Journal of Computers Vol. 33 No. 6, December 2022, pp. 21-36
doi: 10.53106/199115992022123306002

An Non-Intrusive Load Event Detection Approach Based on CEEMDAN-WTD-F Test

Ling-Zhi Yi[1], Xi-Meng Liu[1*], Guo-Yong Zhang[2], Hui-Na Song[3], Ning Liu[1]

[1] College of Information Engineering, Xiangtan University & Hunan Engineering Research Center of Multi-energy Cooperative Control Technology, Xiangtan, Hunan, China
ylzwyh@xtu.edu.cn, {542494402, 1355228562}@qq.com
[2] Key Laboratory of Disaster Prevention and Reduction for Power Grid Transmission and Distribution Equipment & State-Grid Hunan Electric Power Company Disaster Prevention and Reduction Center, Changsha, Hunan, China
zgyhust@163.com
[3] Wasion Group Limited, Changsha, Hunan, China
16299972@qq.com

Received 6 March 2022; Revised 11 June 2022; Accepted 6 July 2022

Abstract. To improve the perception of the switch state of the electrical equipment and realize the identification of the non-intrusive load switching process more accurately, a non-intrusive load event detection method based on the CEEMDAN-WTD-F test is proposed in this paper. Firstly, the adaptive median filter is used to reduce the noise fluctuation of power data of electric equipment and the discrete sequence derivative is used to reduce turn-on transition time for individual loads. Then, based on the principle of decomposition denoising and statistical testing, Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) and Wavelet Threshold De-noising (WTD) are used to process the input signal, and then the F test is used to determine whether an event occurs. Then, the proposed detection method is verified on the BLUED public data set, and a comparison experiment is carried out with the existing detection algorithm. The experimental results fully show that the proposed method has a lower missed detection rate and false detection rate than other detection algorithms, and the detection precision rate, recall rate, and F1 value are higher than 97%.

Keywords: non-intrusive load monitoring, event detection, Complete Ensemble Empirical Mode Decomposition with Adaptive Noise, Wavelet Threshold De-noising, F-test

1 Introduction

The early research on Non-Intrusive Load Monitoring (NILM) can be traced back to the 1980s and was first proposed by George Hart of MIT [1]. Compared with traditional monitoring methods, it is only necessary to install non-intrusive monitoring equipment at the electricity entrance of power users, and through the data of the entire building, the type, switching situation, and operation situation of each equipment can be identified [2]. These power consumption situations have high application value and can bring benefits to power companies and users [3]. For power companies, NILM can be used to realize the de-tailed perception of load types, improve the operational reliability, and make dynamic electricity price and power decisions more scientifically and rationally. For electricity users, the equipment details obtained through load monitoring can be fed back in time, which can detect the working status of electrical appliances to eliminate faults, and plan electricity consumption more reasonably. At the same time, it also protects the privacy of users [4-6].

NILM includes event detection, feature extraction, and load identification [7]. Event detection is an important step in non-intrusive monitoring technology. So far, sever-al methods have been used for event detection. The Cumulative Sum (CUSUM) detection algorithm proposed in the literature [8], determined whether a load switching event had occurred by calculating whether the change of power in the sliding window exceeded a set threshold. The Generalized Likelihood Ratio (GLR) algorithm proposed in the literature [9], and the Goodness Of Fit (GOF) algorithm proposed in the literature [10] were commonly used event detection algorithms, but the window sizes of the reference set and the current set must be fixed. Reference [11] improved the probability ratio theory based on CUSUM and proposed an event detection algorithm based on bilateral CUSUM, but for different situations, it was necessary to manually adjust the model parameters in advance. Reference [12] proposed an improved sliding window event detection algorithm, which calculated the standard deviation of the power sequence within the sliding window to detect load switching. In reference [13] proposed to detect of switching events by calculating the variance and absolute average deviation in the sliding window. Reference [14] applied cepstrum analysis

* Corresponding Author

to event detection and added a voting mechanism based on the chi-square goodness-of-fit method.

However, there are problems such as large power fluctuations caused by the simultaneous operation of a variety of electrical equipment, and a long transition time when some equipment is turned on, which will lead to multiple repeated detections of the same event.

In view of the above problems, to avoid false detection or missed detection and improve the detection accuracy and recall rate. In this paper, a non-intrusive load event detection method based on the CEEMDAN-WTD-F test is proposed. First, the two main factors and processing methods that affect the detection of non-intrusive switching events are analyzed in detail. The adaptive median filtering is used to reduce the volatility of electrical equipment, and the discrete sequence derivatives are used to deal with electrical equipment with a long transition time which is to reduce the shortcomings of insufficient detection accuracy caused by these factors. Then, based on the principle of de-noising and statistical testing, CEEMDAN is used to decompose the load power data. Analyze and filter the correlation coefficient of each component, select the component whose correlation coefficient is greater than 0.5, and perform wavelet threshold processing. After that, these components are reconstructed to obtain a new signal, and the switching event is detected by the F test; then, the parameters are debugged through specific experiments, including the wavelet basis function and the median filter strength, and the optimal parameters are selected as the experimental parameters of event detection. Finally, the non-intrusive load event detection method proposed in this paper is experimentally verified and analyzed by BLUED [15], and the effectiveness of the proposed method is further verified by comparing it with different event detection algorithms. At the same time, to further verify the broad practicability of the proposed method, a comparative analysis is also carried out in different datasets REDD and UK-DALE. After testing, the method can ensure fast and concise operation, and the event detection accuracy, recall, and detection F1 value can reach more than 97%.

2 Related Work and Our Contributions

Load event detection is an important step in non-intrusive load monitoring techniques and is an important prerequisite for load classification and identification. Identifying at which moments load state transitions occur lays the foundation for subsequent load feature extraction and load identification. So far, a number of methods have been used for event detection.

However, in the existing load event detection methods, it is determined whether an event occurs by detecting the change of the current or power curve, but in real life, there are many kinds of electrical equipment working at the same time or when the high-power load When connected, due to its circuit structure, voltage oscillation, electrical interaction, and other reasons, there will be long-term and large-scale power fluctuations and some equipment will have a long transient transition time when it is turned on, which will lead to multiple repeated detection of the same event. Considering these situations, the event detection will lead to false detection and missed detection of the event, so that the user cannot effectively perceive the state of the electrical equipment, which seriously affects the user's comfort and safety.

Aiming at the problems above, an event detection method based on CEEMDAN-WTD-F test is proposed. Firstly, the adaptive median filter is used to reduce the volatility of the electrical equipment, and the discrete sequence derivative is used to deal with the electrical equipment with a long transition time, to reduce or avoid the shortcomings of insufficient detection accuracy caused by these factors; CEEMDAN decomposes the load power data. Analyze and filter the correlation coefficient of each component, select the main correlation component, and perform wavelet threshold processing. After that, these components are reconstructed to obtain a new signal, and the F test is used to detect the switching events; then, the parameters are debugged through specific experiments, including the wavelet basis function and the median filter strength, and the optimal parameters are selected as the event detection parameters. Experimental parameters and experimental analysis are performed on the BLUDE dataset.

The main contributions of this paper are as follows:

1. The adaptive median filtering is used to reduce the volatility of electrical equipment, and the discrete sequence derivatives are used to deal with electrical equipment with a long transition time which is to reduce the shortcomings of insufficient detection accuracy caused by these factors.

2. Applying CEEMDAN-WTD to the power signal processing of event detection can effectively deal with noise, and in the process of restoring the initial signal, it can reduce the loss of signal, and lay the foundation for the later accurate identification of different states. The overall volatility of the information after denoising processing is smaller than that before processing, and the fluctuation is on both sides of the zero point and is relatively uniform, which reduces the standard deviation after processing and is not easily affected by extreme values.

3 Data Pre-processing

3.1 Adaptive Median Filtering Handles Data Volatility

The Median Filter (MF) [16] is a method to filter out the noise of the impulse signal. MF is one of the commonly used methods in smoothing signal processing of salt and pepper noise. It has a good effect of filtering noise on sequence signals, and can also reduce edge blur and pre-serve the characteristics of borders, so it can ensure no edge blur in some scenes, suitable for the field of digital signal processing.

The basic principle of MF is to replace the value of a certain point in the digital sequence with the median value of all point values in the adjacent window so that the sur-rounding values are close to the actual value, and then the isolated noise is removed. For signal $P_n = (1,2,\cdots,N)$, a sliding window of y (y is an odd number) points is usually used, and the window is treated as a neighborhood $(P_x, P_{x+1}, \cdots, P_{x+y})$, the values in it are arranged in or-der from small to large, and then the median value of the sequence is used instead of its center position. When y is an even number, the average of the two numbers in the middle of the sequence is taken as the median of the sequence.

The Adaptive Median Filter (AMF) [17] adds the ability to adaptively adjust the filter window size based on MF, and dynamically change the window size of the median filter according to the preset conditions. At the same time, taking into account the effect of removing salt and pepper noise and protecting the boundary, it has a better effect on suppressing noise. Fig. 1 shows the waveform of the active power of the 37th group in phase A before and after adaptive median filtering.

By analyzing the waveforms before and after processing in Fig. 1, it can be seen that the power changes after processing have the expected results, the peak value of the power is greatly reduced, and the stable value of the power signal is all lower than 200W. The graph before and after processing changes greatly, which weakens the peak value of the power signal, the fluctuation amplitude of the signal when the signal is stable, and the influence of the noise signal on the event detection result, and avoids the wrong detection of the event.

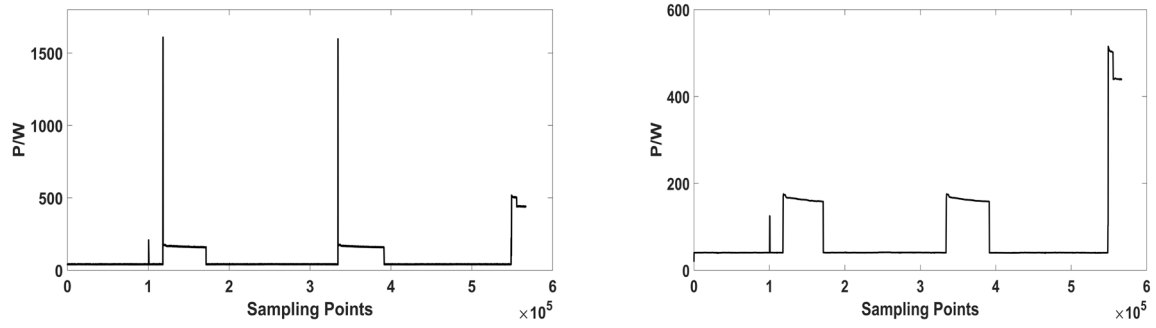


Fig. 1. Power of group 37 before and after AMF treatment

3.2 Discrete Sequence Derivative Processing Long Turn-on Transition Time

When the switching state or working model of the electrical load changes, its electrical behavior will also change, which will lead to deviations in the collected data. It can be seen from Table 1 that the transition time of the turn-on instant of most loads is not more than 0.2 seconds. Therefore, when the time parameter of the event detection algorithm is set to 0.2 seconds, the current working state and the switching events can be detected for most loads. However, some devices have longer state transition times, such as the on-time transition states of range hood, fridge, and AC, with on-transition times of 2.4 seconds, 0.6 seconds, and 0.4 seconds. At the same time, the turn-on process of these loads is a long and complex transient period, and the power of the equipment also fluctuates significantly during the transition period. For such a long-transition event load, when the ordinary event detection algorithm is used for detection, the change of continuous active power will continue to trigger detection, and it will always be recognized as an event occurs until the power reaches a stable state.

Table 1. Transition time for common appliance loads to turn on

Load	Hood	Fridge	AC	TV	Kettle	Fan	Bulb	Microwave
Duration(s)	2.4	0.8	0.4	0.2	0.15	0.1	0.1	0.1

To eliminate these loads with excessively long turn-on transient transition time, the waveforms of these electrical equipment transients should be carefully analyzed. It is found that the first type of loads containing motors, such as refrigerators and vacuum, will have a peak in the active power waveform during the transient state of turn-on, and then drop to a steady-state. The second category of loads with heating elements, such as microwave and electric kettles, whose turn-on transient power rises directly to a steady-state without showing a significant peak.

By observing the waveform at the turn-off time, different from the turn-on time, the transient power wave-form at the turn-off time has only one variation mode in the experiment, it is simply reduced to zero, without any special active power characteristics, so this article only considers the processing of the opening process.

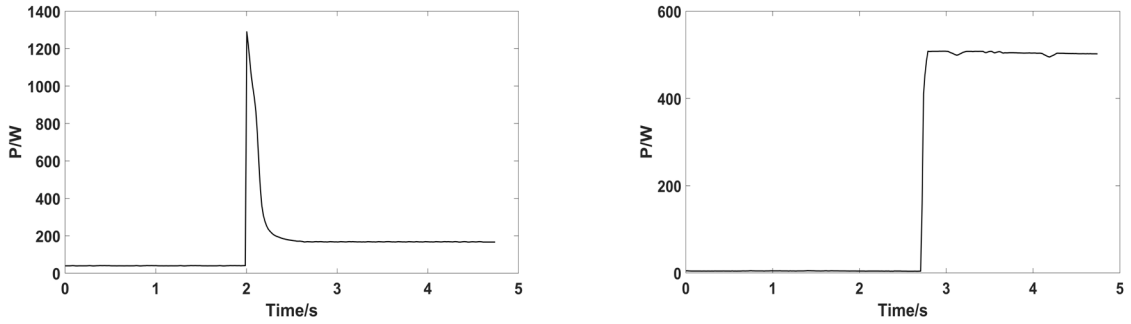
Therefore, based on the above characteristics of the electrical equipment waveform, this paper uses an additional discrete sequence derivative to eliminate the peak value at the turn-on time, prevent repeated detection, and ensure that the event is only detected once during the entire transient period. The discrete sequence derivative can transform the processing of digital signals into discrete signals through sampling. The value of a point is approximated by Taylor series expansion:

$$f(x_{t+1}) = f(x_t) + f'(x_t)h + \frac{f''(x_t)}{2}h^2 + \cdots \quad . \tag{1}$$

$$f'(x_t) = \frac{f(x_{t+1}) - f(x_t)}{h} + O(h) \quad . \tag{2}$$

$$f''(x_t) = \frac{f(x_{t+2}) - 2f(x_{t+1}) + f(x_t)}{h^2} + O(h) \quad . \tag{3}$$

Where $f(x_t)$ denotes the t-th selected electrical parameter, $f'(x_t)$ and $f''(x_t)$ denotes the first and second order derivatives of electrical parameters, h denotes equal spacing between two points. And the derivative describes the direction of change of electrical parameters as well as the rate and slope of the total electrical parameter load data waveform (Fig. 2).



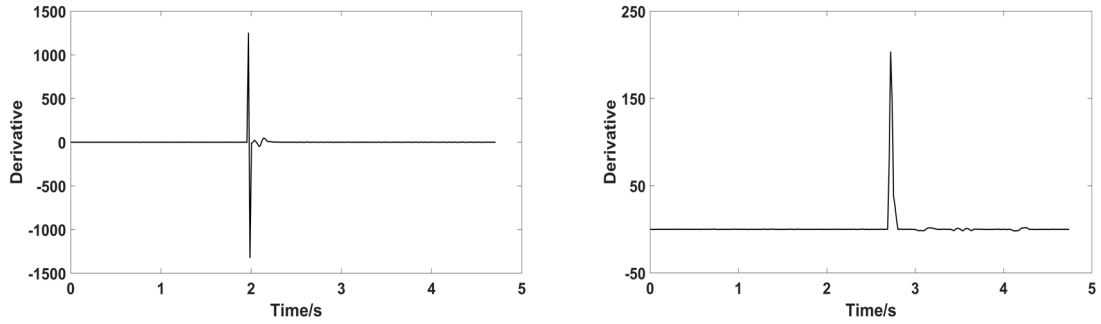


Fig. 2. AC and Microwave turn-on transient power waveform and derivative curve

4 CEEMDAN-WTD-F Event Detection Method

4.1 Principle of CEEMDAN-WTD

The CEEMDAN [18] is to decompose the original signal into IMF arranged in the order of frequency from high to low, then add the corresponding Gaussian white noise with an adaptive performance in each stage of the IMF component. The IMF obtained after the first stage of processing is averaged to obtain a unique margin, and the processing is repeated in turn until the signal decomposition is completed. It can effectively handle the transfer of noise from high frequency to low frequency, and in the process of restoring the initial signal, it can stabilize the amplitude of the high-frequency signal to avoid modal aliasing, thereby reducing the loss of high-frequency signal.

The WTD produces a series of wavelet coefficients by passing the signal through a wavelet transform, then a suitable threshold is selected, the wavelet coefficients are processed, and coefficients smaller than the threshold are discarded. Finally, the processed wavelet coefficients are wavelet inverted to obtain the reconstructed denoised signal.

The specific steps of CEEMDAN-WTD are as follows: assume that $x(t)$ is the original signal; $v(t)$ is the Gaussian white noise whose magnitude satisfies the standard normal distribution; ε is the amplitude factor of the Gaussian white noise, and the n-th amplitude factor is denoted as ε_n. In finding the IMF component of each order, add the white noise with adaptive amplitude to adjust $\varepsilon_n = \beta \cdot std(r_n(t))$, β is taken to be 0.2, $std()$ denotes the standard deviation at the time of calculation; $r_n(t)$ is the n-th residual component; $E_n()$ is the n-th IMF component obtained by EMD decomposition of the original signal; $IMF_n(t)$ is the n-th modal component produced by CEEMDAN decomposition of the signal.

1. The new signal is obtained by adding Gaussian white noise to the original signal. The new signal is averaged over N experiments so that the first IMF component obtained by decomposition using the EMD method can be expressed as:

$$IMF_1\ t\ = \frac{1}{N}\sum_{i=1}^{N} E_1(x(t)+\varepsilon_0 \cdot v_i(t))\ . \tag{4}$$

2. The first residual component $r_1(t)$ is obtained by removing $IMF_1(t)$ from the original data $x(t)$:

$$r_1\ t\ = x\ t\ - IMF_1\ t\ . \tag{5}$$

3. Treat $r_1(t)$ as the new original data and also add white noise $v_i(t)$, then calculate the components of the second order IMF_2. And also find the second remaining component as:

$$IMF_2\ t\ =\frac{1}{N}\sum_{i=1}^{N}E_1\ r_1\ t\ +\varepsilon_1 E_1\ v_i\ t\ . \tag{6}$$

$$IMF_{n+1}\ t\ =\frac{1}{N}\sum_{i=1}^{N}E_n\ r_n\ t\ +\varepsilon_n E_n\ v_i\ t\ . \tag{7}$$

4. Repeat the above steps to obtain the n+1-th $IMF_n(t)$ and n-th residual components $r_n(t)$:

$$IMF_{n+1}\ t\ =\frac{1}{N}\sum_{i=1}^{N}E_n\ r_n\ t\ +\varepsilon_n E_n\ v_i\ t\ . \tag{8}$$

$$r_n\ t\ =r_{n-1}\ t\ -IMF_n\ t\ . \tag{9}$$

If the residual component cannot be decomposed, the above process is interrupted and the signal is:

$$x\ t\ =\sum_{i=1}^{n}IMF_i+r_n\ t\ . \tag{10}$$

5. The correlation coefficient of IMF is screened to determine whether it is the main component. Among them, the correlation coefficient greater than or equal to the set value t=0.5 is defined as the main component, and the rest is defined as the non-main component.

6. The wavelet coefficients are obtained by decomposing the dominant components into wavelets with appropriate wavelet functions and thresholds. The wavelet coefficients are then quantized by selecting the appropriate threshold function for the decomposition. The expressions of the hard and soft thresholding functions are:

$$hard(x,l)=\begin{cases}x\ ,|x|>l\\0\ ,|x|\le l\end{cases}\ . \tag{11}$$

$$soft(x,l)=\begin{cases}\mathrm{sgn}(x)(x-l)\ ,|x|>l\\0\qquad\qquad\ ,|x|\le l\end{cases}\ . \tag{12}$$

7. The processed wavelet coefficients are analyzed, and the wavelet coefficients larger than the wavelet threshold are selected for reconstruction. The components are then reconstructed to obtain a new signal.

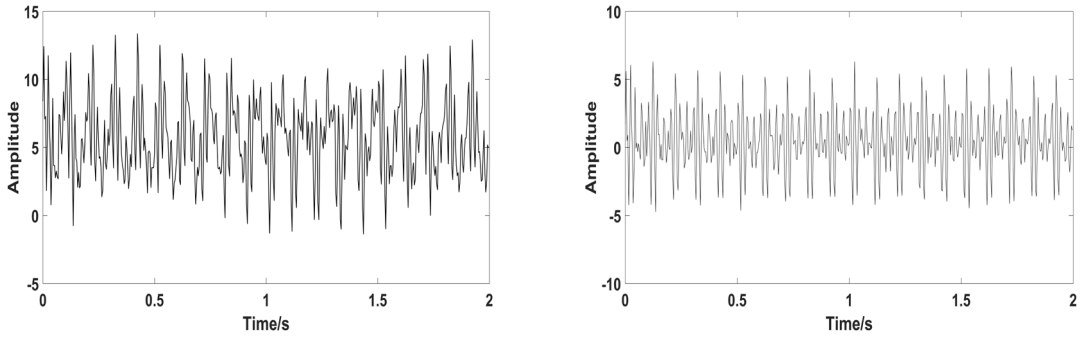


Fig. 3. Original waveform and decomposed denoised waveform

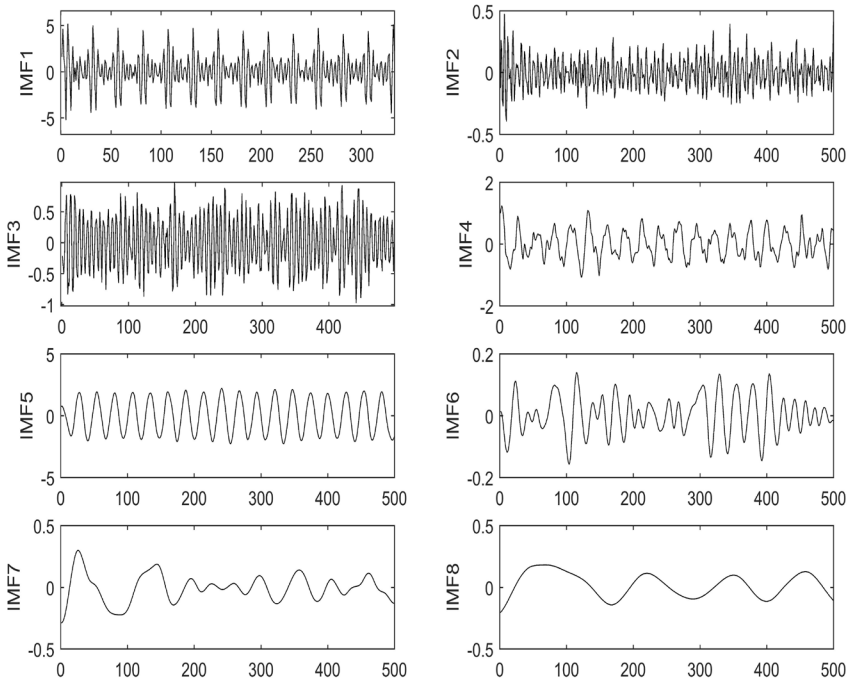


Fig. 4. IMF component waveform of CEEMDAN decomposition

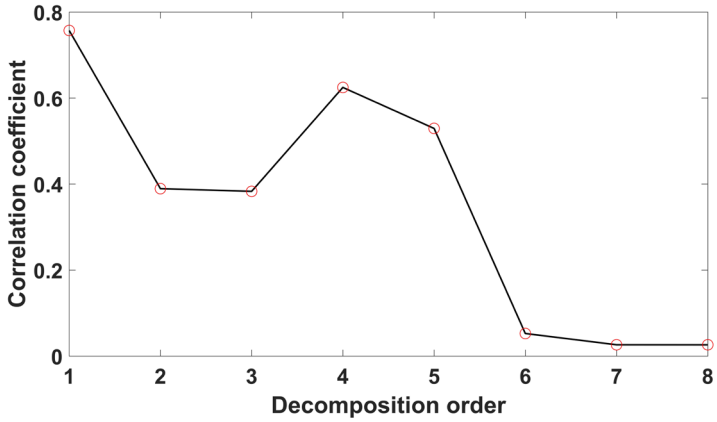


Fig. 5. Correlation coefficient

Fig. 3 shows the original power waveform and the waveform after the CEEMDAN-WTD decomposition and denoising process. This paper randomly intercepts the active power waveform from the BLUED data set with 4000 sampling points. By comparing Fig. 3, the waveforms before and after CEEMDAN-WTD decomposition and denoising processing can be analyzed. It can be concluded that the overall fluctuation of the information after denoising processing is smaller than that before processing. It is relatively uniform so that the standard deviation after processing is reduced, and it is not easily affected by extreme values. Compared with the original sequence, the sequence used after processing has a smaller standard deviation, which weakens the influence of extreme values and further reduces the probability of event misjudgment.

Fig. 4 shows the 8 IMF components obtained by the adaptive decomposition of the original signal after CEEMDAN. By calculating the autocorrelation function and correlation coefficient of each order of the 8 IMF components obtained by decomposition, the main components and non-main components are accurately located.

Fig. 5 shows the IMF autocorrelation coefficients of each order.

4.2 F-test

F-test is a hypothesis test method commonly used in statistics to test the variance change of two data series. By comparing the homogeneity of variance of two groups of data, it is determined whether there is a significant difference between the two groups of data. According to the results of the F test, non-intrusive stress event detection studies can be carried out.

The basic principle of the F-test is as follows: set the two groups of samples are $X_1 = (x_1, x_2, x_3 \cdots x_n)$ $Y_1 = (y_1, y_2, y_3 \cdots y_m)$. So the mean and variance of the two groups of samples are:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i, \overline{Y} = \frac{1}{m}\sum_{j=1}^{m} y_j \ . \tag{13}$$

$$S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n} x_i - \overline{X}^{\ 2}, S_Y^2 = \frac{1}{m-1}\sum_{j=1}^{m} y_j - \overline{Y}^{\ 2} \ . \tag{14}$$

F is defined as the ratio of the variances of the two data samples, the expression for F is:

$$F = \frac{S_X^2}{S_Y^2} : F\ n-1, m-1 \ . \tag{15}$$

Where *n*-1 denotes the degree of freedom of the numerator and *m*-1 denotes the degree of freedom of the denominator. And n and m respectively represent the number of sampling points of the two groups' data.

4.3 Steps Based on the CEEMDAN-WTD-F Test

The event detection flow chart based on the CEEMDAN-WTD-F test is shown in Fig. 6. The specific steps are:

Step 1. Input the raw data.

Step 2. Data preprocessing. Divide the events into 65 groups in chronological order, and perform adaptive median filtering processing on the power to filter out the accuracy of event detection due to data fluctuations.

Step 3. Discrete Sequence Derivative Operations. The discrete sequence derivation is used to process the smoothness of the filtered data to reduce the instability caused by the long transition time of equipment startup.

Step 4. The processed signal is decomposed and denoised by CEEMDAN-WTD and the reconstructed information is extracted.

Step 5. Calculate the primary variance. Using the statistical knowledge, calculate the variance S_X^2 of the reconstructed signal.

Step 6. The maximum deviation is eliminated. Check the recombination sequence, find out the maximum residual, that is, the maximum and minimum deviations, and remove them, and the remaining data form a new sequence;

Step 7. Calculate the quadratic variance. Do the variance calculation again and note as S_Y^2 .

Step 8. F-test. According to the variance $S_X^{\ 2}$ and S_Y^2 found in step 6 and step 8 respectively, calculate the F. And compared with F_α at a given significance level α .

Step 9. Determine whether there is a load event. If $F > F_\alpha$, the load data corresponding to the maximum deviation of the rejection is a load event; otherwise, it is not a load event.

Step 10. Record the number of detected events, and judge whether it is a correctly detected load event. If yes, add 1 to TP, otherwise, add 1 to FP, and calculate the PRE, REC, and F1 values at the same time.

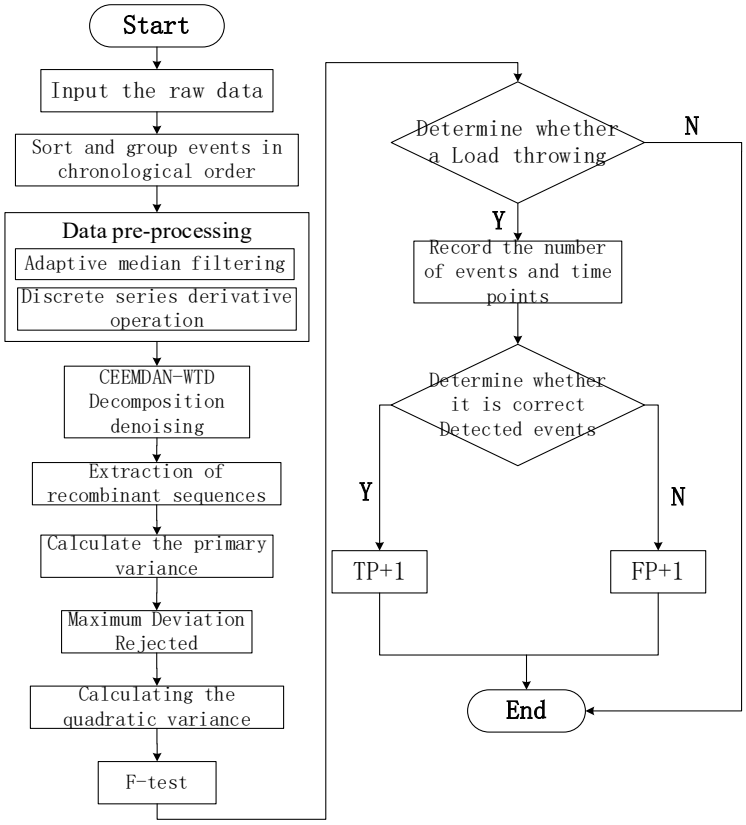


Fig. 6. Flow chart based on CEEMDAN-WTD-F test

5 Experimental Verification and Results Analysis

5.1 Measurement Indicators

Since there are only two results of event detection at each moment: occurrence or non-occurrence of load switching events, non-intrusive event detection is essentially a binary classification problem, and the output results have only two types of models, positive and negative. Therefore, a confusion matrix commonly used in performance evaluation in the two-class problem is introduced, and the detection results are divided into four categories.

Among them, TP indicates that an event has occurred and the detection result is an event; FP indicates that no event has occurred but the detection result is an event; TN indicates that no event has occurred and the detection result is no event, and FN indicates that an event has occurred but the detection result is no event. During the limited number of load switching operations and the testing process, the value of TN must be much larger than the other three types. To avoid the influence caused by the imbalance of sample categories, the selected metrics should not depend on the value of the TN category. Therefore, this paper uses the precision rate (PRE), the recall rate (REC), and the F1 value to evaluate the detection results.

$$PRE = \frac{TP}{TP + FP} . \tag{16}$$

$$REC = \frac{TP}{TP + FN} . \tag{17}$$

$$F1 = 2 \times \frac{PRE \times REC}{PRE + REC} . \tag{18}$$

The PRE represents the ratio of the actual events in all the detected load switching events, which measures the detection accuracy of the event detection method. The REC represents the proportion of load events detected in all load switching events, and it measures whether the event detection method can comprehensively detect the entire data set. The F1 value is the harmonic mean of precision and recall, which comprehensively measures the score of the event detection algorithm on the two indicators of precision and recall.

5.2 Experimental Parameter Tuning

Based on the previous experiments and the proposed method, the influence of wavelet base function, wavelet threshold function, and filter strength S of AMF on the event detection results is analyzed. To determine the best parameter values for the experimental results, assay analysis was performed on the BLUED dataset.

This paper uses the public data set BLUED to train and evaluate the proposed method, which records the electricity usage of a real home for a week, using sensors with a sampling frequency of 12khz. This dataset collects switching events of multiple loads of various types, only 871 of them are used in the experiment of this paper, and all events are divided into 65 groups according to the sampling time sequence.

Table 2. Detection results of different wavelet basis functions

Function	Detected	TP	FP	FN	PRE	REC
Sym	936	861	75	10	0.919	0.989
Coif	941	851	90	20	0.904	0.977
Bior	932	840	92	31	0.901	0.964
Db	955	853	98	18	0.897	0.979

Firstly, the effect of different wavelet basis functions on event detection is analyzed, and the filtering strength S is fixed. Then the experiments are carried out by using Sym wavelet, Coif wavelet, Bior wavelet, and Db wavelet respectively. The experimental results are shown in Table 2.

According to the experimental results of the four wavelet basis functions in Table 3, the Sym wavelet basis function has the highest identification PRE and REC. At the same time, considering the characteristics of input power and the purpose of denoising, Sym wavelet and the general threshold method (sqtwolog) are selected for specific analysis.

Table 3. Sym series wavelet detection results for threshold functions

Function		Detected	TP	FP	FN	PRE	REC	F1
	S	936	861	75	10	0.919	0.989	0.953
Sym1								
	H	925	860	65	11	0.929	0.987	0.957
	S	911	858	53	13	0.942	0.985	0.963
Sym2								
	H	897	857	40	14	0.955	0.984	0.969
	S	893	856	37	15	0.959	0.983	0.971
Sym3								
	H	886	855	31	16	0.965	0.982	0.973
	S	884	854	30	17	0.966	0.980	0.972
Sym4								
	H	881	853	28	18	0.968	0.979	0.973
	S	878	852	26	19	0.970	0.978	0.974
Sym5								
	H	876	851	25	20	0.971	0.977	0.974
	S	874	850	24	21	0.973	0.976	0.974
Sym6								
	H	872	849	23	22	0.974	0.975	0.974

The above table records the event detection results of the wavelet base functions of the Sym series under the corresponding soft and hard threshold functions. Through comparative analysis, it can be seen that the PRE and REC are more balanced when the wavelet function is Sym3 and the threshold function is the hard threshold. And suitable for event detection.

Table 4. Records of test results with different filter strengths

S	Detected	TP	FP	FN	PRE	REC	F1
7	934	862	75	9	0.919	0.989	0.954
19	923	861	65	10	0.929	0.988	0.958
27	917	860	57	11	0.938	0.987	0.962
35	904	859	45	12	0.950	0.987	0.968
43	898	858	40	13	0.955	0.985	0.970
61	882	856	26	15	0.971	0.983	0.977
75	877	855	22	16	0.975	0.982	0.978
95	870	851	19	20	0.978	0.977	0.978
109	858	847	11	24	0.987	0.972	0.979
113	855	846	9	25	0.989	0.971	0.980

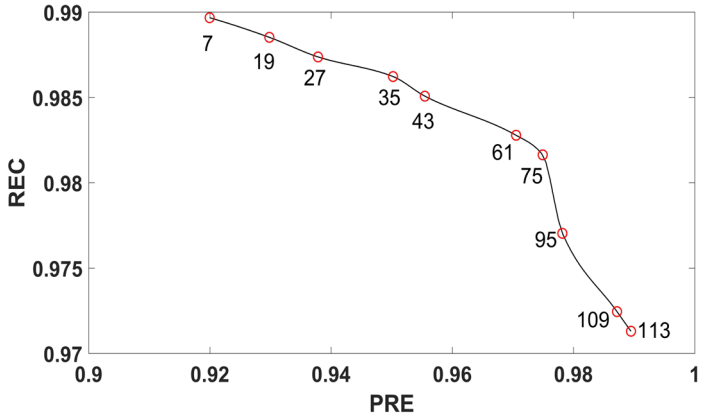


Fig. 7. P-R curves with different filter strengths

When the wavelet function is Sym3 and the threshold function is the hard threshold, the influence of the filter strength *S* of *AMF* on the event detection results is analyzed. Table 4 records the event detection results under different filter strengths S and draws the P-R curve as shown in Fig. 7.

By analyzing the detection results under different filter strengths, it can be seen that when the filter strength *S* is about 75, the most ideal and balanced event detection effect can be obtained, and the precision and recall rates can reach close to 98%. Then analyze the detection effect of each group when $S = 75$.

5.3 Analysis of Test Results

Due to a large number of events in the dataset, it is impossible to show all the detection conditions. To fully demonstrate the performance effect of the detection method proposed in this paper, this paper randomly selects the detailed detection results of several groups of events and displays the detection results, as shown in Table 5 and Fig. 8.

From the actual event detection results of several random groups in Table 5, it can be seen that the PRE and F1 values for each group can exceed 90% and are close to 1, and the REC value is also high. For the analysis of 871 actual event detection results in the data set: the method proposed in this paper detected a total of 877 events, of

which 855 were correctly detected, 16 were missed, and 22 were falsely detected. The PRE is 97.5%, the REC is 98.2%, and the F1 value is 97.8%. The detection accuracy is very high, and there are enough correctly detected events to meet the high-precision detection requirements of events. The recall rate of detection is high, which proves that the number of missed detections of this method is small, and there is a high F1 value, which provides an efficient and novel method for the detection of non-intrusive load switch events.

Table 5. Specific test results of CEEMDAN-WTD-F

Group	Actual	Detected	TP	FP	FN	PRE	REC	F1
21	21	21	21	0	0	1	1	1
22	42	43	41	2	1	0.953	0.976	0.965
23	12	10	10	0	2	1	0.833	0.909
24	11	11	11	0	0	1	1	1
25	5	5	5	0	0	1	1	1
26	8	8	8	0	0	1	1	1
27	14	14	14	0	0	1	1	1
All	871	877	855	22	16	0.975	0.982	0.978

As can be seen in Fig. 8, the sampling point offset between the actual event and the detected event is small and can maintain the same sampling point, indicating that the detection error is small. The small false and missed detection results and the high detection accuracy indicators fully demonstrate the effectiveness of the proposed method.

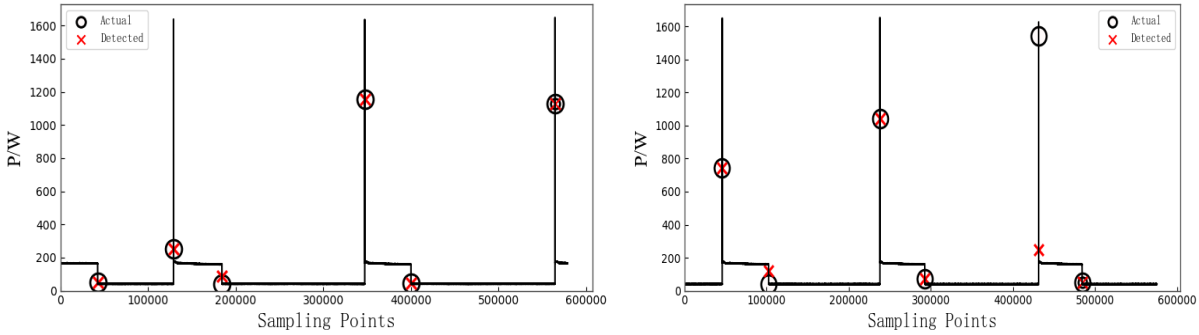


Fig. 8. Graph of test results for some groups

Fig. 9 shows the PRE, REC, and F1 values for each group when $S = 75$.

As can be seen from the below figure, among the total 65 sets of event data, the PRE of 50 sets of detected events has reached 100%, the PRE of 8 sets of events in the remaining 15 sets is greater than 90%, and the PRE of the smallest set of events is also above 65%. For REC, 55 groups of REC detected events reached 100%, and the accuracy rates of the remaining 10 groups of events were all above 80%, and most of them were more than 90% accurate. At the same time, the F1 values of these 65 groups are high, and most of them are close to or equal to 100%. These three evaluation indicators show that the accuracy of the event detection method proposed in this paper is high enough, and the number of false detections and missed detections is small.

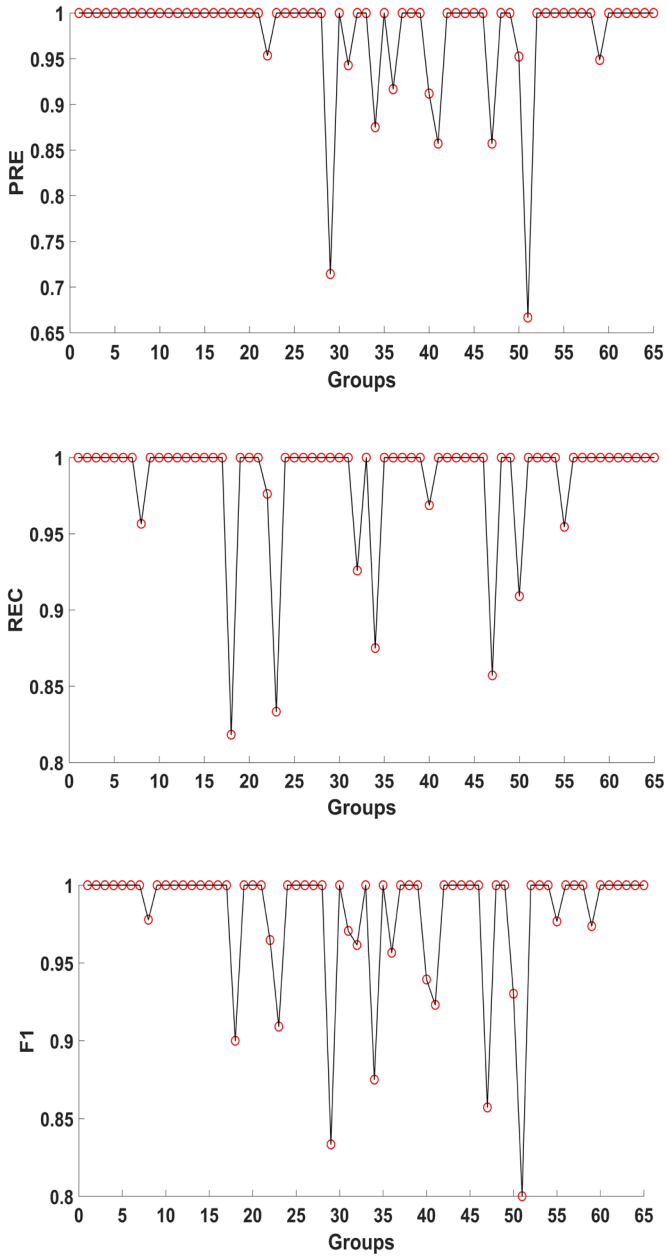


Fig. 9. PRE, REC, F1 results per group ($S = 75$)

5.4 Algorithm Comparison Analysis

Firstly, the detection method proposed in this paper is compared with several traditional event detection algorithms such as F-test, CUSUM, and GLR. Then compare the performance with several new algorithms: voting $\chi 2$ GOF in literature [14], and MAD-SW in literature [13]. All comparison experiments were carried out in the BLUED data set. The comparison results of PRE, REC, and F1 are shown in Table 6 as the experimental analysis results. Fig. 10 shows some of the detection results of various algorithms.

Comparing with the three traditional event detection methods, it can be found that there are too many false detections when using F-testing events; when using CUSUM detection, when the power threshold is 10, the REC is high, but the number of FPs is also high, there is a high false detection. At the same time, when the CUSUM method is used, the number of missed detections and false detections in event detection is relatively large; when the GLR method is used, the REC is very high, but the PRE is only 0.922, indicating that the detection accuracy

is not high and there are many false detections.

Compared with several advanced event detection methods: voting χ2 GOF, and MAD-SW, it can be found that this event detection method is similar to the MAD-SW method in recall rate REC, but has better performance in precision rate and F1 (increase 3% and 1.2%, respectively). Compared with the voting χ2 GOF method, it can achieve better performance in various evaluation indicators (REC, PRE, and F1 are increased by 0.4%, 3.2%, and 1.8%, respectively). On REC, our method is slightly lower than MAD-SW, but both PRE and F1 are significantly higher than MAD-SW.

Table 6. Detection results of different methods

Algorithm		TP	FP	FN	PRE	REC	F1
F-test		842	266	29	0.759	0.967	0.851
CUSUM	P=10	856	194	15	0.815	0.983	0.891
	P=15	852	165	19	0.838	0.978	0.902
	P=20	846	142	25	0.856	0.971	0.910
	P=25	830	127	41	0.867	0.953	0.908
	P=30	806	117	65	0.873	0.925	0.899
GLR		860	73	11	0.922	0.987	0.953
χ2 GOF		852	52	19	0.942	0.978	0.960
MAD-SW		865	50	9	0.945	0.989	0.966
This paper		855	22	16	0.975	0.982	0.978

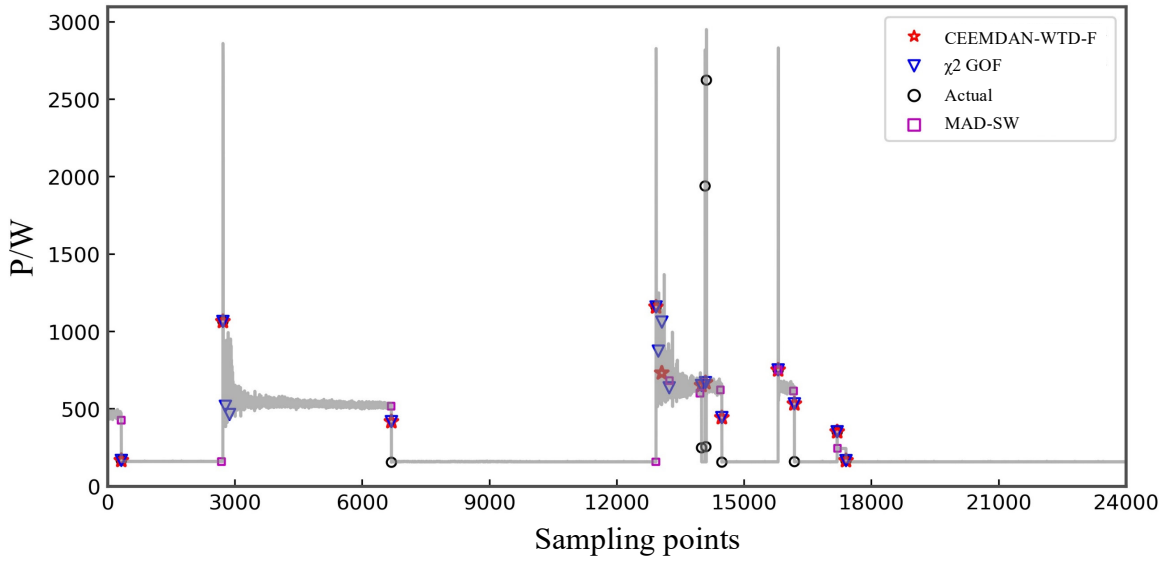


Fig. 10. Experimental comparison results of different detection methods

Throughout the comparison of several methods, only the CEEMDAN-WTD-F test method proposed in this paper has all the indicators close to the accurate value, the accuracy rate PRE is very high, reaching 97.5%, and the number of false detections and missed detections are very low, the number of events detected at the same time is closest to the actual number of events. To sum up, the test method based on CEEMDAN-WTD-F proposed in this paper can significantly reduce the false detection rate and improve the detection accuracy. Problems with the false positive rate or high missed call rate.

The above experiments are all performed on the BLUED dataset. To further test the validity of the CEEMDAN-WTD-F test, comparative experiments are conducted in BLUED, REDD, and UK-DALE, respectively. Since each data set contains a large number of load switching events, for the convenience of comparison, this paper only selects 100 actual events in each data set.

Table 7. Experimental results on different data set

Data set	Detecte	TP	FP	FN	PRE	REC
BLUED	101	97	4	3	0.961	0.970
REDD	99	96	3	4	0.969	0.962
UK-DALE	102	97	5	3	0.951	0.971

It can be seen from Table 7 that the event detection method proposed in this paper has good event detection accuracy for each data set. This can further illustrate the effectiveness of the proposed method.

6 Conclusion

Event detection is a basic link in unsupervised non-intrusive load monitoring technology, and its accuracy will affect the effect of the entire load decomposition and identification. To improve the accuracy of non-intrusive load monitoring, this paper proposes a load event detection method based on the CEEMDAN-WTD-F test, which is summarized as follows:

1. With CEEMDAN, the decomposition process can effectively overcome the mode-mixing problem, and the reconstruction error is almost zero, with greatly reduced computational cost. It can effectively handle the transfer of noise from high frequency to low frequency, and in the process of restoring the initial signal, it can stabilize the amplitude of the high-frequency signal to avoid modal aliasing, thereby reducing the loss of high-frequency signal.

2. Apply the wavelet threshold denoising method to the power signal processing of event detection can avoid the interference of noise to the power signal to a certain extent, and lay a foundation for the later accurate identification of different states.

3. The load event detection method proposed in this paper has simple parameters, is easy to adjust, and has a strong adaptive ability for different types of load events.

4. In terms of detection effect, compared with multiple event detection algorithms, the method proposed in this paper has the best PRE and F1 values and also has a higher REC. Therefore, this method has a better detection effect in event detection.

However, when the interval for changing the state of various loads is smaller than the sampling frequency since the working power of different loads is too different, the low-power equipment may be missed during event detection. At the same time, in order to realize the engineering application of non-intrusive load monitoring and decomposition, and how to further optimize the detection accuracy of this method and reduce its computational complexity, it will continue to be studied and improved in the follow-up work.

7 Acknowledgement

This work was supported by the National Natural Science Foundation of China (61572416), Hunan Province Natural Science Fund (2020JJ6009) and Hunan postgraduate innovation project and State key Laboratory of disaster Prevention and Mitigation of Power Transmission and Transformation equipment of Power Grid.

References

[1] G.W. Hart, Nonintrusive appliance load monitoring, Proceedings of the IEEE 80(12)(1992) 1870-1891.
[2] H.Y. Liu, S.B. Shi, X.H. Xu, A non-intrusive load identification method based on RNN model, Power System Protection and Control 47(13)(2019) 162-170.
[3] S.S. Hosseini, K. Agbossou, S. Kelouwani, Non-intrusive load monitoring through home energy management systems, A comprehensive review, Renewable and Sustainable Energy Reviews 79(2017) 1266-1274.
[4] P.M. Papadopoulos, V. Reppa, M.M. Polycarpou, Scalable distributed sensor fault diagnosis for smart buildings, Journal of Automatica Sinica 7(3)(2020) 638-655.
[5] N. Batra, A. Singh, K. Whitehouse, If you measure it, can you improve it? exploring the value of energy disaggregation, in: Proc. Buildsys'15 proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built, 2015.
[6] S. Ghosh, A. Chatterjee, D. Chatterjee, Improved non-intrusive identification technique of electrical appliances for a smart residential system, IET Generation, Transmission & Distribution 13(5)(2019) 695-702.
[7] Z. Lan, B. Yin, T. Wang, A non-intrusive load identification method based on convolution neural network, in: Proc. 2017

IEEE Conference on Energy Internet and Energy System Integration (EI2), 2017.
[8] M. Lu, Z. Li, A hybrid event detection approach for non-intrusive load monitoring, IEEE Transactions on Smart Grid 11(1) (2019) 528-540.
[9] H. Rashid, P. Singh, V. Stankovic, Can non-intrusive load monitoring be used for identifying an appliance's anomalous behaviour? Applied energy 238(2019) 796-805.
[10]A. Ruano, A. Hernandez, J. Ureña, NILM techniques for intelligent home energy management and ambient assisted living: A review, Energies 12(11)(2019) 2203.
[11]L.L. Niu, H.J. Jia, Transient Event Detection Algorithm for Non-intrusive Load Monitoring, Automation of Electric Power Systems 35(9)(2011) 30-35.
[12]X.F. Song, Residential Electricity Consumption Behavior Analysis Based on Non-intrusive Load Monitoring, North China Electric Power University (Beijing), 2019.
[13]A.U. Rehman, T.T. Lie, B. Valles, Event-Detection Algorithms for Low Sampling Nonintrusive Load Monitoring Systems Based on Low Complexity Statistical Features, IEEE Transactions on Instrumentation and Measurement 69(3)(2020) 751-759.
[14]L. De Baets, J. Ruyssinck, C. Develder, On the Bayesian optimization and robustness of event detection methods in NILM, Energy and Buildings 145(2017) 57-66.
[15]K. Anderson, A. Ocneanu, Benitez, BLUED: A fully labeled public dataset for event-based non-intrusive load monitoring research, in: Proceedings of the 2nd KDD Workshop on Data Mining Applications in Sustainability (SustKDD), 2012.
[16]T.S. Huang, G.J. Yang, G.Y. Tang, A fast two dimensional median filtering algorithm, IEEE Trans. Acoustics Speech Signal Process 27(1)(1979) 13-18.
[17]K. Han, Z. Wang, Z. Chen, Fingerprint image enhancement method based on adaptive median filter, in: Proc. 2018 24th Asia-Pacific Conference on Communications (APCC), 2018.
[18]Y. Xu, M.Z. Luo, T. Li, ECG Signal De-noising and Baseline Wander Correction Based on CEEMDAN and Wavelet Threshold, Sensors 17(12)(2017) 2754.