

Research Article

State Characteristic Clustering for Nonintrusive Load Monitoring with Stochastic Behaviours in Smart Grids

Ruotian Yao ¹, Hong Zhou ¹, Dongguo Zhou ¹ and Heng Zhang²

¹School of Electrical Engineering and Automation, Wuhan University, Wuhan 430072, China

²Flight Automatic Control Research Institute, Xi'an 710065, China

Correspondence should be addressed to Hong Zhou; hzhouwuh@whu.edu.cn and Dongguo Zhou; dgzhou1985@whu.edu.cn

Received 7 August 2020; Accepted 20 July 2021; Published 31 July 2021

Academic Editor: Francesco Lo Iudice

Copyright © 2021 Ruotian Yao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Integrating the nonintrusive load monitoring (NILM) technology into smart meters poses challenges in demand-side management (DSM) of the smart grid when capturing detailed power information and stochastic consumption behaviours, due to the difficulties in accurately detecting load operation states in real household environments with the limited information available. In this paper, a state characteristic clustering (SCC) approach is presented for promoting the performance of event detection in NILM, which makes full use of multidimensional characteristic information. After identifying different types of state domains in an established multidimensional characteristic space, we design a sliding window difference search method (SWDS) to extract their initial clustering centre. Meanwhile, the mean-shift updating and iterating procedures are conducted to find the potential terminal stable state according to the probability density function. The above control strategy considers the transient events and stable states in a time-series dataset simultaneously, which thus allows the exact state of complex events to be obtained in a fluctuating environment. Moreover, a multisegment computing scheme is applied for fast computing in the state characteristic clustering process. Experiments of three different cases on both our real household dataset and REDD public dataset are provided to reveal the higher performance of the proposed SCC approach over the existing related methods.

1. Introduction

Recent years have witnessed an increasing attention to energy management in the smart grid, which forms a two-directional information flow between the supply side and demand side. Load monitoring is an effective method to obtain users' stochastic consumption behaviours in the demand side [1]. Suffering from the high cost of the intrusive load monitoring (ILM) approach, which needs to install sensors on every appliance, the nonintrusive load monitoring (NILM) technology can provide detailed load consumption information for both customers and utilities through uniform module [2]. Due to the increasing numbers of flexible loads, distributed generators, electric vehicles, etc. in the future smart grids and microgrids [3, 4], most utility companies are prone to encourage users to participate in demand response to maintain the load balance. Therefore, integrating the NILM algorithm into smart meters becomes an inevitable trend, in which utility companies can realize

efficient demand-side management (DSM) based on the demand behaviour of energy consumers [5].

Most NILM approaches on the demand side demonstrated that tracking the consumption of main appliances in residential buildings can help consumers to minimize power consumption [6]. On the other hand, for the supply side, obtaining consumers' multidimensional consumption information enables the precise prediction of stochastic energy demand [7], the optimization decision [8], and power-sharing control of microgrids [9]. To this end, an original NILM framework was proposed in [10]. Following this line, a lot of works had contributed to the development of the algorithm for accurate load monitoring. Many different optimization approaches are presented, which mainly relied on pattern recognition, signal processing, cloud computing, and machine learning for processing voltage, current, harmonics, and other behaviour features [11–14]. Among them, related results can be classified into event-based and state-based approaches. The latter case includes the hidden

Markov model (HMM) and its variants [1, 15, 16], whereas the former case generally depends on the transient state and the change signals due to the increasing aggregate power induced by the turning on load [17]. Moreover, the majority of machine learning (including neural network) based methods applied for NILM require a training set for load modelling, which are generally regarded as supervised learning methods. However, the standard dataset for training needs to be collected in advance for every specific appliance in households.

In order to directly obtain the operating state of loads, modified event detection methods are applied for recent event-based approaches [18–22]. For an event detection algorithm, the intuitive idea is to determine transient events by using a trigger threshold in detecting power signals [18]. To obtain the load waveform from the aggregate signal, a two-step iterative shrinkage threshold algorithm in the high-frequency domain was presented in [19]. Shi et al [20] introduced a hybrid similar time window algorithm to perform demand prediction in a lower-resolution dataset. The generality of the algorithm benefited from the improved cross-prediction approach, without knowing historical data. Also, for low sampling NILM systems, two novel event detection algorithms, variance and mean absolute deviation, are proposed [21]. By balancing the optimal window width and optimal performance, the aggregated active power data of real-world dataset can be captured based on a sliding window. Among the existing approaches, the switch signature will be extracted first, and then effective features will be stored into the database for further load identification [23, 24].

However, event-based approaches are dependent on the transient and stable state events; thus, missing any event may lead to the deviation of disaggregation results. The challenge of high precision event acquisition mechanism is how to identify complex event states in a real-world residential environment with fluctuation. Following this line, to tackle the problems of the optimal combination of the extractor and classifier, an intelligent event-driven NILM approach based on the convolutional neural network is proposed. The energy consumption behaviour and working condition of appliances in a residential dataset can be identified effectively [22]. By combining an unsupervised event detection method with an additive factorial HMM, energy disaggregation results are obtained by online cloud computing [14]. As a sequential analysis method, the cumulative sum (CuSum) method can detect small deviations in the process by calculating the deviations between the sample value and the target value.

CuSum is commonly used for change point detection in the non-Bayesian setting when both prechange and post-change probability density functions are known [25]. Furthermore, in order to solve the problem of quickest detection of dynamic events in systems, a Network-CuSum algorithm based on the breadth-first search and a thresholding approach were presented in [26]. However, because of the accumulative property, even small fluctuations can lead to an increase (or decrease) in the cumulative sum. Hence, the approaches, especially those extracting power signal events as features, are susceptible to noises, such as jitter and

outlier. Better results can be achieved by using thresholding and filtering algorithms under a fluctuation environment, but the basic idea should be improved further for the accurate NILM system.

Most unsupervised methods can be applied directly to time-series data without training datasets. Therefore, recent studies have attempted to identify load operating states by using clustering algorithms. The graph spectral clustering approach implements the power consumption forecast of a set of houses according to the current and previous appliance states [27]. In order to capture the timestamp of a transient event between two adjacent stable states, a modified density-based spatial clustering of applications with noise (DBSCAN) algorithm is proposed [28]. Moreover, multilayer perceptron classifiers are built for nonlinear loads by using harmonic characteristics. A stable state segmentation detector and a linear discriminant classifier group are used in [29] for load detection by considering the global characteristic similarity. As an iterative and nonparametric algorithm, mean shift and its variants are widely used in data clustering and denoising. By iterating through the probability density space, the clustering centre of the data can be located. Moreover, mean shift can identify clusters with different shapes, sizes, and densities due to its unsupervised property. In [30], an improved semisupervised kernel mean-shift clustering implements an automatic estimation of mean-shift parameters and automatically recovers an unknown number of clusters. In the mean-shift iteration process, a threshold parameter needs to be set as a stopping criterion to terminate the iteration process. Therefore, Aliyari Ghassabeh and Rudzicz [31] proposed a modified mean shift to guarantee the convergence of the data sequence without the stopping criterion.

Applying load detection approaches of NILM to real residential environments is a challenge for many algorithms. A fundamental problem is how to detect complex transient events (such as continuously varying transient states) in fluctuating environments effectively. Obviously, some threshold-based approaches (such as the trigger threshold method) do not have the property of detecting slowly changing events and long slope load transition stages. The approaches as mentioned above mainly employ the active/reactive power as the unique characteristic in the NILM system. To the authors' knowledge, there is very limited work on event detection using the mean-shift clustering iterative method in the multidimensional characteristic space. However, in the multidimensional space, clustering centres of stable states are more centralized than those of some transition states. Also, some state transition processes that cannot be reflected by one-dimensional data series can be clearly demonstrated in the multidimensional space. It is also a challenge to detect small-power devices when multiple loads are operating simultaneously. Motivated by these, we propose a state characteristic clustering (SCC) approach for demand side with stochastic behaviours, in which all data points collected by the NILM module will form several clusters in a multidimensional characteristic space. The proposed state domain definition is different from the often-adopted stable states of load events. A load event indicates

that an appliance is turned on or off, and the process in which the appliance runs normally and smoothly is called a stable state. However, the state domain defined in this paper represents the clustering region in the multidimensional characteristic space. Multidimensional steady-state data can be grouped in a specific domain around the cluster centre, while transient states are presented as a series of discrete points due to differences in time and space during load switching. Therefore, by the sliding window difference search (SWDS) method, initial cluster centres and state domains will be identified, respectively. Moreover, adjacent clusters can be shifted and merged by the mean-shift process as a terminal stable domain. Therefore, stochastic load events can be further captured. The main contributions of this paper are in the following aspects:

- (i) Although a hybrid detection approach and a subtractive clustering approach are used in [32] and [33], respectively, only aggregated active/reactive power data are often considered to detect the load event of residential appliances. Therefore, to avoid the false alarm of using a single power characteristic for event detection, multiple characteristics collected by the NILM module are used in this paper to detect load events from multiple perspectives. We then established a state domain in the multidimensional characteristic space, under which the distribution characteristics of data can be highlighted clearly in groups and the clusters can be therefore divided into stable domains and transition domains according to the dataset.
- (ii) The calculation burden of the algorithm will increase due to the large amount of high-frequency data [19]. For threshold-based methods [18], continuously varying transient events cannot be detected effectively in high-fluctuation environments. Therefore, the SWDS is proposed to extract the initial cluster centre of each original domain. Different from the threshold-based method and the CuSum method [26], the detection performance on low sampling rate time-series data can be improved by configuring appropriate SWDS parameters. In addition, due to the cumulative characteristics of CuSum, there will be persistent false alarms, resulting in a low detection accuracy. The proposed method can efficiently detect these complex states from 0.5 Hz aggregated power data in real household datasets.
- (iii) Inspired by the mean shift, we present the updating and iteration scheme for state domains in the multidimensional space. To be specific, the initial centres will shift to a new cluster centre by updating and iterating according to the data density, and similar cluster centres will be merged into one domain. Accurate load events and states can be therefore determined synchronously once the original state domains are updated. Since it will take a specific time for SWDS

and mean-shift procedure, a multisegment computing scheme is utilized for fast computing.

The outline of this paper is as follows: In Section 2, the definition of the state domain and multidimensional characteristic space are presented. Section 3 proposes the SCC approach, where the SWDS and the fundamental of mean-shift clustering procedure are described in detail. Section 4 gives three cases to verify the effectiveness of the proposed SCC approach, and the performance improvements of SCC on two datasets are proved by comparisons with other popular methods. Finally, conclusions are drawn in Section 5.

2. Establishment of State Domain in a Multidimensional Characteristic Space

The characteristic of a load can be defined by the multidimensional power features in a load identification system. Power features play an essential role in determining the type of load. The active power, reactive power, current harmonics, and other features, which are generated by the original voltage and current, are essential electrical signatures since the appliance is turned ON or OFF. Moreover, features such as running time, usage frequency, and operating time distribution also reflect the operating characteristic of a load. Hence, an event of a load can be considered as a signal switching (electrical or nonelectrical signals) in a state domain. An example of the transition of the state domain in a 3-dimensional space is shown in Figure 1.

Consider a d -dimensional nonzero dataset Ω^d with n elements, $\Omega^d = [\omega_1, \omega_2, \omega_3, \dots, \omega_n]$, where $\omega_i \in R^d$ denotes a multidimensional characteristic vector. The state domain can be represented as a cluster, so a load event can be detected by finding the cluster centre of a data block. Moreover, a stable domain may shift from the current position to a new state when a load event occurs. By identifying the two stable domains, the transient process associated with these two states is represented by their transfer trajectory, which is defined as a transition domain.

When all loads are running normally, we consider the current state is a stable state with an initial cluster centre ω_a . Therefore, we define that the dataset at this time interval is contained within a stable domain Φ_A . Then, when a load is switched from an OFF state to an ON state (or vice versa), all electrical signals will transfer to another state. In the multidimensional characteristic space, we denote this transfer trajectory as a transition domain Φ_{ab} . Furthermore, the final stable state is Φ_B with a cluster centre ω_b . Then, the set of data points can be given by $\Omega_A = \{\omega | \omega \in \Phi_A\}$, $\Omega_B = \{\omega | \omega \in \Phi_B\}$, and $\Omega_{ab} = \{\omega | \omega \in \Phi_{ab}\}$, respectively. Each point in the state domain is distributed in a limited area. On the contrary, Φ_{ba} represents the transition from state Φ_B to state Φ_A .

Generally speaking, a whole operating period of an appliance consists of at least a transition domain and two stable domains. The stable domain represents the regular operation of the appliance, and its power data vary within a specific range of

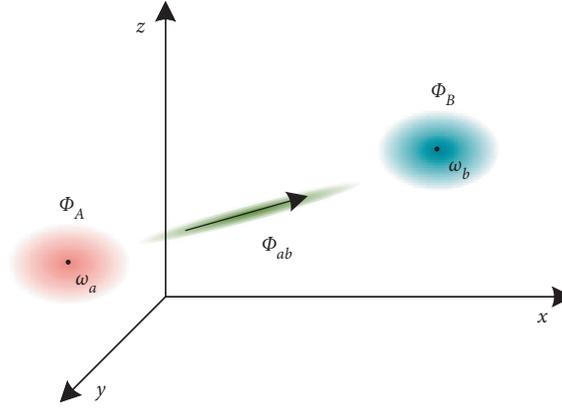


FIGURE 1: An example of stable domain and its transfer trajectory in a 3-dimensional characteristic space.

fluctuations. Also, the power data change during the state transition process. Notably, the transition state and the stable state usually appear alternately in a time-series dataset.

Figure 2 illustrates the changes of active power P , reactive power Q , and 3^{rd} current harmonic H of an air-conditioner during an operating period. It can be seen that the transient state and the stable state occur alternately ($\Phi_A - \Phi_{ab} - \Phi_B - \Phi_{bc} - \Phi_C$). In the transient state Φ_{ab} (sampling points 27–45), active power and 3^{rd} current harmonics slowly rise when the air-conditioner is started because of the operation of the compressor. At the same time, the reactive power decreases. In the stable state Φ_B (sampling points 45–155), the load reaches a stable running state with small fluctuations. In the last stage, all three characteristic curves show abrupt declines when the load is turned off, and the state returns to its initial state.

Missing or multiple detections in the slow-varying transient process often pose a challenge to accurate load monitoring systems. By mapping the multiple power data into a characteristic space, small fluctuations during the transition process can be effectively captured. As seen in Table 1, active power distribution of the air-conditioner in Figure 2 is analysed in detail. For the active power P , it can be seen that the standard deviation of transition domains is higher than that of stable domains (as shown in bold in Table 1). Therefore, we present an appropriate method to find the number of clusters and their initial centres accurately.

3. Proposed State Characteristic Clustering Approach

This section first describes the SWDS algorithm, which is employed in the determination of the initial cluster centre, and then presents the SCC approach and the iteration process of clustering points. Moreover, in order to make the algorithm more suitable for practical demand-side energy management, the multisegment computing scheme is used to increase the execution time.

3.1. The Sliding Window Difference Search. Finding appropriate initial cluster centres can improve the precision of the SCC approach effectively. If a point that belongs to a transition domain is selected, the shifting process of the clustering algorithm will be affected. Therefore, this paper presents an SWDS algorithm to extract the initial distribution characteristic of stable domain data by detecting the switching signatures of electrical loads.

For a time-series power data, define a sliding window Y (i) with width m :

$$Y_i = [\omega_{p,i}, \omega_{p,i+1}, \dots, \omega_{p,i+m-1}]. \quad (1)$$

Then, the mean value Y_i of characteristic p in the sliding window Y (i) is

$$\bar{Y}_i = \frac{\omega_{p,i} + \omega_{p,i+1} + \dots + \omega_{p,i+m-1}}{m}. \quad (2)$$

Assume that the difference between i and j of characteristic p can be represented as

$$\psi_{i,j} = |\omega_{p,i} - \omega_{p,j}|, \quad (3)$$

where $i < j \leq m$. Thus, we define a background noise parameter θ according to the dataset; that is, the difference of the data in the sliding window can be limited to an appropriate range.

To initialize the cluster centre, let $\eta_{i,s}$ be the difference between two cluster centres; therefore, we have

$$\eta_{i,s} = |\bar{Y}_i - \bar{Y}_s|. \quad (4)$$

Let γ represent as the minimum number of load events in the current data series. Thus, if $\eta_{i,s} > \gamma$, then the event in the current sliding window can be detected. Consequently, a stable domain ω_i will be selected while the following two conditions meet simultaneously:

$$\begin{cases} \psi_{i,j} < \theta, \\ \eta_{i,s} > \gamma. \end{cases} \quad (5)$$

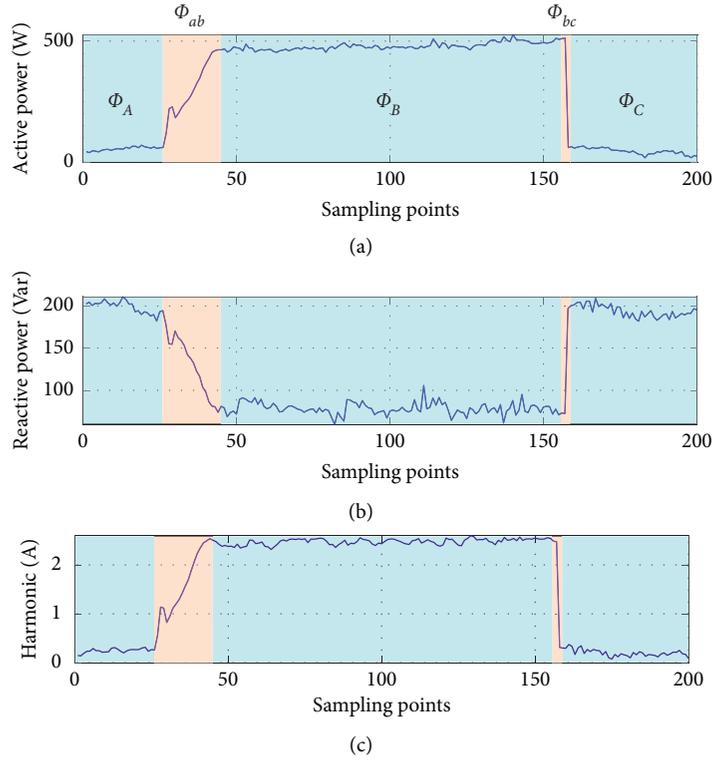


FIGURE 2: An operating period of an air-conditioner, where Φ_A , Φ_B , and Φ_C are stable states and Φ_{ab} and Φ_{bc} are transient states.

Figure 3 shows an example of the SWDS process. As we can see, a window with width m slides over the data series and a load event occurs on $Y(i)$. By calculating $\psi_{i,j}$ and $\eta_{i,s}$, i can be set as the initial cluster centre in this data block.

3.2. Mean-Shift-Based State Clustering. As a density-based clustering approach, the mean-shift clustering algorithm does not need any prior conditions and statistical parameter estimation. The extremum of probability density function can be calculated quickly by the gradient iteration and then converge to a cluster centre with the highest probability density. The main process of state clustering is as follows.

Consider a multidimensional characteristic space B_h with a fixed bandwidth h in a nonzero dataset Ω ; thus, the objective function can be written as

$$B_h(x) = \{y | \|y - x\|^2 \leq h^2\}, \quad (6)$$

where x is the centre of the multidimensional space, y represents a data point of the space, and h is the bandwidth that determines the search range for maximum density points. Usually, a smaller bandwidth has higher detecting performance for the loads with small power. However, a bandwidth that is too small will be sensitive to the power fluctuation and easily cause over classification.

For a d -dimensional space with n data points, the sample point density estimator by mean-shift process can be expressed as [30]

$$\hat{f}_h(x) = \frac{1}{nh^d} \sum_{x_i \in B_h} K\left(\left\|\frac{x - x_i}{h}\right\|^2\right), \quad (7)$$

where x_i represents the i -th data and $K(x)$ represents the radial basis kernel function. For a profile function $k(x)$, $K(x)$ is given as follows:

$$K(x) = c_{k,d} k(\|x\|^2), \quad (8)$$

where $c_{k,d}$ is the normalized constant. In this paper, the *Epanechnikov* kernel $K_E(x)$ is used by considering to keep the least-mean-square error.

$$K_E(x) = \begin{cases} \frac{1}{2}(d+2)(1-\|x\|^2), & \|x\| \leq 1, \\ 0, & \text{else,} \end{cases} \quad (9)$$

$$k_E(x) = \begin{cases} 1-x, & 0 \leq x \leq 1, \\ 0, & \text{else.} \end{cases}$$

Next, the initial centre will shift to a new cluster centre by the mean-shift process according to the data density. To obtain the maximum density point, let

$$\nabla \hat{f}(x) = 0. \quad (10)$$

Thus, the derivative of the objective function could be described as

TABLE 1: State characteristic distribution of the air-conditioner during an operating period.

Domain	Section	Number of sampling points	Standard deviation (W)	Average power (W)
Stable	Φ_A	27	8.21	21.8
Transition	Φ_{ab}	18	117.20	132.09
Stable	Φ_B	110	15.36	483.77
Transition	Φ_{bc}	2	319.14	286.42
Stable	Φ_C	43	12.20	44.05

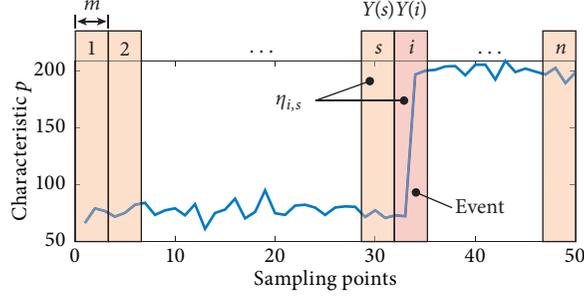


FIGURE 3: Cluster centre initialization process by SWDS.

$$\nabla \hat{f}_{h,k}(x) = \frac{2c_{k,d}}{nh^{d+2}} \sum_{x_i \in \mathcal{S}_h} (x - x_i) k' \left(\left\| \frac{x - x_i}{h} \right\|^2 \right). \quad (11)$$

Now, we reduce that

$$\nabla \hat{f}_{h,g}(x) = \frac{2c_{k,d}}{nh^{d+2}} \left[\sum_{x_i \in \mathcal{S}_h} g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right) \right] m_{h,g}(x), \quad (12)$$

for given $g(x) = -k^0(x)$, where

$$m_{h,g}(x) = \frac{\sum_{i=1}^n x_i g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right)}{\sum_{i=1}^n g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right)} - x, \quad (13)$$

is the updated direction of the new cluster centre.

3.3. Update and Iteration Procedure. By repeated iterations, the mean value moves in the direction of the highest density area. The iteration can be expressed as follows:

$$x_{k+1} = x_k + m_{h,g}(x_k). \quad (14)$$

Finally, the peak point x_p of the probability density function and its $B_h(x)$ can be obtained, which represents the current stable domain centre and the current high-dimensional space, respectively. More in detail, the main process of state characteristic clustering is as follows:

- (i) Select one or more points in the sampled data domain as an initial cluster centre.
- (ii) Set an iteration termination threshold e and the high-dimensional space bandwidth h according to the characteristics of power fluctuation, where e is determined by the allowable error range.

(iii) Calculate the mean-shift vector $m_{h,g}(x)$ of $B_h(x)$ in each high-dimensional space according to equation (13).

(iv) Move $B_h(x)$ along to the mean-shift vector $m_{h,g}(x)$ by equation (14).

(v) If $m_{h,g}(x) > e$, then record the density centre of this new stable domain and continue the iterative process. If $m_{h,g}(x) < e$, then it indicates that this high-dimensional space belongs to the same stable domain as the previous one, so as to merge the two cluster centres and end the iteration.

It should be noted that in order to make use of time-series information and reduce the interference of irrelevant data during the mean-shift process, this paper sets the clustering range of each initial cluster centre by SWDS.

Since the units and dimensions of each characteristic data are different, it is necessary to normalize the data. Let $\omega_{p,i}$ represent the value of characteristic p of point i , then the data can be normalized as follows:

$$x_i^p = \begin{cases} \frac{\omega_{p,i} - \omega_{p,\min}}{\omega_{p,\max} - \omega_{p,\min}}, & \omega_{p,i} \in [\omega_{p,\min}, \omega_{p,\max}], \\ 1, & \omega_{p,i} > \omega_{p,\max}, \\ 0, & \omega_{p,i} < \omega_{p,\min}, \end{cases} \quad (15)$$

where $\omega_{p,\min}$ and $\omega_{p,\max}$ represent the maximum and minimum values of characteristic p in the current dataset, respectively.

The measuring environment of power data will affect the selection of bandwidth h . Therefore, we have prerecorded

the d -dimensional characteristic of each appliance in stable domains. After normalization, h can be further defined as

$$h = \sqrt{\frac{1}{N} \sum_{x_i \in \Omega^d} \|x_i - \bar{x}\|^2}, \quad (16)$$

where \bar{x} represents the density centre obtained by taking the mean value of points in the space.

3.4. The Multisegment Computing Scheme. In practice, we can usually get at least one data point for each domain by using SWDS. According to the characteristics of these data points, the distribution of state domains can be obtained. For the width m of a sliding window, it determines the amount of data to be detected. Therefore, based on this property, we conclude that for a larger m , the SWDS can reduce error detections by power fluctuations in the detecting process. On the contrary, a smaller window width can improve the detecting performance in the case of load events occurring at a small time interval.

Therefore, in real-case scenario, there exists a trade-off between the optimal window width and the best detecting performance. Furthermore, for a large amount of data, in order to improve the processing efficiency of the proposed method, the multisegment computing scheme is used in this paper. Figure 4 shows a simple demonstration of the proposed SCC approach with the multisegment computing scheme.

We conducted separate clustering for each w sampled data and then obtained the power characteristic of each data block. First, the minimum length of each block should be defined appropriately. It should be noted that w is much larger than m in real-world dataset, and each data block will contain multiple SWDS and mean-shift processes. Therefore, SWDS detection results are not unique. Then, for each detected initial cluster centre, mean shift will be processed separately to determine the final cluster centre. After all the stable/transition is determined, the relative key data point will be recorded and stored in the form of numerical ranges along time series. Note that if the number of stable domains $M > 2$, then it will be regarded as there are events in this block. Then, turn to the next data block and repeat the process. By combining with SWDS and mean-shift clustering, the state characteristic clustering can be calculated faster in blocks along with time series.

4. Experimental Results and Discussion

In this part, experiment results of three different cases are carried out to demonstrate the performance of the proposed SCC approach. Here, household appliance dataset, which includes air-conditioner, rice cooker, microwave oven, induction cooker, electric kettle, TV, refrigerator, electric fan, and desktop computer, is stored in the database. Each appliance was turned ON/OFF for several times to verify the effectiveness of the proposed method. Notably, the NILM data sampling module in our lab, which collects the total active power, reactive power, and harmonics data, uploads data every two seconds. Besides, the proposed SCC is

compared with existing algorithms in each case to further verify the proposed method in terms of effectiveness.

The event detection process is taken sequentially with time, and the purpose is to identify the occurrence and termination of load events accurately. The trigger threshold (TT) method is a quick and simple method, especially for high-power resistive appliances [18]. It uses a threshold parameter to capture appliance switch signals. As we mentioned in the previous section, event points in the dataset can be extracted by comparing to the sum of the minimum cumulative sum statistics. Therefore, the bilateral-CuSum (BC) method with known prechange and post-change distributions will be effective for multistate events [26]. For comparison, the TT, BC, and proposed SCC methods are used to verify the performance of event detection in different cases, respectively.

4.1. The Detection of Continuously Varying State. In this case, the work mainly focuses on a load with a continuously varying transient state. For instance, the power curve of an air-conditioner increases slowly since it is turned on. To our knowledge, this is a challenge for many existing event detection algorithms because the power changes slowly rather than a transient signal.

Figure 5 illustrates the starting process of the air-conditioner. It can be seen that there exist three states: the stable state without any load running (Φ_A), transient state, and stable state with the air-conditioner running (Φ_B). Roughly speaking, the active power changes about 450 W between two stable domains Φ_A and Φ_B . The actual transient event occurred at point 105. Then, after undergoing a slow climb about 15 points, the state reaches Φ_B domain.

In order to reduce the influence caused by data fluctuation, here, the threshold parameter for judging the load event is set as $P_{\min} = 70$ W for each algorithm. That is, a power change over P_{\min} is considered as a valid load event. For the BC method, the minimum mutation parameter Δ_{\min} is set to 200 and the noise parameter θ is set to 20. For the proposed SCC method, the significant parameters mentioned in Section 3.2 are shown in Table 2.

For the proposed SCC, Figure 6 illustrates the state transition process with a continuously varying state in a P - Q - H 3-dimensional space. Note that all power characteristics are normalized before the state clustering process. Figure 6(a) shows an overview of the state transition process. It can be seen that the two domains Φ_A and Φ_B are very far apart in this 3-dimensional space. The data points connecting the two domains belong to a continuously varying event. It is obvious that this long state transition process can only be shown in a multidimensional characteristic space. For stable domain Φ_A in Figure 6(b), SWDS first determines the initial cluster centre (black triangle) and then shifts to the terminal cluster centre (red triangle) by the mean-shift process, where the terminal cluster is the region with highest data density. Moreover, we use the green line to mark the shift trajectory of the cluster centre.

Table 3 shows the event detection results of the TT, BC, and proposed SCC methods. The comparison with ground

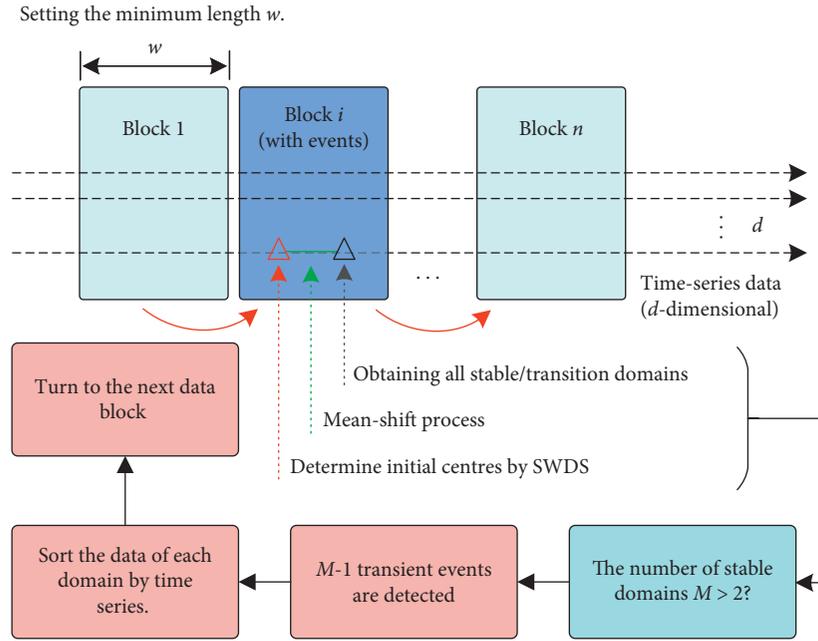


FIGURE 4: The multisegment computing mechanism of SCC by block calculation.

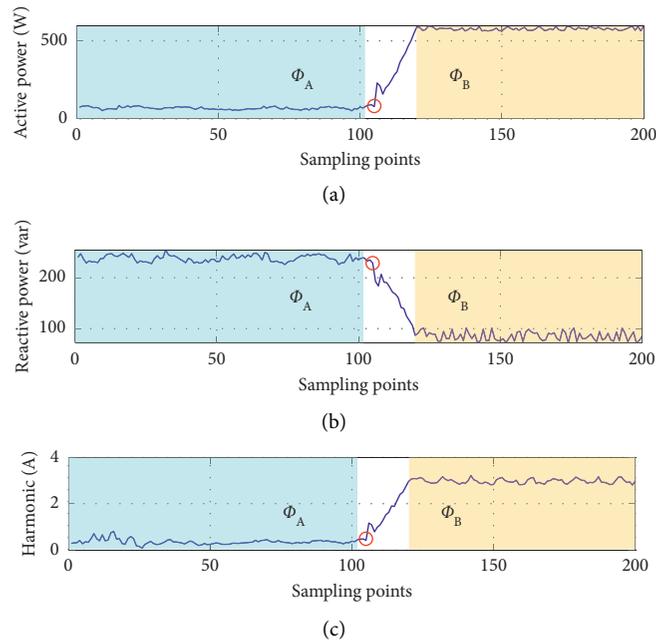


FIGURE 5: Three states of an air-conditioner during the whole operating process, where the continuously varying transient state between Φ_A and Φ_B lasted 50 seconds.

truth (GT) data shows that the TT method fails to detect the event. This complex state cannot be detected by just thresholds. The BC method captures two load events at 111 and 116 points, respectively, that lead to multidetection. In contrast, the proposed SCC method avoids false detection from the load undergoing the long transient process and gets the almost correct starting point. Therefore, the proposed SCC method in multidimensional space has higher detection accuracy for such continuously varying events.

4.2. Detection Performance of Small Power Loads under a Fluctuating Environment. In this case, it mainly demonstrates the proposed method in separating the load with small power loads from a fluctuating environment with other loads overlapping. Figure 7 illustrates the case that an induction cooker with 2000 W active power is turned to ON state at sampling point 10 and then a computer runs at point 26. From the visual of inspection, the computer almost cannot be found just through the active power curve. It is

TABLE 2: Some significant parameters in the SCC method.

Parameter	Symbol	Value
Bandwidth	h	0.05
Data block length	ω	200
Iteration termination threshold	e	0.01
Noise parameter	θ	30
Minimum active power	P_{\min}	70
Minimum spacing between domains	Γ	70
Window width	m	3

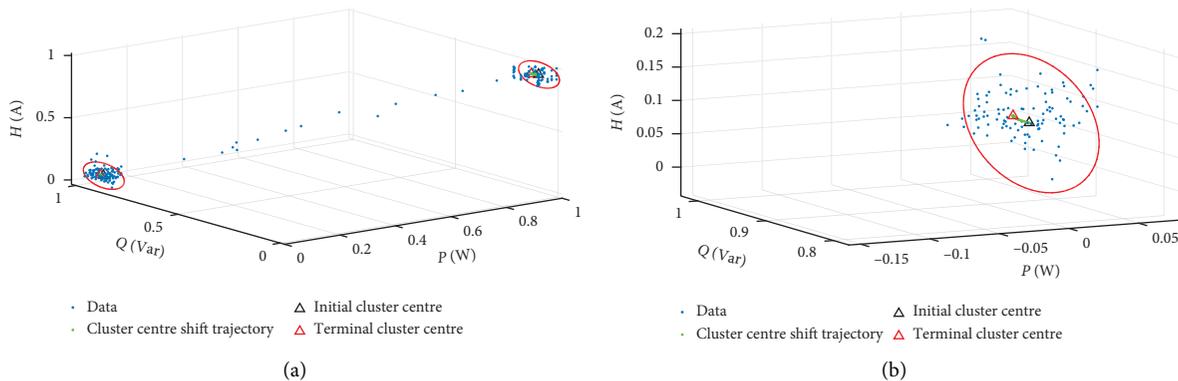
FIGURE 6: The SCC process of an air-conditioner with a continuously varying state in a (P) - (Q) - (H) 3-dimensional space (after normalization). (a) The overview of state transition process. (b) The cluster centre of stable domain Φ_A and its shift trajectory.

TABLE 3: Event detection results of three methods for continuously varying state.

Methods	Event points
GT	105
TT	-
BC	111 and 116
SCC (proposed)	105

because that the induction cooker produced a 70 W fluctuation, while the power of the computer is just 70–90 W. At this point, the load events of the computer are easy to overlap with the background fluctuation. Nevertheless, the proposed SCC method can distinguish these two load events by multiple characteristics in the 3-dimensional space. As seen in Figure 7, red circles denote the boundary of the B_h by mean-shift clustering.

For more understanding, Figure 8 illustrates the whole process of state clustering. First, there are five initial stable domains named Φ_A – Φ_E , which are detected by SWDS. In addition, the domain Φ_C has four initial cluster centres and Φ_E contains two initial cluster centres. For domains Φ_A , Φ_B , and Φ_D , they just have a corresponding initial cluster centre. As can be seen in Figure 8, the initial centres of the domain Φ_C are merged into a single cluster centre by using mean-shift clustering, and the same as the initial centres of the domain Φ_E , while other domains retain their original cluster centre. Therefore, the number of cluster centres is $M=5$, corresponding to 5 stable domains. Therefore, this case has $M-1$ transient events. Finally, four load events can be detected according to those clusters.

Furthermore, to demonstrate that the proposed SCC method is less sensitive to the parameter setting, here, experiments on different parameter settings are carried out for comparison. In order to measure the efficiency of each event detection algorithm, the accuracy metric F-measure in machine learning is introduced. Generally, the event detection results can be divided into four categories: true positive (TP), which indicates the algorithm detect load events correctly; true negative (TN), which indicates that no events occur and the algorithm gives no alarms; false positive (FP), which indicates that the algorithm detects an event but does not actually occur; and false negative (FN), which indicates that the algorithm did not correctly detect the events that occurred. Moreover, the F-measure (FM) is a metric by combining the multiple-detection rate and missing detection rate. Therefore, a higher FM represents a better performance.

Tables 4–7 show the F-measure results of threshold P_{\min} , minimum mutation parameter Δ_{\min} , minimum spacing Γ , and bandwidth h on TT, BC, and SCC methods, respectively.

Table 4 shows the influence of different trigger threshold parameters. The TT method, which is judged by a single

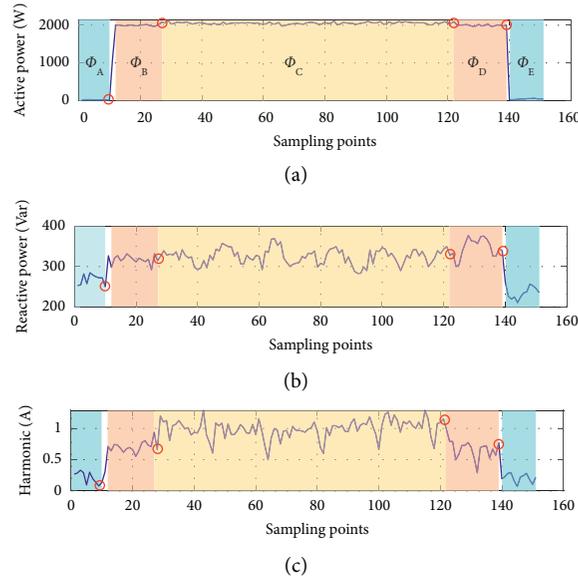


FIGURE 7: A computer is turned on in a noisy environment with an induction cooker running, where Φ_C represents its stable domain.

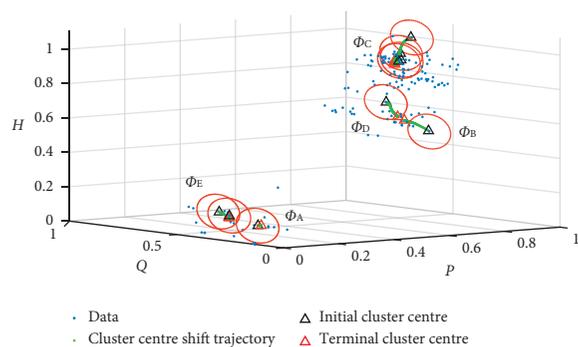


FIGURE 8: The state transition and cluster centre merging process under a fluctuating environment (after normalization).

point, may make small power loads' power signature confuse with background noises. Table 5 shows that the optimal value of the parameter in bilayer CuSum should be properly set by accurate statistics on the power in the database. Also, the bilateral-CuSum method is sensitive to the power of a single point at the beginning of a transient event, and the accuracy of FM declines quickly after the optimal value. Tables 6 and 7 reflect that the proposed method has a wide range of optimal parameter settings. Therefore, the performance of the SCC method has its advantage in terms of the parameter configuration for load detection.

4.3. Real Household Scenario with Long-Term Stochastic Behaviour. In this case, the user's stochastic behaviours during the actual operation of multiple appliances are considered. Therefore, ten appliances in real household scenario are randomly switched 100 times for a long-term operation test. The aggregate time-series data containing about 4000 sampling points were recorded in MySQL

database previously. Similar to the first two cases, in order to show the performance evaluation of the proposed SCC method, the TT and BC methods are used for comparison. Here, the optimal parameters are selected according to the tables in Section 4.2. In the following, the results will be discussed in detail.

Figure 9 shows the distribution of data points in this time-series dataset, where Figure 9(a) shows the aggregate active power curve during this long-term operation. Red points represent the SWDS result of detected stable state. Figure 9(b) shows the overall distribution of data points in a P - Q - H 3-dimensional space (without normalization). It can be found that the stable points are almost correctly detected by the proposed method.

Taking the data block (2400–2800) as an example, Figure 10(a) indicates the cluster centre's shift and merging process. The adjacent initial cluster centres are finally converged to one stable domain, corresponding to load events. Finally, each stable domain is determined by a certain cluster centre. As seen in Figure 10(b), the original initial

TABLE 4: The influence of different thresholds P_{\min} on the performance of the TT method.

P_{\min}	TP	FP	FN	FM
60	4	9	0	0.46
70	4	5	0	0.61
80	4	3	0	0.72
90	4	3	0	0.72
100	3	2	1	0.66
110	2	1	2	0.57

TABLE 5: The influence of different minimum mutation parameters Δ_{\min} on the performance of the BC method.

Δ_{\min}	TP	FP	FN	FM
60	4	7	0	0.53
70	4	3	0	0.72
80	4	1	0	0.89
90	4	0	0	1.00
100	3	0	1	0.86
110	2	0	2	0.67
120	2	0	2	0.67
130	2	0	2	0.67

TABLE 6: The influence of different minimum spacings Γ between domains on the performance of the SCC method.

Γ	TP	FP	FN	FM
30	4	4	0	0.67
40	4	2	0	0.80
50	4	2	0	0.80
60	4	0	0	1.00
70	4	0	0	1.00
80	4	0	0	1.00
90	4	0	0	1.00
100	2	0	2	0.67

TABLE 7: The influence of different bandwidths h on the performance of the SCC method.

h	TP	FP	FN	FM
0.01	3	1	1	0.75
0.02	4	0	0	1.00
0.03	4	0	0	1.00
0.04	4	0	0	1.00
0.05	4	0	0	1.00

cluster centre is replaced by the terminal cluster centre after executing the SCC approach. For ease of presentation, we normalized the original power data into $[0, 1]$.

In this case, there are a total of 16 initial cluster centres after SWDS processing. Furthermore, results show that there are 13 stable domains and 12 transition domains, which means that 12 events are detected. For the ease of description, the statistical results of this environmental scenario are shown in Figure 11, where each block length is $w = 400$. The SCC results are almost consistent with the GT data.

Moreover, the FM performance of the proposed SCC method and the comparison results with other methods are shown in Table 8.

It can be seen from Table 8 that the FN of the TT method is low because the threshold is set according to the actual

power information. Events will be recorded when the power value exceeds P_{\min} . However, when the power fluctuates greatly, or the power rises slowly, there is a phenomenon that a single event will be judged as a group of events; thus, the FP is higher than other methods. The FP of the bilateral-CuSum algorithm is also high because of the limit of window width m , which will cause false alarms on slowly rise events easily. Because massive noise can be easily confused with small power signatures, its FN is also higher than other methods. For the SCC in this paper, both FP and FN are relatively low, because this method has a good detection performance for long-term transient events. Also, small changes in power signals can be more easily captured in the multidimensional spaces. As an approach with multiple characteristic properties, SCC can accurately detect small switching signatures

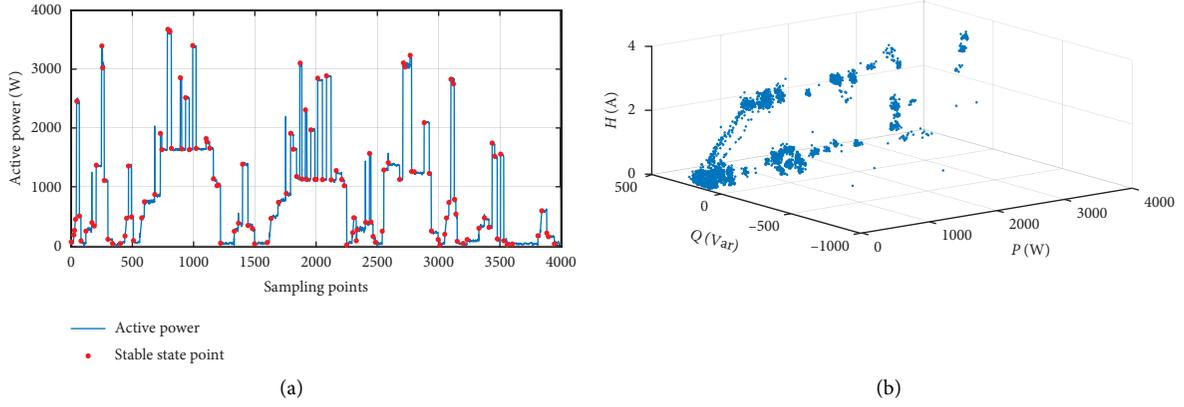


FIGURE 9: The distribution of data points generated by 10 appliances with stochastic behaviours. (a) The aggregate active power curve during a long-term operation, where red points represent the stable state. (b) Data points in the (P) - (Q) - (H) 3-dimensional space.

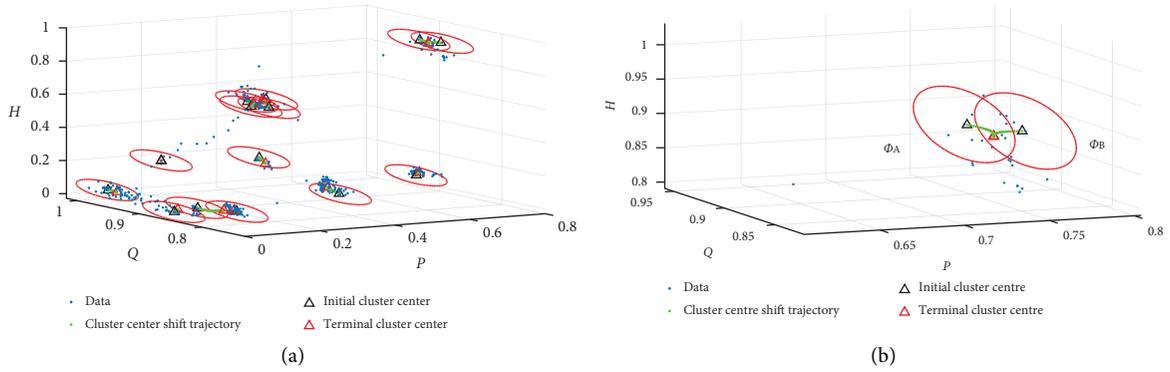


FIGURE 10: The SCC process in data block 2400–2800 (after normalization). (a) The detailed state domains in the data block by executing SCC. (b) A merging process of adjacent cluster centres in 2710–2780.

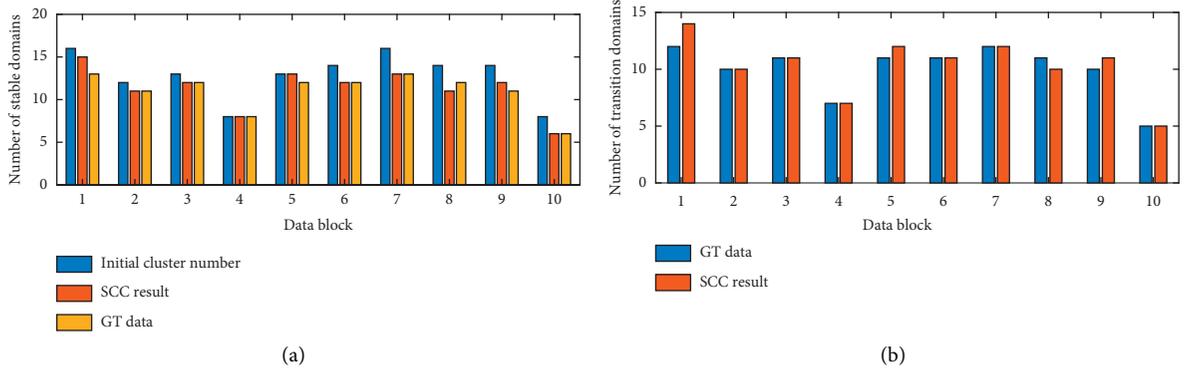


FIGURE 11: Comparison of SCC results and ground truth data in this stochastic environmental scenario with 10 data blocks.

in the fluctuation environment. Therefore, the proposed SCC reduces the number of false detection or missing detection in reality. However, it should be noted that when the event occurs at block boundaries, the next stable state may not be detected accurately.

Furthermore, to verify the effectiveness of the proposed method, the precision (P) and recall (R) of all three methods mentioned above are computed. They can be described as

$$P = \frac{TP}{TP + FP}, \quad (17)$$

$$R = \frac{TP}{TP + FN}. \quad (18)$$

Figure 12 draws the P-R curve of these methods. By calculating the equilibrium point when $P = R$, it can be seen

TABLE 8: The F-measure (FM) performance of 3 methods in this environmental scenario.

Method	TP	FP	FN	FM
TT	99	28	1	0.87
BC	90	20	10	0.85
SCC (proposed)	98	5	2	0.97

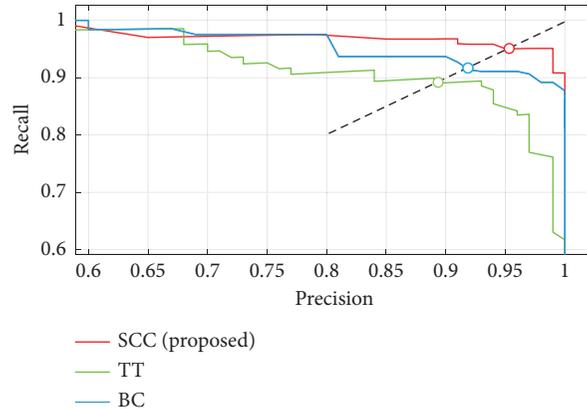
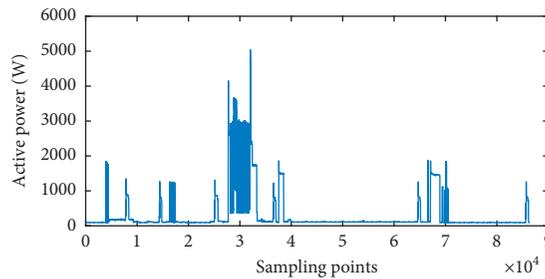
FIGURE 12: P - R performance of each method in the stochastic environmental scenario.

FIGURE 13: The aggregate power curve of house 3 in REDD.

that the proposed SCC in this paper performs better than others, where the algorithm with larger value has a better performance.

4.4. Validation with Data from REDD Public Dataset.

This case aims to validate the effectiveness of the proposed SCC method on a widely used public dataset. Moreover, the experimental results are expanded by comparing the proposed SCC with other existing methods.

The reference energy disaggregation data set (REDD) is one of the most used public datasets for NILM system [34]. It includes the low-frequency data at a 1 Hz sampling rate of all six houses and two sets of high-frequency data at 15 kHz sampling rate. In order to obtain multidimensional power profiles, here, we perform fast Fourier transform (FFT) on high-frequency data to obtain active/reactive power signals

every 2 seconds and 1–10 order current harmonics. Figure 13 shows the aggregate power curve of house 3 on April 22, 2011.

The house 3 in REDD, which we used in this section, contains 20 kinds of electric appliances. Therefore, we change the original experimental parameters according to the power characteristic of REDD. Some significant parameters for REDD dataset are shown in Table 9.

Then, we use the above parameters to test the selected data series of house 3, and the results are shown in Table 10.

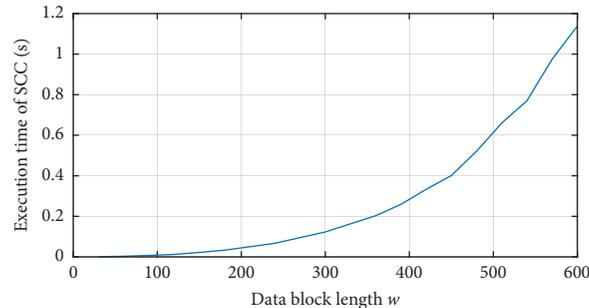
As can be seen in Table 10, the TT and BC methods have much more false alarms than that of the SCC method, resulting in low overall FM accuracy. Therefore, these test results suggest that our proposed method based on state clustering is insensitive to the dataset difference. The SCC has better detection performance on both REDD and our own dataset than other existing works.

TABLE 9: Some significant parameters for REDD dataset.

Parameter	Symbol	Value
Bandwidth	h	0.05
Data block length	ω	60
Iteration termination threshold	e	0.01
Noise parameter	θ	10
Minimum active power	P_{\min}	70
Minimum spacing between domains	Γ	70
Window width	m	3

TABLE 10: Experimental results on REDD dataset.

Method	TP	FP	FN	Recall	Precision	FM
TT	147	44	11	0.9304	0.7696	0.8424
BC	154	35	4	0.9747	0.8148	0.8876
SCC (proposed)	131	2	27	0.8291	0.9850	0.9003

FIGURE 14: Execution time analysis of different data blocks (w).

4.5. Execution Time Analysis. The NILM system is mostly used for demand-side management in the smart grid. Therefore, these are expected to be real-time methods (i.e., calculations completed before the next data block is collected). Therefore, the execution time of the proposed SCC method is analysed in this section.

The execution time of SCC is mainly determined by the preprocessing time, response delay, and state clustering time. The preprocessing time includes the time needed to calculate power RMS, current harmonics, etc. Response delay time includes data block acquisition time, sliding window width, and sliding frequency. It is worth saying that the SCC runs independently in each data block; therefore, in order to improve the execution speed, the parameter configuration of block length is essential. To further analyse the execution time of SCC, in this paper, we use different data block lengths w to detect a 24 h time-series household data and record the average execution time of each process.

Figure 14 shows the effect of different data blocks w on execution time. The results show that even in a data block with a 600-point length, the SCC execution time is only 1.14 s. Therefore, the proposed SCC method can carry out real-time load monitoring in real residential

environments. Overall, for an accurate demand response process, the execution time can be further reduced by changing the data block length and data sampling rate.

5. Conclusions

Based on the sliding window difference search and mean shift, a state characteristic clustering approach for demand-side nonintrusive load monitoring is proposed. The proposed SCC approach can achieve the operating state of several electrical appliances in a household effectively, which utilizes multidimensional aggregated data collected by an NILM module. In this approach, SWDS is used to find the initial cluster centre of each state domain, and then adjacent clusters can be shifted or merged as one stable domain in a multidimensional space. Multidimensional power features can reflect the real operation characteristic of loads and are more reliable than a single power feature, especially in a high-fluctuation environment. Three cases on both our real household dataset and REDD dataset show that the proposed SCC method can improve the detection performance of load events. Also, by comparing with two existing methods, the results demonstrate that the proposed SCC method has a higher F-measure accuracy in a complex real

residential environment and public dataset. Besides, for real-time demand response in the future smart grid, we proposed a multisegment computing scheme to improve the execution time.

In the following work, a more reliable model with complex states and appliances will be further considered and then applied to achieve load identification. Meanwhile, business environments with office equipment are also worth investigating.

Data Availability

The private household data used to support the findings of this study are included within this paper. The encrypted multidimensional power data were not given in full since they belong to a smart meter installed in a customer's premise.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under grant no. 61873195.

References

- [1] R. Bonfigli, E. Principi, M. Fagiani, M. Severini, S. Squartini, and F. Piazza, "Nonintrusive load monitoring by using active and reactive power in additive factorial hidden markov models," *Applied Energy*, vol. 208, pp. 1590–1607, 2017.
- [2] M. S. Piscitelli, S. Brandi, and A. Capozzoli, "Recognition and classification of typical load profiles in buildings with non-intrusive learning approach," *Applied Energy*, vol. 255, Article ID 113727, 2019.
- [3] J. Lai, X. Lu, X. Yu, and A. Monti, "Cluster-oriented distributed cooperative control for multiple ac microgrids," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 11, pp. 5906–5918, 2019.
- [4] J. Lai, X. Lu, X. Yu, and A. Monti, "Stochastic distributed secondary control for ac microgrids via event-triggered communication," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 4, pp. 2746–2759, Jul. 2020.
- [5] F. Luo, G. Ranzi, W. Kong, Z. Y. Dong, S. Wang, and J. Zhao, "Non-intrusive energy saving appliance recommender system for smart grid residential users," *IET Generation, Transmission & Distribution*, vol. 11, no. 7, pp. 1786–1793, 2017.
- [6] S. Ghosh, A. Chatterjee, and D. Chatterjee, "Improved non-intrusive identification technique of electrical appliances for a smart residential system," *IET Generation, Transmission & Distribution*, vol. 13, no. 5, pp. 695–702, 2019.
- [7] J. A. Mueller and J. W. Kimball, "Accurate energy use estimation for nonintrusive load monitoring in systems of known devices," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 2797–2808, 2018.
- [8] D. García, I. Díaz, D. Pérez, A. A. Cuadrado, M. Domínguez, and A. Morán, "Interactive visualization for nilm in large buildings using non-negative matrix factorization," *Energy and Buildings*, vol. 176, pp. 95–108, 2018.
- [9] J. Lai and X. Lu, "Nonlinear mean-square power sharing control for ac microgrids under distributed event detection," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 1, pp. 219–229, 2020.
- [10] G. W. Hart, "Prototype nonintrusive appliance load monitor," MIT Energy Laboratory Technical Report, and Electric Power Research Institute Technical Report, Electric Power Research Institute, Washington, DC, 1985.
- [11] S. Naderian and A. Salemnia, "Method for classification of pq events based on discrete gabor transform with fir window and t2fk-based svm and its experimental verification," *IET Generation, Transmission & Distribution*, vol. 11, no. 1, pp. 133–141, 2017.
- [12] M. Aiad and P. H. Lee, "Energy disaggregation of overlapping home appliances consumptions using a cluster splitting approach," *Sustainable Cities and Society*, vol. 43, pp. 487–494, 2018.
- [13] B. Buddhahai, W. Wongseeree, and P. Rakkwamsuk, "A non-intrusive load monitoring system using multi-label classification approach," *Sustainable Cities and Society*, vol. 39, pp. 621–630, 2018.
- [14] M. A. Mengistu, A. A. Girmay, C. Camarda, A. Acquaviva, and E. Patti, "A cloud-based on-line disaggregation algorithm for home appliance loads," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 3430–3439, 2019.
- [15] W. Kong, Z. Y. Dong, D. J. Hill, J. Ma, J. H. Zhao, and F. J. Luo, "A hierarchical hidden markov model framework for home appliance modeling," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3079–3090, 2018.
- [16] S. Makonin, F. Popowich, I. V. Bajic, B. Gill, and L. Bartram, "Exploiting hmm sparsity to perform online real-time non-intrusive load monitoring," *IEEE Transactions on Smart Grid*, vol. 7, no. 6, pp. 2575–2585, 2016.
- [17] M. G. Cominola and A. E. Rizzoli, "A hybrid signature-based iterative disaggregation algorithm for non-intrusive load monitoring," *Applied Energy*, vol. 185, pp. 331–344, 2017.
- [18] J. Liang, S. K. K. Ng, G. Kendall, and J. W. M. Cheng, "Load signature study—Part I: basic concept, structure, and methodology," *IEEE Transactions on Power Delivery*, vol. 25, no. 2, pp. 551–560, 2010.
- [19] X. Wu, X. Han, and K. X. Liang, "Event-based non-intrusive load identification algorithm for residential loads combined with underdetermined decomposition and characteristic filtering," *IET Generation, Transmission & Distribution*, vol. 13, no. 1, pp. 99–107, 2019.
- [20] X. Shi, H. Ming, S. Shakkottai, L. Xie, and J. Yao, "Non-intrusive load monitoring in residential households with low-resolution data," *Applied Energy*, vol. 252, p. 113283, 2019.
- [21] U. Rehman, T. T. Lie, B. Valles, and S. R. Tito, "Event-detection algorithms for low sampling nonintrusive load monitoring systems based on low complexity statistical features," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 3, pp. 751–759, 2020.
- [22] D. Yang, X. Gao, L. Kong, Y. Pang, and B. Zhou, "An event-driven convolutional neural architecture for non-intrusive load monitoring of residential appliance," *IEEE Transactions on Consumer Electronics*, vol. 66, no. 2, pp. 173–182, 2020.
- [23] S. Lin, L. Zhao, F. Li, Q. Liu, D. Li, and Y. Fu, "A nonintrusive load identification method for residential applications based on quadratic programming," *Electric Power Systems Research*, vol. 133, pp. 241–248, 2016.
- [24] X. Wu, D. Jiao, K. Liang, and X. Han, "A fast online load identification algorithm based on vi characteristics of high-

- frequency data under user operational constraints,” *Energy*, vol. 188, Article ID 116012, 2019.
- [25] P. S. Jain, S. Bhashyam, and A. P. Kannu, “Algorithms for change detection with sparse signals,” *IEEE Transactions on Signal Processing*, vol. 68, pp. 1331–1345, 2020.
- [26] S. Zou, G. Fellouris, and V. V. Veeravalli, “Quickest change detection under transient dynamics: theory and asymptotic analysis,” *IEEE Transactions on Information Theory*, vol. 3, no. 65, pp. 1397–1412, 2019.
- [27] C. Dinesh, S. Makonin, and I. Bajic, “Residential power forecasting using load identification and graph spectral clustering,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 66, no. 11, pp. 1900–1904, 2019.
- [28] Z. Zheng, H. Chen, and X. Luo, “A supervised event-based non-intrusive load monitoring for non-linear appliances,” *Sustainability*, vol. 10, no. 4, p. 1001, 2018.
- [29] J. Yu, Y. Gao, Y. Wu, D. Jiao, C. Su, and X. Wu, “Non-intrusive load disaggregation by linear classifier group considering multi-feature integration,” *Applied Sciences*, vol. 9, no. 17, p. 3558, 2019.
- [30] S. Anand, S. Mittal, O. Tuzel, and P. Meer, “Semi-supervised kernel mean shift clustering,” *IEEE Trans Pattern Anal Mach Intell*, vol. 36, no. 6, pp. 1201–1215, 2014.
- [31] Y. Aliyari Ghassabeh and F. Rudzicz, “Modified mean shift algorithm,” *IET Image Processing*, vol. 12, no. 12, pp. 2172–2177, 2018.
- [32] M. Lu and Z. Li, “A hybrid event detection approach for non-intrusive load monitoring,” *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 528–540, 2020.
- [33] N. Henao, K. Agbossou, S. Kelouwani, Y. Dube, and M. Fournier, “Approach in nonintrusive type i load monitoring using subtractive clustering,” *IEEE Transactions on Smart Grid*, vol. 8, no. 2, pp. 812–821, 2017.
- [34] J. Z. Kolter and M. J. Johnson, “REDD.: a public data set for energy disaggregation research,” in *Proceedings of the SustKDD Workshop Dataon Data Mining Applications in Sustainability*, vol. 1, pp. 1–6, San Diego, CA, USA, 2011.