{(\tau)} = \{y_{i,1}^{(\tau)}, y_{i,2}^{(\tau)}, \dots, y_{i,j}^{(\tau)}\}$

$$y_{i,j}^{(\tau)} = \frac{1}{\tau} \sum_{f=0}^{\tau-1} x_{f+i+\tau(j-1)} \quad j = 1, 2, \dots, [N/\tau]; i = 1, 2, \dots, \tau \quad (8)$$

where, $[N/\tau]$ expresses the rounding of N/τ , τ is the scale factor, $\tau = 1, 2, \dots$; it is obvious that the coarse-grained sequence is the original sequence when $\tau = 1$. When $\tau > 1$, we have a different time series $z_i^{(\tau)} | (i = 1, 2, \dots, \tau)$, the length of $z_i^{(\tau)}$ is $[N/\tau]$. Only $z_1^{(\tau)}$ is considered in the MPE method, coarsely grained to a sequence whose length is $[N/\tau]$. Figure 7 shows the coarse granulation process when $\tau = 3$. The calculation of the permutation entropy of the above multi-scale time series is IMPE.

$$IMPE(x, \tau) = \frac{1}{\tau} \sum_{i=1}^{\tau} PE(z_i^{(\tau)}) \quad (9)$$

3.2.3. The Feature Extraction of IMPE

To calculate the IMPE, firstly, two parameters need to be determined: the embedding dimension m and embedding time λ . If the value of m is too small, the reconstructed phase point contains inadequate state information and the algorithm cannot reflect the dynamic characteristics of the time series. If the value of m is too high, the time series will be homogenized and the algorithm will lose the ability to express the subtle changes of the time series [49]. The value of m should be 3–7 [50], because this range can reveal the essential characteristics of the time series [51]. The value of λ has little influence on the calculation of permutation entropy and the phase space points contain the most information of the time series when $\lambda = 1$ [52,53]. Hence, the value of λ in this paper is 1. The calculated result of the change of permutation entropy when m is evaluated from 1 to 9 and $\lambda = 1$ is shown in Figure 8. In the case of $m < 3$, the permutation entropy cannot detect the dynamic characteristics of the current signal. In the case of $m > 7$, the value of the permutation entropy is too small, which homogenizes the time series and causes the loss of the ability to measure the dynamics and complexity of the time series. Therefore, m is set to 6.

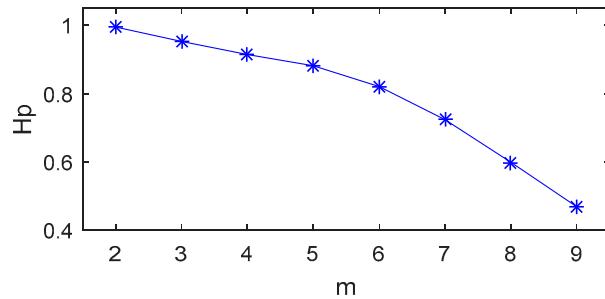


Figure 8. The curve of the calculated value of permutation entropy changing with the embedded dimension.

Using the group of phase-space parameters above, the 8 scales of the IMPE under different electrical load types are calculated, as shown in Table 2 and Figure 9. When the scales are between 1 and 3, the IMPE of the induction cooker, incandescent lamp, hairdryer, laptop, electric drill and electric oven during normal operations is greater than that during the serial arc fault conditions. The reason for this is that during serial arc fault conditions, the fluctuation amplitude of the signal is larger and the trend of the current signal in the local region is stronger than that during normal operations. In other words, there are more data segments which are monotonous during serial arc fault conditions. This leads to the inhomogeneity of the probability distribution of the permutation pattern of the symbol sequences in the phase space. Therefore, the value of IMPE during serial arc fault conditions is lower than that during normal operations. During normal operations, the IMPE of the vacuum cleaner is lower than that during serial arc fault conditions. The reason for this is that the trend of the current signal in the local region is weaker during serial arc fault conditions, which leads to the distribution of the permutation pattern of the symbol sequences being more homogeneous in the phase space. The energy distribution of the wavelet packet components later also explains the intrinsic reason for the distribution of IMPE under different operation conditions and different load types.

Table 2. The permutation entropy of different load types during normal operations and serial arc fault conditions.

Load Type	Operating Condition	Scale ₁	Scale ₂	Scale ₃	Scale ₄	Scale ₅	Scale ₆	Scale ₇	Scale ₈
Induction cooker	Normal	0.8209	0.8248	0.8017	0.7749	0.7416	0.7293	0.7523	0.6635
	Arc	0.6264	0.7586	0.7662	0.7715	0.7381	0.7345	0.7664	0.6553
Incandescent lamp	Normal	0.9220	0.8891	0.8516	0.8072	0.7838	0.7417	0.7207	0.6772
	Arc	0.8468	0.8343	0.8114	0.7928	0.7795	0.7458	0.7048	0.6934
Hairdryer	Normal	0.8893	0.8417	0.8015	0.7697	0.7491	0.7124	0.6929	0.7049
	Arc	0.7951	0.7831	0.7727	0.7712	0.7518	0.7219	0.7101	0.6912
Notebook computer	Normal	0.8311	0.8514	0.8123	0.7732	0.7546	0.7342	0.6927	0.6816
	Arc	0.7186	0.7710	0.7617	0.7689	0.7395	0.6855	0.6893	0.6838
Electric drill	Normal	0.8715	0.8744	0.8376	0.7853	0.7589	0.7403	0.7012	0.6953
	Arc	0.7681	0.8272	0.8168	0.7917	0.7646	0.7342	0.7075	0.7010
Electric oven	Normal	0.9168	0.8786	0.8243	0.8091	0.7698	0.7175	0.7174	0.6905
	Arc	0.7056	0.7439	0.7683	0.7820	0.7512	0.7355	0.7108	0.6972
Vacuum cleaner	Normal	0.8104	0.7945	0.7916	0.7799	0.7396	0.7247	0.7114	0.6985
	Arc	0.9020	0.8619	0.8239	0.7817	0.7314	0.7286	0.7008	0.7049

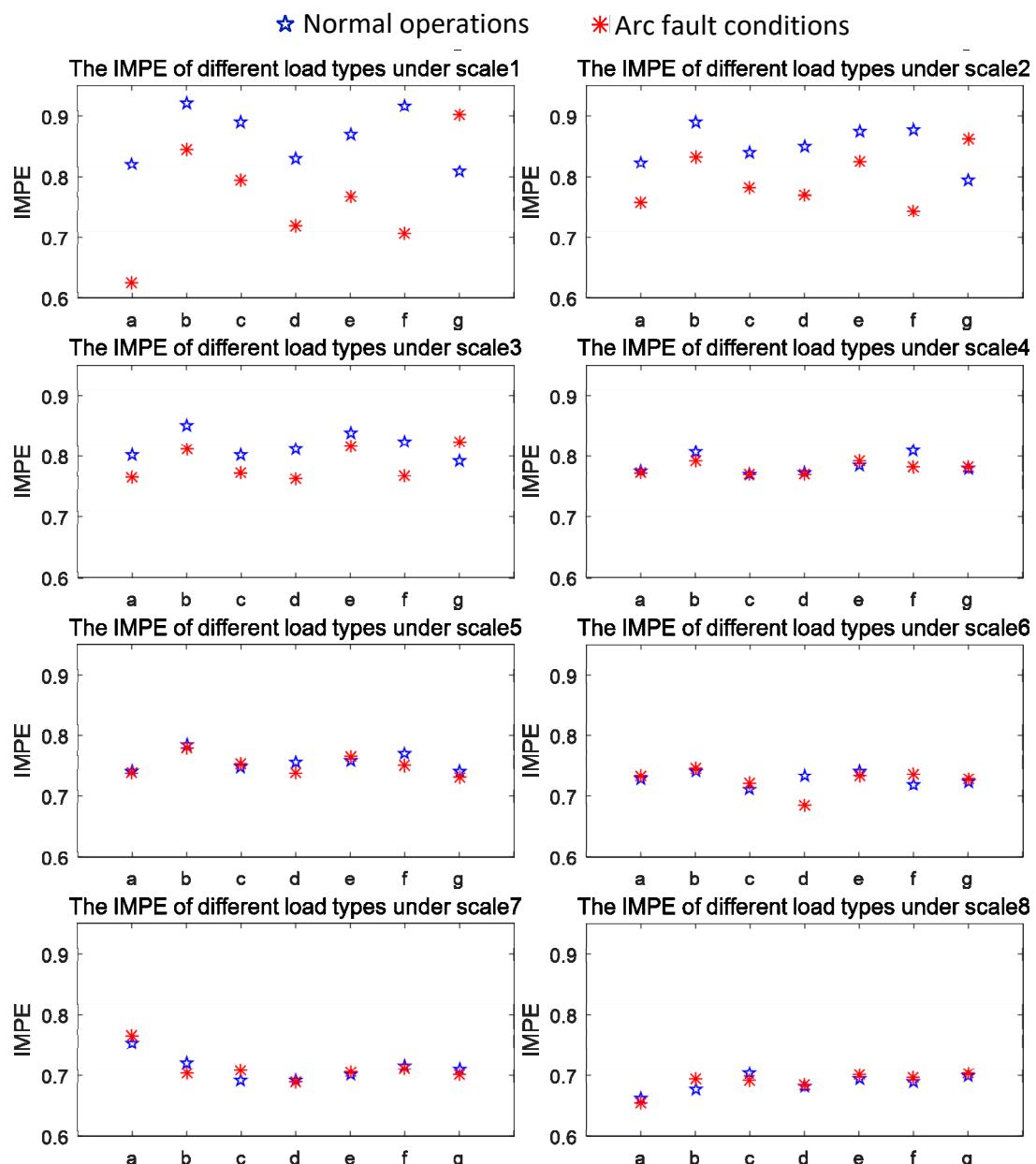


Figure 9. The improved multi-scale permutation entropy (IMPE) of different load types during normal operations and serial arc fault conditions after filtering. (a) Induction cooker. (b) Incandescent lamp. (c) Hairdryer. (d) Notebook computer. (e) Electric drill. (f) Electric oven. (g) Vacuum cleaner.

However, the values of IMPE after the 3rd scale do not conform to the regularity shown in the first 3 scales. The values of IMPE after the 3rd scale during normal operations is close to that during serial arc fault conditions, which are impossible to distinguish from (that is, whether they are from serial arc fault conditions or normal operations). The reason for this is because the oversize time scale makes the time series excessively coarser and the signal loses the main information, making it difficult to achieve effective feature extraction. Therefore, the IMPE in the first 3 scales is adopted as part of the features for the classification later.

3.3. WPT

WPT [39] is an improvement of WT and WPT is able to divide the signal more flexible by further decomposing high-frequency information compared to WT. Thus, the resolution of the high-frequency

information of WPT is improved. The diagram of the three-layer wavelet packet decomposition of the signal is shown in Figure 10. Where H represent the high-frequency components and L represents the low-frequency components. The signal is decomposed into 8 frequency bands after three-layer wavelet packet decomposition and single signal reconstruction.

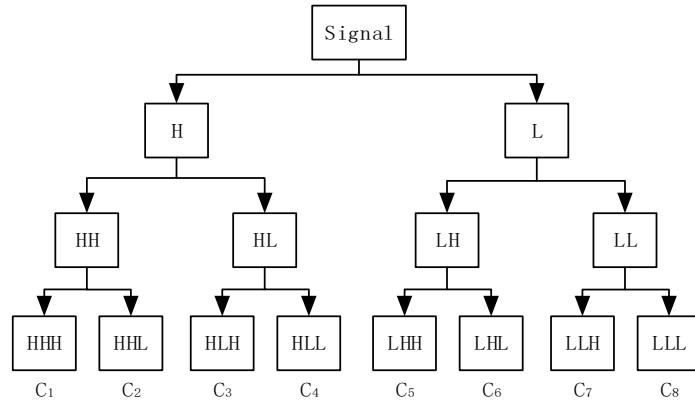


Figure 10. The decomposed diagram of the three layers wavelet.

In general, the wavelet packet components C_j ($j = 1, \dots, 2^n$) are obtained after the signal is decomposed into the 2^n frequency band by n -layer WPT. Let $E_j = |C_j|^2$ be the power of C_j . $e_j = E_j/E$ is the percent of the energy of C_j in the whole signal energy $E = \sum_{i=1}^n E_i$. Thus, $\sum e_j = 1$ and the definition of the wavelet packet energy-entropy is shown in Equation (10). In essence, the energy-entropy of WPT reflects the uniformity of the energy distribution of the wavelet packet components. When the energy distribution of the wavelet components is more uniform, the entropy value H is larger.

$$H = -\sum_{j=1}^{2^n} e_j \log_2(e_j) \quad (10)$$

The energy of the signal in different frequency bands will change when the serial arc fault occurs, which results in the changing of the energy of different WPT components and the energy-entropy of WPT will change at the same time. Therefore, the energy-entropy of WPT can be considered a part of the fault feature for serial arc fault detection. In this paper, the db3 wavelet basis is used to realize the 2-layer wavelet packet decomposition for the current signal. Therefore, 4 wavelet packet components (C_1, C_2, C_3 and C_4) are obtained. The percent of the energy of C_j in the whole signal energy and the wavelet packet energy-entropy are shown in Table 3, Figures 11 and 12.

Table 3. The percent of the energy of C_j in the whole signal energy and the wavelet packet energy-entropy.

Load Type	Operating Condition	C1	C2	C3	C4	Wavelet Energy-Entropy
Induction cooker	Normal	0.9027	0.0586	0.0067	0.0320	0.5805
	Arc	0.9612	0.0236	0.0039	0.0113	0.2866
Incandescent lamp	Normal	0.7541	0.1622	0.0266	0.0570	1.1076
	Arc	0.9172	0.0445	0.0117	0.0266	0.5286
Hairdryer	Normal	0.6285	0.1042	0.0577	0.2096	1.4709
	Arc	0.8822	0.0637	0.0133	0.0408	0.6837
Notebook computer	Normal	0.9100	0.0512	0.0135	0.0252	0.5612
	Arc	0.9725	0.0192	0.0024	0.0059	0.2132
Electric drill	Normal	0.8855	0.0668	0.0180	0.0296	0.6709
	Arc	0.9403	0.0357	0.0100	0.0140	0.4078
Electric oven	Normal	0.2544	0.2140	0.2488	0.2828	1.9930
	Arc	0.9520	0.0262	0.0063	0.0155	0.3445
Vacuum cleaner	Normal	0.9265	0.0380	0.0103	0.0253	0.4832
	Arc	0.8149	0.1504	0.0106	0.0241	0.8509

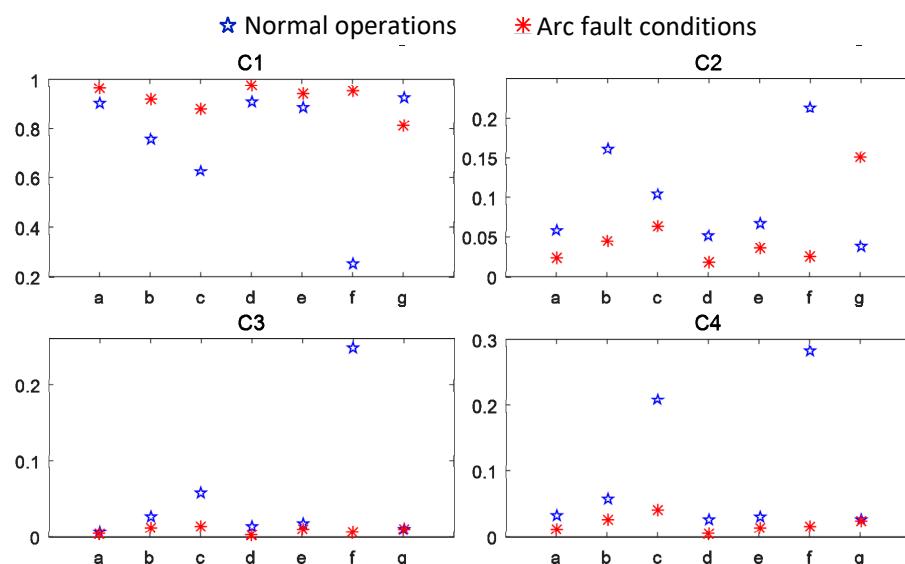


Figure 11. The percent of the energy of C_j in the whole signal energy of different load types during normal operations and serial arc fault conditions (Y axis indicates the percent of the energy of C_j in the whole signal energy). (a) Induction cooker. (b) Incandescent lamp. (c) Hairdryer. (d) Notebook computer. (e) Electric drill. (f) Electric oven. (g) Vacuum cleaner

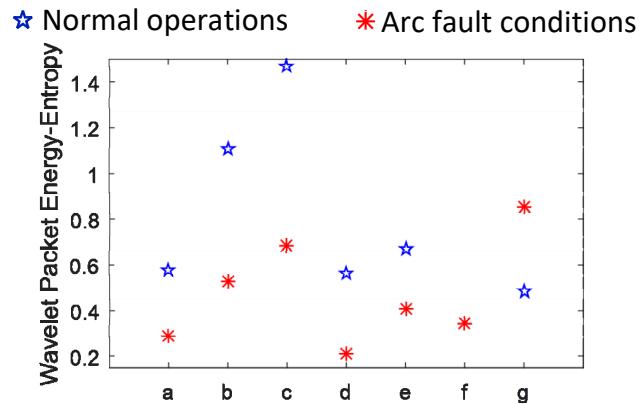


Figure 12. The wavelet packet energy-entropy of different load types during normal operations and serial arc fault conditions. (a) Induction cooker. (b) Incandescent lamp. (c) Hairdryer. (d) Notebook computer. (e) Electric drill. (f) Electric oven. (g) Vacuum cleaner.

When the arc fault occurs under load types such as induction cookers, incandescent lamps, hair dryers, notebook computers, electric drills or electric ovens, the energy of the current signal (the signal in Figure 6) is more concentrated in component C1 (the low-frequency band). As a result, the value of the wavelet packet energy-entropy during normal operations is less than that during serial arc fault conditions. When the load type is vacuum cleaner, the energy of the C1 component during serial arc fault conditions is less than that during normal operations, but the energy of component C2 and C3 is bigger than that during normal operations, which makes the energy distribution of the whole wavelet packet component more uneven. Additionally, the value of the wavelet packet energy-entropy during normal operations is bigger than that during serial arc fault conditions. Therefore, the wavelet packet energy and wavelet packet energy-entropy can effectively distinguish between the serial arc operation condition and normal operations at different loads, which makes it suitable to be part of the characteristics for later detection. When the load is vacuum cleaner, the high-frequency component is stronger during serial arc fault conditions compared to that during normal operations, which will lead to the increase in the inhomogeneity of the permutation pattern's distribution of the symbol sequences in the phase space. This reveals the reason why the value of IMPE will decrease during serial arc fault conditions compared to that during normal operations.

4. The Detection of Serial Arc Fault

4.1. RF

RF is an ensemble algorithm that is composed of multiple decision trees [41] and the detection result of RF is determined by decision tree voting. The construction of the random forest classifier is mainly divided into the following three steps:

- 1) Adopting the bootstrap resampling method to extract the J training dataset from the original data set and generating J decision trees;
- 2) At each internal node of the tree, f features are randomly selected from F features ($f \leq F$) as candidate features. According to the principle of the minimum non-purity of nodes, an optimal feature is selected from f candidate features for the splitting and growth of the node;
- 3) The decision trees obtained by training are used to constitute the random forest classifier, which classifies the new dataset according to the voting results of the decision trees.

4.2. Analysis of detection results

In this paper, the 8 dimensional features obtained based on the training set are used to train the RF and then the RF is utilized to realize the detection of arc faults. The prior methods, BPNN and LSSVM, are used as the comparison.

The number of decision trees in RF affects the performance of the detection [41]. The more decision trees there are, the higher the diversity of the classifiers that is guaranteed. After the number of decision trees reaches a certain number, the performance tends to be stable, but the training time and the complexity of algorithm increases. If the number of decision trees is too small, it will lead to poor performance and a large detection error. First of all, the influence of the number of decision trees on the detection accuracy is studied in order to select the appropriate number of decision trees. The change range of the number of decision trees is set as 1–40 and the step size is 1. Figure 13 is the change curve of the training time (Figure 13a) and detection accuracy (Figure 13b) with the increase in the number of decision trees. As shown in Figure 13a, the training time gradually increases with the increase in the number of decision trees. In general, the training time and the number of decision trees present the relationship of a linear function. As shown in Figure 13b, the detection accuracy increases along with the increase in the number of decision trees in the first half of the curve. When the number of decision trees reaches 23, the detection accuracy is 96.42%. As the number of decision trees continues to increase, the detection accuracy is stabilized between the range of 96% and 97%, but the training time will increase in the form of a linear function. Therefore, the number of decision trees is 23 by comprehensively considering the detection accuracy and training time.

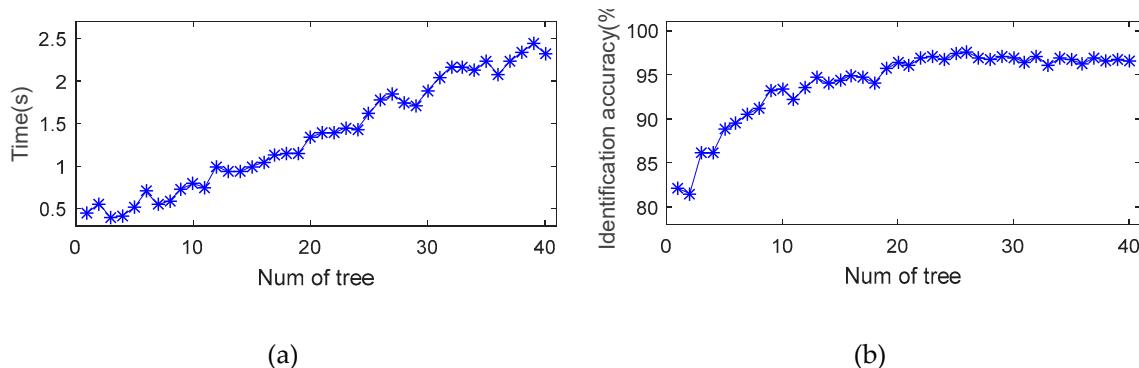
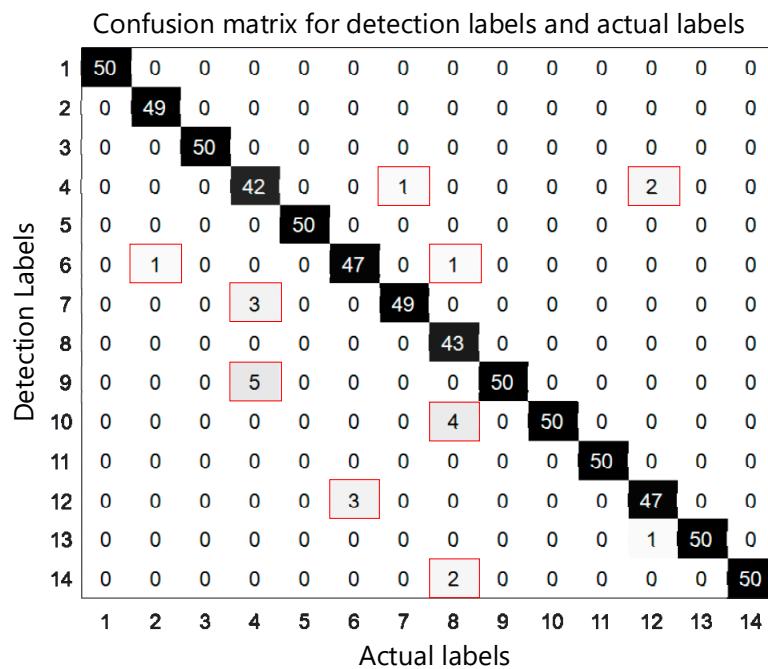


Figure 13. The curve of training time-13 (a) and detection accuracy-13 (b) with the increase in the number of trees.

The trained RF was used to detect the test set. The allocation of the label and detection results of various loads in different operating conditions are shown in Table 4. Figure 14 is the confusion matrix of the detection results. In Figure 14, the horizontal axis and vertical axis represent the actual labels and the labels of detection, respectively. For example, there are 45 samples in the test set which are accurately detected when the actual label is 4. A total of 3 samples and 5 samples in the test set are detected into label 7 and label 9, respectively. In other words, 8 samples are mistakenly detected.

Table 4. The detection results of the random forest (RF).

Load Type	Operating Condition	Label	Number of Test Set	Number of Correctly Detected	Detection Accuracy/%
Induction cooker	Normal	1	50	50	100
	Arc	2	50	49	98
Incandescent lamp	Normal	3	50	50	100
	Arc	4	50	42	84
Hairdryer	Normal	5	50	50	100
	Arc	6	50	47	94
Notebook computer	Normal	7	50	49	98
	Arc	8	50	43	86
Electric drill	Normal	9	50	50	100
	Arc	10	50	50	100
Electric oven	Normal	11	50	50	100
	Arc	12	50	47	94
Vacuum cleaner	Normal	13	50	50	100
	Arc	14	50	50	100

**Figure 14.** The confusion matrix of the test data in Random Forest (RF).

We are able to come to the following conclusion:

- 1) The total detection accuracy is over 90%, reaching 96.71% (677/700), indicating that the trained RF is able to effectively detect the arc fault of different load types;
- 2) During normal operations: the detection accuracy reached 100% when the load type was one of the following: induction cooker (label 1)、incandescent lamp (label 3)、hairdryer (label 5)、electric hand drill (label 9)、electric oven (label 11) and vacuum cleaner (label 13). Only one sample was misjudged—the notebook computer (label 7)—and this condition is detected as the normal operation of the hairdryer, with a detection accuracy of 98%. The total detection accuracy under different loads during normal operations is 99.71%. This shows that the trained RF works well in identifying the different conditions of each load based on the high-dimensional fault features extracted;
- 3) During the serial arc fault conditions: the detection accuracy is 100% when the load types are electric drills (label 10) and vacuum cleaners (label 14). The accuracy is 98%、84%、94%、86% and 94% when the load is the induction cooker, incandescent lamp, hairdryer, notebook computer and electric oven, respectively. The total detection accuracy during serial arc fault conditions is 93.71%,

which is lower than that during normal operations. This is because the arc is a kind of gas discharge phenomenon with an unstable state and the random fluctuation of the current in the circuit is enhanced simultaneously during serial arc fault conditions. Thus, the process of feature extraction and state detection will be disturbed. Consequently, the total detection accuracy of the loads during serial arc fault conditions is significantly lower than that during normal operations.

4.3. Comparison with Prior Methods

In order to further verify the performance of the trained RF, this section gives a comparison with the prior methods: BPNN and LSSVM. BPNN and LSSVM are trained and tested using the same dataset as RF.

BPNN [30] is composed of an input layer, hidden layer and output layer. The neurons of each two adjacent layers are connected by a weight and the neurons in the same layer are not connected. A back-propagation algorithm is used to adjust the weights between neurons in the network during the training stage and the output of the network is able to approximate the actual output. In this paper, the features are 8 dimensions and there are 14 states based on the 7 load types. Therefore, the number of neurons in the input layer and output layer are 8 and 14 respectively. The learning rate and the number of hidden layer nodes are undetermined parameters. The number of nodes in the hidden layer affects the performance of the neural network. If the number of hidden layer nodes is too few, the system error of the network is independent to the training sample and it will result in the loss of its generalization ability. If the number of hidden layer nodes is too many, the complexity of the network will increase. Therefore, the training time will be too long and there will be an "overfitting" phenomenon. If the learning rate is too low, the training cost will increase. If the learning rate is too high, the training will not converge or even diverge.

SVM is a method based on statistical theory, which uses structural risk minimization (SRM) to construct the optimal separable hyperplane by mapping linearly indivisible data to high-dimensional space through the inner product kernel. LSSVM is an improved form of SVM [29]. LSSVM transforms the quadratic constraint problem of the SVM inequality optimization into linear problem solving and greatly increases the computational efficiency. A single LSSVM is able to solve the binary classification problem. The arc fault detection in this paper belongs to the multiple classification problem. Here, one-versus-one coding is used to construct multiple LSSVMs to realize multiple classification. Since there are 14 working states for the 7 loads in this paper, it is necessary to build $14 \times (14-1)/2 = 91$ LSSVMs. The kernel parameter sig^2 and penalty factor gam in the LSSVM affect the detection performance [29]: sig^2 mainly affects the distribution complexity of the sample data in the high-dimensional feature space. The role of gam is mainly reflected in the adjustment of the confidence range and risk ratio in the eigenspace.

Here, in order to avoid blindly selecting the parameters, the cross-validation method is adopted to optimize the parameters of BPNN and LSSVM. The number of hidden layer nodes and the learning rate of BPNN are 190 and 0.16, respectively. Due to the limited space in this paper, the parameters of 91 LSSVMs are not listed.

Table 5 shows the detection result of different classifiers. Figures 15 and 16 show the detection confusion matrix of LSSVM and BPNN, respectively. The following can be known:

Table 5. The detection result of the different classifiers.

Classifier	The Number of Test Data	The Number of Correctly Detected	Detection Accuracy under Normal Operations (%)	Detection Accuracy under Serial arc Fault Conditions (%)	Detection Accuracy (%)
BPNN	700	621	94.57	80	88.71
LSSVM	700	644	95.14	89.71	92.43
RF	700	677	99.71	93.71	96.71

Confusion matrix for detection labels and actual labels														
Detection Labels	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	46	0	0	0	0	3	0	0	0	3	0	3	0	0
	0	48	0	0	0	0	0	8	0	0	0	0	0	0
	0	0	50	1	0	0	0	0	0	0	0	0	0	0
	0	0	0	34	0	0	6	0	2	0	0	2	0	0
	0	0	0	0	50	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	47	0	0	0	0	0	11	0	0
	0	0	0	12	0	0	40	1	3	0	0	1	0	0
	0	2	0	0	0	0	0	36	0	0	0	0	0	0
	0	0	0	3	0	0	2	0	45	0	0	0	0	0
	4	0	0	0	0	0	0	4	0	47	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	50	0	0	0
	0	0	0	0	0	0	0	1	0	0	0	28	0	0
	0	0	0	0	0	0	2	0	0	0	0	5	50	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	50

Figure 15. The confusion matrix of the test data (BPNN).

Confusion matrix for detection labels and actual labels														
Detection Labels	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	49	0	0	0	0	3	0	0	0	0	0	0	0	0
	0	49	0	0	0	0	0	3	0	0	0	0	0	0
	0	0	50	0	0	1	0	0	0	0	0	0	0	0
	0	0	0	36	0	1	13	1	0	0	0	2	0	0
	0	0	0	0	50	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	43	0	0	0	0	0	0	0	0
	0	0	0	11	0	0	36	0	1	0	0	0	0	0
	0	1	0	0	0	0	0	42	0	1	0	0	0	0
	0	0	0	3	0	0	0	0	48	0	0	0	0	0
	0	0	0	0	0	0	1	0	49	0	1	0	1	0
	0	0	0	0	0	0	1	0	0	50	0	0	0	0
	0	0	0	0	0	2	0	1	0	0	0	46	0	0
	0	0	0	0	0	0	0	1	0	0	0	1	50	0
	0	0	0	0	0	0	2	0	0	0	0	0	0	49

Figure 16. The confusion matrix of the test data (LSSVM).

1) The total detection accuracy of BPNN and LSSVM is 88.71% and 92.43%, respectively, and the total detection accuracy of RF is 96.71%. RF has the best detection performance, followed by LSSVM, and BPNN has the worst detection performance;

2) During normal operations, the number mistakenly detected by BPNN was 19 and the detection accuracy was 94.57%. The number mistakenly detected by LSSVM was 17 and the accuracy was 95.14%. The number mistakenly detected by RF was 1 and the detection accuracy was 99.71%. Compared to LSSVM and BPNN, RF is more effective in avoiding false alarms (the normal operation being misreported as an arc fault);

3) During serial arc fault conditions, the number mistakenly detected by BPNN was 60 and the detection accuracy was 80%. The number mistakenly detected by LSSVM was 39 and the detection accuracy was 89.71%. The number mistakenly detected by RF was 22 and the detection accuracy was 93.71%. During serial arc fault conditions, the detection accuracy of the three classifiers is lower than that during normal operations, but RF also has the highest detection accuracy; the LSSVM and BPNN accuracies are less than 90%;

Based on the above analysis, the performance of the RF arc fault detection outperforms the prior methods, such as LSSVM and BPNN.

4.4. The Experiments of Transient Events

The transient events in this section include the start operations and stop operations of different load types. For the transient events beyond the operating conditions in Table 5, we train the RF using the training set by ignoring the load type. If the sample was detected as the normal operation, this sample was correctly detected. If the sample was detected as the serial arc fault condition, this sample was erroneously detected. We record 20 samples of each load type during the start operations or stop operations and the record detecting length of the samples is 0.01s.

Figure 17 shows the waveform of different load types during the start operations and stop operations, and the detection results are shown in Table 6. Only 1 sample of the induction cooker during the stop operations was erroneously detected and the total detection accuracy was 99.28%. The randomness and complexity of the current are at the normal level regardless of whether the loads are under the start or stop operations. The serial arc fault detection method that this paper proposed has an excellent performance in avoiding the incorrect detection of different load types during start and stop operations. We can come to the conclusion that the reliability of this method is trustworthy.

Table 6. Detection results of RF.

Load Type	Operating Condition	Number of Samples	Number of Correctly Detected	Detection Accuracy/%
Induction cooker	Start	20	20	100
	Stop	20	19	95
Incandescent lamp	Start	20	20	100
	Stop	20	20	100
Hairdryer	Start	20	20	100
	Stop	20	20	100
Notebook computer	Start	20	20	100
	Stop	20	20	100
Electric drill	Start	20	20	100
	Stop	20	20	100
Electric oven	Start	20	20	100
	Stop	20	20	100
Vacuum cleaner	Start	20	20	100
	Stop	20	20	100

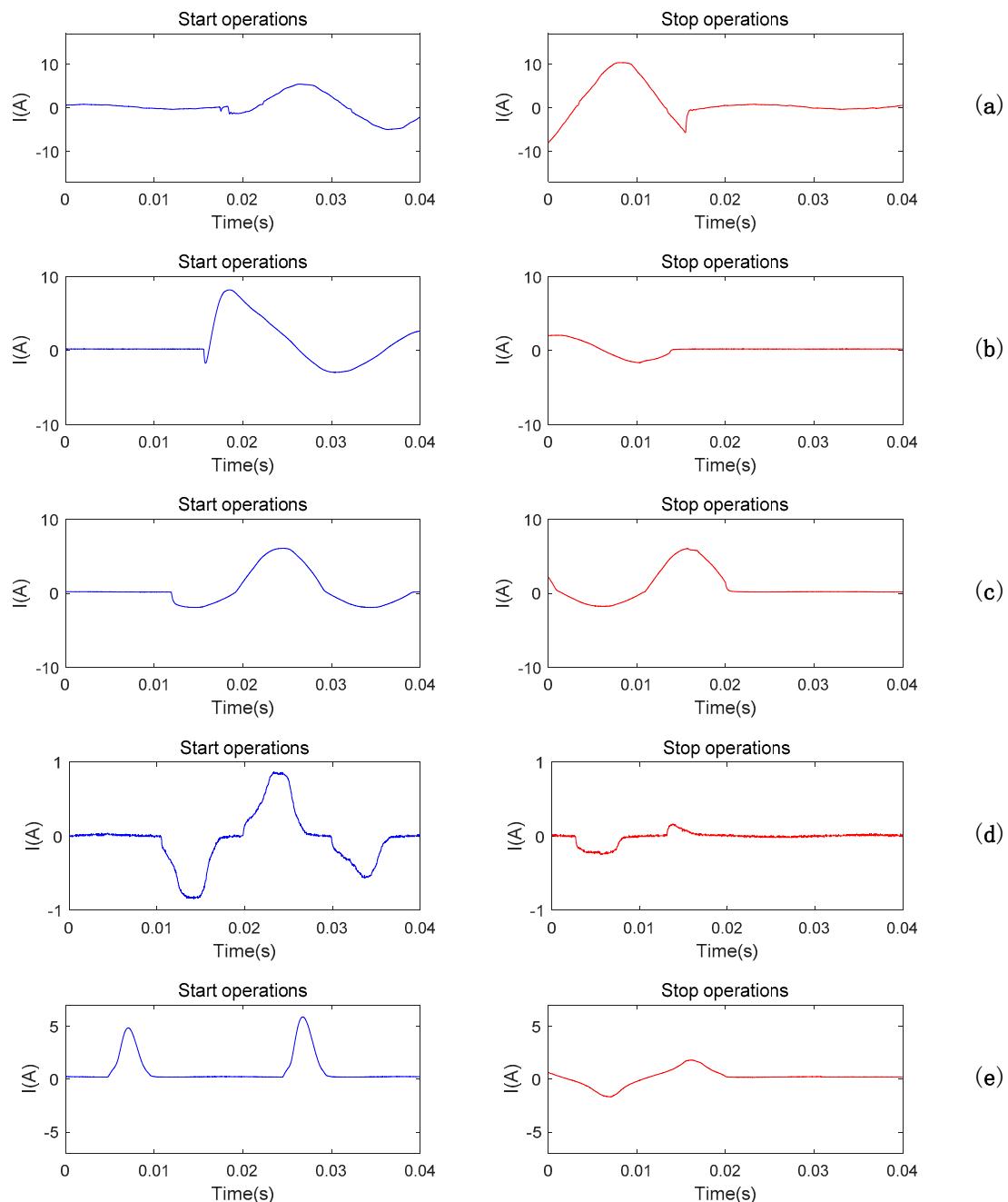


Figure 17. Cont.

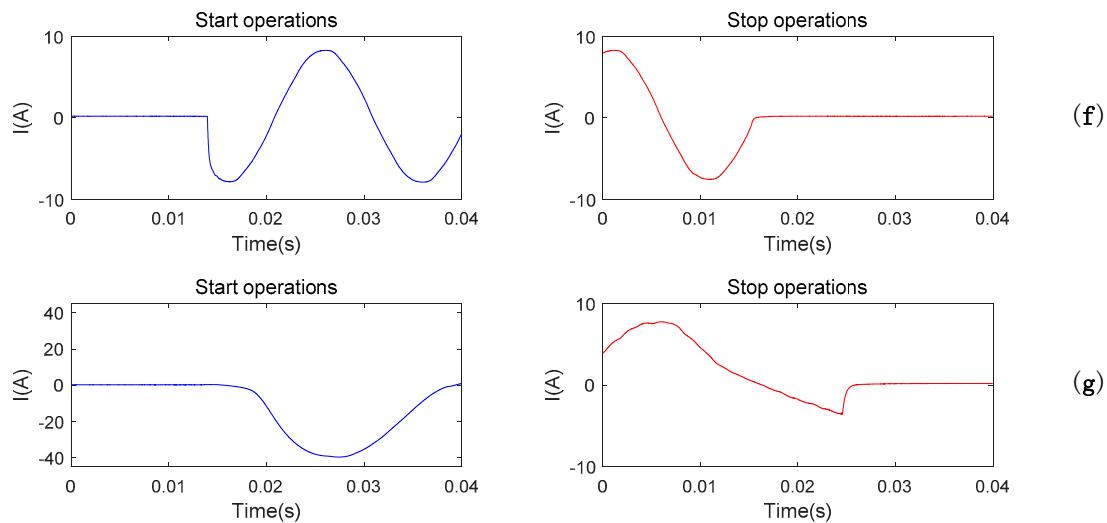


Figure 17. The waveform of different load types during the start operations and stop operations. (a) induction cooker. (b) incandescent lamp. (c) hairdryer. (d) notebook computer. (e) electric drill. (f) electric oven. (g) vacuum cleaner.

5. Conclusions

Arc faults are an important cause of electric fires, which brings a great challenge to residential electricity safety. In this paper, a novel arc fault detection method is designed to accurately detect the arc faults of various load types. The main content is summarized as follows:

1) The characteristic of the arc is mainly reflected in the high-frequency part of the current signal. Based on the matrix construction method, the high-pass filtering of the current signal is realized by using SVD. The matrix construction method is simple and effective and it is without any predefined parameters

2) This paper proposes the application of the IMPE in arc fault feature extraction for the first time. The high-dimensional fault features consist of IMPE, wavelet packet energy and wavelet packet energy-entropy, which is able to reflect the complexity of the signal and the distribution characteristics of the signal in different frequency bands. Additionally, the characteristics of different features during normal operations and serial arc fault conditions are analyzed in detail;

3) This paper presents the application of RF to arc fault detection for the first time. In this paper, based on the high-dimensional fault features extracted, the trained RF detects the normal operations and serial arc fault conditions of different load types effectively. The effectiveness of the RF designed in this paper is proved through comparative experiments, RF is of better performance in arc fault detection compared with BPNN and LSSVM. Whenever the loads are in the state of the start operations or stop operations, the method that this paper proposed is able to effectively avoid incorrect detections.

Author Contributions: Z.Y. designed the research, analyzed the data, implemented the algorithm and wrote the manuscript. L.W., Z.Y., Y.Z. and Y.G. revised the manuscript. All authors have read and approved the final manuscript.

Funding: This work has been supported by National Natural Science Foundation of China (No.51877102).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. United States Fire Administration National Fire Data Center. Residential building fires (2013–2015). *Top. Fire Rep. Ser.* **2017**, *16*, 1–17.
2. Fire Department of Ministry of Public Security. *China Fire Services* (2016); Yunnan People's Publishing House: Kunming, China, 2016; pp. 396–407.

3. United States Fire Administration National Fire Data Center. Residential building fires (2009–2011). *Top. Fire Rep. Ser.* **2014**, *14*, 1–11.
4. Lee, R.H. The Other Electrical Hazard: Electric Arc Blast Burns. *IEEE Trans. Ind. Appl.* **1982**, *246–251*. [[CrossRef](#)]
5. Gregory, G.D.; Scott, G.W. The arc-fault circuit interrupter: An emerging product. *IEEE Trans. Ind. Appl.* **1998**, *34*, 928–933. [[CrossRef](#)]
6. Restrepo, C.E. Arc fault detection and discrimination methods. In Proceedings of the Electrical Contacts-2007 53rd IEEE Holm Conference on Electrical Contacts, Pittsburgh, PA, USA, 16–19 September 2007; IEEE: New Orleans, LA, USA, 2007; pp. 115–122.
7. Lippert, K.J.; Domitrovich, T.A. AFCIs—From a standards perspective. *IEEE Trans. Ind. Appl.* **2014**, *50*, 1478–1482. [[CrossRef](#)]
8. Gregory, G.; Wong, K.; Dvorak, R. More about arc-fault circuit interrupters. *IEEE Trans. Ind. Appl.* **2004**, *40*, 1006–1011. [[CrossRef](#)]
9. Jiang, B.-F.; Wang, L. Research on Arc Model in AC Low-voltage Electrical Wire Fault. *Proc. Chin. Soc. Univ. Electr. Pow. Syst. Autom.* **2009**, *4*, 50.
10. Wang, Q.L.; Wang, B.W.; Guan, L.H.; Zhao, Z.Z. Model and experiment of low-voltage AC series fault arc. *Chin. Soc. Univ. Electr. Pow. Syst. Autom.* **2018**, *30*, 26–35.
11. Lan, H.-L.; Zhang, R.-C. Study on the feature extraction of fault arc sound signal based on wavelet analysis. *Proc. Chin. Soc. Univ. Electr. Pow. Syst. Autom.* **2008**, *4*, 57–62.
12. Kim, C.J. Electromagnetic Radiation Behavior of Low-Voltage Arcing Fault. *IEEE Trans. Deliv.* **2009**, *24*, 416–423. [[CrossRef](#)]
13. Zhang, R.C.; Yang, J.; Du, J.H. Study on In-Process Detection and Diagnosis of Faults Arc Based on Early Sounds Signature and Intermittent Chaos. *Key Eng. Mater.* **2008**, *381*, 611–614. [[CrossRef](#)]
14. Meyar-Naimi, H.; Hasanzadeh, S.; Sanaye-Pasand, M. Discrimination of arcing faults on overhead transmission lines for single-pole auto-reclosure. *Int. Trans. Electr. Energy* **2013**, *23*, 1523–1535. [[CrossRef](#)]
15. Zhang, L.P.; Miao, X.R.; Shi, D.Y. Control, Research on low voltage arc fault recognition method based on EMD and ELM. *Electric. Mach. Control* **2016**, *20*, 54–60.
16. Jia, K.; Ren, Z.; Bi, T.; Yang, Q. Ground fault location using the low voltage side recorded data in distribution systems. *IEEE Trans. Ind. Appl.* **2015**, *51*, 1. [[CrossRef](#)]
17. Ukil, A. Detection of direction change in prefault current in current-only directional overcurrent protection. In Proceedings of the Conference of the IEEE Industrial Electronics Society, Florence, Italy, 23–26 October 2016; pp. 3829–3833.
18. Fan, R.; Liu, Y.; Huang, R.; Diao, R.; Wang, S. Precise Fault Location on Transmission Lines Using Ensemble Kalman Filter. *IEEE Trans. Deliv.* **2018**, *33*, 3252–3255. [[CrossRef](#)]
19. Lu, Q.; Wu, H.; Wang, S.; Kong, D. *Research on Arc-Fault Detection Based on Difference-Root Mean Square Method*; Low Voltage Apparatus: Shanghai, China, 2013; pp. 6–10.
20. Cai, X. Research on Arc Fault Detection and Protection Technique in the AC Supply System. Master’s Thesis, Nanjing University of Aeronautics and Astronautics, Nanjing, China, 2013.
21. Zhang, S.; Zhang, F.; Liu, P.; Han, Z. Series Arc Fault Identification Method Based on Energy Produced by Wavelet Transformation and Neural Network. *Trans. China Electrotech. Soc.* **2014**, *29*, 290–295.
22. Kawady, T.A.; Elkalashy, N.I.; Ibrahim, A.E.; Taalab, A.-M.I. Arcing fault identification using combined Gabor Transform-neural network for transmission lines. *Int. J. Electr. Syst.* **2014**, *61*, 248–258. [[CrossRef](#)]
23. Liu, Y.-W.; Wu, C.-J.; Wang, Y.-C. Detection of serial arc fault on low-voltage indoor power lines by using radial basis function neural network. *Int. J. Electr. Syst.* **2016**, *83*, 149–157. [[CrossRef](#)]
24. Lu, Q.W.; Wang, T.; Li, Z.R.; Wang, C. Detection Method of Series Arcing Fault Based on Wavelet Transform and Singular Value Decomposition. *Trans. China Electrotech. Soc.* **2017**, *17*, 212–221.
25. Guo, F.Y.; Deng, Y.; Wang, Z.Y.; You, J.L.; Gao, H.X. Series Arc Fault Characteristics Based on Gray Level-Gradient Co-Occurrence Matrix. *Trans. China Electrotech. Soc.* **2018**, *33*, 71–81.
26. Wang, Y.; Zhang, F.; Zhang, S. A New Methodology for Identifying Arc Fault by Sparse Representation and Neural Network. *IEEE Trans. Instrum. Meas.* **2018**, *67*, 2526–2537. [[CrossRef](#)]
27. Ma, S.H.; Bao, J.Q.; Cai, Z.Y.; Meng, C.Y. A novel arc fault identification method based on information dimension and current zero. *Proc. CSEE* **2016**, *36*, 2572. (In Chinese)

28. Yang, K.; Zhang, R.C.; Yang, J.H.; Du, J.H.; Chen, S.H.; Tu, R. Series arc fault diagnostic method based on fractal dimension and support vector machine. *Trans. China Electrotech. Soc.* **2016**, *31*, 70–77.
29. Cheng, Q.; Tezcan, J.; Cheng, J. Confidence and prediction intervals for semiparametric mixed-effect least squares support vector machine. *Recognit. Lett.* **2014**, *40*, 88–95. [\[CrossRef\]](#)
30. Gu, H.; Zhang, F.; Wang, Z.; Ning, Q.; Zhang, S. In Identification method for low-voltage arc fault based on the loose combination of wavelet transformation and neural network. In Proceedings of the 2012 Power Engineering and Automation Conference, Wuhan, China, 14–16 September 2012; pp. 1–4.
31. Zhang, S.; Zhang, F.; Liu, P.; Han, Z. Identification of Low Voltage AC Series Arc Faults by using Kalman Filtering Algorithm. *Elektronika ir Elektrotechnika* **2014**, *20*, 51–57. [\[CrossRef\]](#)
32. Surrel, Y. Design of algorithms for phase measurements by the use of phase stepping. *Appl. Opt.* **1996**, *35*, 51. [\[CrossRef\]](#)
33. Kuai, M.; Cheng, G.; Pang, Y.; Li, Y. Research of Planetary Gear Fault Diagnosis Based on Permutation Entropy of CEEMDAN and ANFIS. *Sensors* **2018**, *18*, 782. [\[CrossRef\]](#)
34. Zheng, J.; Pan, H.; Yang, S.; Cheng, J. Generalized composite multiscale permutation entropy and Laplacian score based rolling bearing fault diagnosis. *Mech. Syst. Process.* **2018**, *99*, 229–243. [\[CrossRef\]](#)
35. Kreuzer, M.; Kochs, E.F.; Schneider, G.; Jordan, D. Non-stationarity of EEG during wakefulness and anaesthesia: Advantages of EEG permutation entropy monitoring. *J. Clin. Monit. Comput.* **2014**, *28*, 573–580. [\[CrossRef\]](#)
36. Ravelo-García, A.G.; Navarro-Mesa, J.L.; Casanova-Blancas, U.; Martin-Gonzalez, S.; Quintana-Morales, P.; Guerra-Moreno, I.; Canino-Rodríguez, J.M.; Hernández-Pérez, E. Application of the Permutation Entropy over the Heart Rate Variability for the Improvement of Electrocardiogram-based Sleep Breathing Pause Detection. *Entropy* **2015**, *17*, 914–927. [\[CrossRef\]](#)
37. Costa, M.; Goldberger, A.L.; Peng, C.K. Multiscale Entropy Analysis of Complex Physiologic Time Series. *Phys. Rev. Lett.* **2007**, *89*, 705–708. [\[CrossRef\]](#)
38. Azami, H.; Escudero, J. Improved multiscale permutation entropy for biomedical signal analysis: Interpretation and application to electroencephalogram recordings. *Biomed. Process. Control* **2016**, *23*, 28–41. [\[CrossRef\]](#)
39. Shuting, W.; Xiong, Z. Teager Energy-entropy Ratio of Wavelet Packet Transform and Its Application in Bearing Fault Diagnosis. *Entropy* **2018**, *20*, 388.
40. Lin, S.; Liu, Z.; Hu, K. Traction Inverter Open Switch Fault Diagnosis Based on Choi–Williams Distribution Spectral Kurtosis and Wavelet-Packet Energy Shannon Entropy. *Entropy* **2017**, *19*, 504. [\[CrossRef\]](#)
41. Cutler, A.; Cutler, D.R.; Stevens, J.R. Random Forests. *Mach. Learn.* **2004**, *45*, 157–176.
42. Wang, Z.; Zhang, Q.; Xiong, J.; Xiao, M.; Sun, G.; He, J. Fault Diagnosis of a Rolling Bearing Using Wavelet Packet Denoising and Random Forests. *IEEE Sens. J.* **2017**, *17*, 5581–5588. [\[CrossRef\]](#)
43. Cai, J.-D.; Yan, R.-W. Fault diagnosis of power electronic circuit based on random forests algorithm. In Proceedings of the 2009 Fifth International Conference on Natural Computation, Tianjin, China, 14–16 August 2009; pp. 214–217.
44. Ko, B.C.; Gim, J.W.; Nam, J.Y. Cell image classification based on ensemble features and random forest. *Electron. Lett.* **2011**, *47*, 638. [\[CrossRef\]](#)
45. Özçift, A. Random forests ensemble classifier trained with data resampling strategy to improve cardiac arrhythmia diagnosis. *Comput. Biol. Med.* **2011**, *41*, 265–271. [\[CrossRef\]](#)
46. Liu, R.; Tan, T. An SVD-based watermarking scheme for protecting rightful ownership. *IEEE Trans. Multimed.* **2002**, *4*, 121–128.
47. Do, V.T.; Chong, U.P. Signal Model-Based Fault Detection and Diagnosis for Induction Motors Using Features of Vibration Signal in Two-Dimension Domain. *Stroj. Vestn-J Mech. E.* **2011**, *57*, 655–666. [\[CrossRef\]](#)
48. Wen, L.; Li, X.; Gao, L.; Zhang, Y. A New Convolutional Neural Network-Based Data-Driven Fault Diagnosis Method. *IEEE Trans. Ind. Electron.* **2018**, *65*, 5990–5998. [\[CrossRef\]](#)
49. Salas, J.; Kim, H.; Eykholt, R. Nonlinear dynamics, delay times, and embedding windows. *Phys. D Phenom.* **1999**, *127*, 48–60.
50. Pompe, B.; Bandt, C. Permutation Entropy: A Natural Complexity Measure for Time Series. *Phys. Rev. Lett.* **2002**, *88*, 174102.
51. Cao, Y.; Tung, W.-W.; Gao, J.B.; Protopopescu, V.A.; Hively, L.M. Detecting dynamical changes in time series using the permutation entropy. *Phys. Rev. E* **2004**, *70*, 046217. [\[CrossRef\]](#)

52. Yan, R.; Liu, Y.; Gao, R.X. Permutation entropy: A nonlinear statistical measure for status characterization of rotary machines. *Mech. Syst. Process.* **2012**, *29*, 474–484. [[CrossRef](#)]
53. Zhang, X.; Liang, Y.; Zhou, J.; Zang, Y. A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM. *Measurement* **2015**, *69*, 164–179. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).