Peer-reviewed version available at Appl. Sci. 2019, 9, 3558; doi:10.3390/app9173558

Article

Non-Intrusive Load Disaggregation by Linear Classifier Group Considering Multi-Feature Integration

Jinying Yu, Yuxin Wu and Chang Su*

School of Electric and Electronic Engineering, North China Electric Power University, Changping District, Beijing 102206, China; 1161200227@ncepu.edu.cn (J.Y.); 1161200120@ncepu.edu.cn (Y.W.); 50102011@ncepu.edu.cn (C.S.)

* Correspondence: 50102011@ncepu.edu.cn; Tel.: +86-010-61771641 (C.S)

Abstract: Non-intrusive load monitoring (NILM) is a core technology for demand response (DR) and energy conservation services. Traditional NILM methods are rarely combined with practical applications, and most studies aim to decompose the whole loads in a household, which leads to low identification accuracy. In this paper, an NILM approach based on multi-feature integrated classification (MFIC) is explored, which combines some non-electrical features such as ON/OFF duration, usage frequency of appliances, and usage period to improve load differentiability. The implementation of MFIC algorithm is consistent with traditional event-based method. The uniqueness of our algorithm is that it designs an event detector based on steady-state segmentation and a linear discriminant classifier group based on multi-feature global similarity. Simulation results using an open-access dataset demonstrate the effectiveness and high accuracy of MFIC algorithm, with the state-of-the-art NILM methods as benchmarks.

Keywords: non-intrusive load monitoring; load disaggregation; linear classifier; demand response

1. Introduction

The increased public awareness of energy conservation in recent years motivates electricity consumers to actively participate in energy management [1]. Demand response (DR) is one of the solutions for demand side management, which response to certain conditions by reducing or shifting loads to a different time period. With the advent of smart grid, residential DR has great research potential. Since different types of appliances have different opportunities and ways to participate in DR, it is crucial to study detailed appliance-level power consumption data. In addition, the visualization of detailed consumption of high-power appliances will help customers to replace some inefficient devices, so as to save energy [2].

Traditional intrusive load monitoring needs to install large numbers of sensors to acquire signal of each appliance. In the process of sensors' installation and maintenance, the power supply needs to be temporarily interrupted, causing inconvenience for users. Since the practicability of intrusive method is poor, Hart innovatively proposed the concept of non-intrusive load monitoring (NILM) in the 1980's [3]. Since its high cost efficiency and less installation effort, NILM is more attractive to customers and utilities. The main idea of NILM is compositing electrical signals at power entrance to track the working status and detailed energy consumption information of individual appliances.

Early studies in NILM aim at detecting state-changing event by identifying distinct electrical features of individual appliances, which are called "load signature" and can be divided into two categories: steady-state and transient state. The most common used steady-state signatures are active and reactive power [4-5]. They are effective in identifying high-power devices, but are challenged to separate low-power appliances due to the possibility of power overlap. Later works extended the steady-state signature to many aspects, such as harmonics [6], current and voltage

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 June 2019 doi:10.20944/preprints201906.0169.v1

Peer-reviewed version available at Appl. Sci. 2019, 9, 3558; doi:10.3390/app9173558

waveforms [7], voltage-current trajectory [8-10], inactive current [11] etc. All of them can effectively disaggregate certain types of appliances. In order to define more accurate load signatures, researchers [12-14] begin to extract features from the period of two stable operations, which is called transient signature. Since transient signatures usually have a relatively shorter duration, the probability of feature overlapping is lower. However, they rely on the samples measured with high rates, so the practicability is limited.

With the large-scale deployment of smart meters, NILM approaches that work with lower sampling rate have drawn increasing attention. Most smart meters installed in practical applications measure and transmit the power signals at a relatively low frequency, generally between 1Hz and 1/900Hz [15]. Consequently, the steady-state signatures become a more suitable choice for NILM and have more reality for application. Low-rate NILM methods can be divided into two categories. One is referred as event-based NILM [16], which implements load monitoring by classifying the signatures related to load events. The other is state-based NILM [17], which realizes load disaggregation through pattern recognition.

Most of state-based NILM methods are based on Hidden Markov Model (HMM) and its variations [18-21] due to the strong ability in modelling the combination of stationary process with continuous valued data over discrete time. Four different extensions of HMM are presented in [20], but they are likely to converge to a local minimum. To address this problem, Hierarchical Dirichlet Process Hidden semi-Markov Mode (HDP-HSMM) is described in [21]. To extend NILM service to new households without further intrusive monitoring, a model fitting algorithm is designed in [22], which adopts iterative k-means to fit a HMM with only one typical duty cycle of device. The main drawback of HMM is that it is heavily dependent on clean transitions from one state to another, while for continuously varying appliances, the performance is poor. To alleviate this problem, a sparse coding method based on structured prediction is developed [23]. Motivated by the success of deep learning, a deep sparse coding is proposed in [24]. However, a typical shortcoming is that for going deeper, more parameters need to be learned. Additionally, state-based algorithms have a common drawback, i.e., long periods of training and high computational complexity, which makes them difficult to be applied to real-time decomposition.

Event-based algorithms have a relatively fixed processing procedure, including event detection, feature extraction and event classification. To obtain accurate identification results, different classification techniques are tried, including k-means [25], k-Nearest neighbour (k-NN) [26], naïve Bayes [27], maximum likelihood [28] and decision tree (DT) [29]. In [28] the maximum likelihood classifier is designed to disaggregate load based on the power profiles, but it only works for single-state loads. Reference [29] relies on graph signal processing (GSP) to perform edge detection, clustering, and pattern matching. However, experimental results show that power fluctuation or close power range of appliances will influence algorithm performance. A novel combined k-means-SVM-based NILM method is developed in [30]. However, event-based methods face a common challenge, that is, most of the existing algorithms only rely on a two-dimensional feature space of active and reactive power for load identification without considering other additional features, such as time and sequence signatures. In addition, the same type appliances in different households have quite different signatures, so it is not suitable to use a unified model to represent them.

The existing NILM methods have been focused on detection of all appliances without considering the applicability of load disaggregation in realistic application, that is, there is no definition of an accurate load space related to the actual application. Since the original load space is too complex, it is impractical to identify all devices based on one-dimension aggregate signal. Consequently, in order to jump out of the dilemma of traditional NILM study, it is necessary to combine it with practical application to define a suitable load space.

In order to address the difficulty of identifying appliances with similar power, a linear discriminant classifier group considering multi-dimensional features is designed in this paper. It is an event-based method, which can work seamlessly with Smart Meter Infrastructure without installing additional acquisition devices. Considering the practical application of this study is to

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 June 2019 doi:10.20944/preprints201906.0169.v1

Peer-reviewed version available at Appl. Sci. 2019, 9, 3558; doi:10.3390/app9173558

provide appliance-level information for DR and energy-saving service, the load space to be monitored can be narrowed down to some controllable and high-power devices.

This work formalizes a load identification technique based on the multi-feature integrated classification (MFIC), where the only input is the time-stamped power readings from the smart meter. The major contributions of this paper are 3-fold, which are listed as follows:

(1) Considering the different operating habits and inherent electrical characteristics of loads, multi-dimensional features are used to model each appliance and improve the load differentiability. In addition, due to the great difference of appliances signatures in different households, this paper uses proprietary model database to replace the uniform feature database.

(2) An event detector based on steady-state segmentation is designed, which has fewer parameters and is independent of the detection window.

(3) Based on the overall similarity of multi-features, a linear discriminant classifier for each appliance is designed, which constitutes a linear discriminant classifier group.

The structure of this paper is given as follows. Section 2 selects multi-dimensional features for load modelling. In addition, a brief analysis of DR and energy-saving services is made to specific the research objective and narrow down the load space. Section 3 elaborates on the problem definition and the complete process of proposed MFIC algorithm. Section 4 presents experiments and their results discussion. The last section concludes the paper and discusses future works.

2. Appliance modelling

2.1 Appliance Behaviour Modelling

The foundation of NILM is to establish an exclusive appliance model library for each user. Therefore, exploiting and defining adequate distinguishable features to model appliance behaviour is an important preparation work for NILM. Since low-frequency electrical signals contain less detailed features, most of the existing low-rate NILM studies rely solely on power metrics to characterize an appliance. However, the power values of some appliances are very close, so the accuracy is low only using power features to identify. Through the analysis of the concrete operation process of each appliance, some distinguishing features can be found. The power consumption of several typical appliances is shown in Figure. 1.

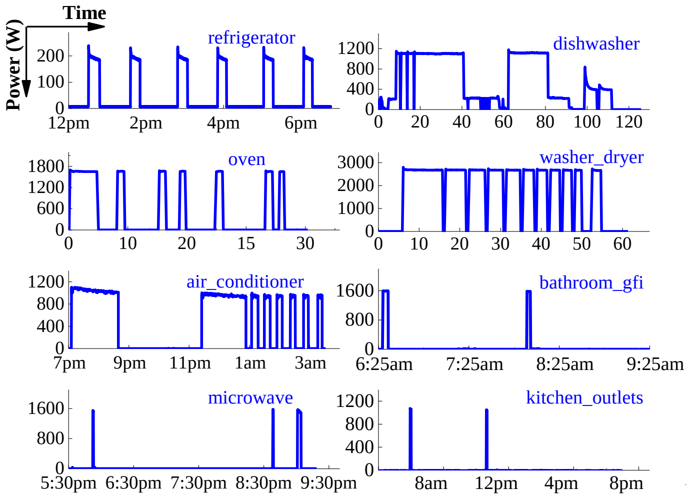

Figure 1. Load profiles of eight typical appliances

Figure.1 presents load profile of the refrigerator between 12pm and 7pm. Refrigerator represents a kind of load with fixed operating period and fixed intermittent time. Without considering the opening of the door, refrigerator operates in a periodic cycle with 15-minutes ON and 51-minutes OFF. A complete operation cycle of the dishwasher can be divided into three main stages: wash, rinse and dry. The operation cycle of dishwasher is determined by its internal

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 June 2019 doi:10.20944/preprints201906.0169.v1

Peer-reviewed version available at Appl. Sci. 2019, 9, 3558; doi:10.3390/app9173558

structure. Oven is also an intermittent running load, but unlike refrigerator, the ON-duration and OFF-duration are not fixed. Generally, the operation time of oven is between 15 minutes and 1 hour, depending on the users' setting. Observing the load profile of washer dryer, it can be found that the first ON time is always the longest, following by several ON/OFF cycles. The number of ON/OFF cycles is related to laundry loads.

In this work, we select the following eight features that can be divided into two categories, i.e., intrinsic features and statistical features. Intrinsic features, also known as electrical features, are generated by the internal structure of appliances. Statistical features reflect users' habits of using specific appliances, which can be called non-electrical features.

Intrinsic features include the following four criteria:

- Active power change. Note that there is an obvious step change in the load profiles given in Figure.1. It refers to the event brought about by the state transition of appliances, and appears as the rising or falling edge. The typical appliances can be divided into three types: single-state, continuous varying and multi-state. The single-state load has a pair of identical rising/falling edges and constant power consumption between them, such as microwave. However, a continuous varying load such as refrigerator generally has a pair of different rising/falling edges and the power consumption during operation is continuously changing. Besides, multi-state load has more than one working stage, such as dishwasher.
- On-duration. It refers to the continuous operating time in a periodic cycle. It is mainly applied to the loads with fixed operating periods, such as refrigerator. As abovementioned, the On-duration of the refrigerator is 15 minutes.
- Off-duration. Similar to the On-duration, Off-duration stands for the continuous standby time in a periodic cycle. For instance, the On-duration of refrigerator is 51 minutes.
- On times. It means the number of turning on contained in each operation cycle.

Statistical features serve to emphasize the user's operating habits for different appliances, including the following four kinds.

- Switching-time. It refers to the possible switching on/off time, which is related to the function of an appliance.
- Usage frequency in one day. It is the possible usage count for a day. Notice that there are many appliances operate in an ON/OFF cycle such as oven, air conditioner and washer dryer. When calculating the usage frequency of these loads in a day, it is necessary to ensure that a complete operation process is counted once.
- Working days in a week. It means the number of days that an appliance may work within a week. A refrigerator is a constant-opening device. But the washing machine is less likely to be used every day.
- Duration of a complete use process. This feature is used to record the duration of an appliance from start to shut down, including all subsequences ON/OFF cycles. It should be noted that if a load has the habit of operating multiple times in a short time, but their final task is identical, so they can be summed up as a complete operation. For instance, the users may use the stove several times during cooking.

2.2 Determine the load space to be monitored

As described in the introduction, it is impractical to identify all types of appliances using only one NILM approach, so it is essential to build a reasonable load space according to the specific application of NILM.

The opportunities for various appliances to participate in DR are different. They are usually divided into three categories: uncontrollable load (UCL), transferable load (TL) and interruptible load (IL). UCL refers to the load that has no energy storage capacity and may be opened at any time. Its power fluctuation range is small and basically has no capacity to transfer load. However, TL has the ability of load transfer, because its using time is more flexible and the total energy consumption is certain, such as the washing machine and the timed electric rice cooker. Consequently, under the premise of ensuring the completion of the work requirements, it can participate in DR by changing

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 June 2019 doi:10.20944/preprints201906.0169.v1

Peer-reviewed version available at Appl. Sci. 2019, 9, 3558; doi:10.3390/app9173558

the running time. Air conditioner and water heater are typical IL, which can be interrupted at certain time to reduce the power consumption. Therefore, these loads can usually be interrupted temporarily without resulting in consumer discomfort. In summary, TL and IL can participate in DR, and can be monitored and managed centrally at the control center. On the contrary, the interruption of UCL is likely to affect user's comfort, so UCL is allowed to be centrally controlled only when emergency occurs.

Above analysis specifies the load space that NILM should focus on when providing services for DR. The main research target is to identify the high-power and adjustable loads, while there is no need to track the appliances with small power. The load space that needs to be monitored is narrowed down, only including some controllable loads (IL and TL) or high energy consumption loads without adjustable ability. The final load space to be monitored is reported in Table 1.

Table 1. Load space to be monitored.

Types	Loads
IL or TL	HVAC (Heating, Ventilation and Air Conditioner), Electric Heat loads (water heater, furnace, oven), electric vehicle charger, washer and dryer, dishwasher, refrigerator
UCL	High-power devices

3. Methodology

3.1 Load Disaggregation Definition

The definition of load decomposition can be expressed as follows: Given the mixed signal collected at the entry point of a house and typical appliance models in it, we need to breakdown the mixed signal into a set of individual components that are attributed to specific appliances.The composite signal is clearly depending on which appliances are switched ON at the given moment, so it is necessary to design a Boolean coefficient $a_{n,m}(k)$, which determines whether the mode m of appliance n is ON at the kth sampling point. Mathematically, the observed mixed signal can be formulated as a linear combination of some unknown appliance load profiles.

$$P(k)=\sum_{n=1}^{N}\sum_{m=1}^{M}a_{n,m}(k)p_{n,m}(k)+e(k) \tag{1}$$

where $P(k)$, k=1,2,3,…,L is the aggregated power signal (L is the number of samples) and $p_{n,m}(k)$ denotes the individual power consumption of appliance n in mode m. N and M are the number of appliances and modes, respectively. $e(k)$ stands for the noise signal and small appliances, including phone chargers, DVD player and laptop computers. In load identification, $e(k)$ is ignored. In other words, the objective in NILM is to decode $P(k)$ and obtain the status of each appliances using a set of appliance models in the house.

In existing research, combinatorial optimization is a common method to solve (1), which finds the optimum appliance status by minimizing the difference between the actual aggregate power and the sum of disaggregated appliance powers, subject to some prior information.

$$a_{n,m}(k)=\underset{a_{n,m}(k)}{\arg\min}\left|P(k)-\sum_{n=1}^{N}\sum_{m=1}^{M}a_{n,m}(k)p_{n,m}(k)\right| \tag{2}$$

3.2. Algorithm overview

There will be many uncertainties in reality, because different appliances may perform similar electrical signature "power". Appliances may not operate at its rated power, because the actual power consumption is proportional to the total load. Therefore, it is difficult to solve the problem in (2) by combinatorial optimization.

The actual electric data displays that it is easy to segment the total signal into some steady-state process by clearly step changes. Therefore, an event-based algorithm is designed in this paper to solve the load disaggregation problem. This method consists of three steps: 1) event

 doi:10.20944/preprints201906.0169.v1

Peer-reviewed version available at Appl. Sci. 2019, 9, 3558; doi:10.3390/app9173558

detection and clustering, 2) event paring and key electrical feature extraction and 3) feature matching. In the first step, we make statistics on the significant changes in active power, which represent that some appliances have changed their status. Then, events with similar power should be grouped, i.e., clustering. After the formation of clusters, events in "positive" clusters should be paired with those in negative clusters. Finally, extract the key features from each positive-negative cluster pair, and match with the appliance models. The flowchart is illustrated in Figure.2.

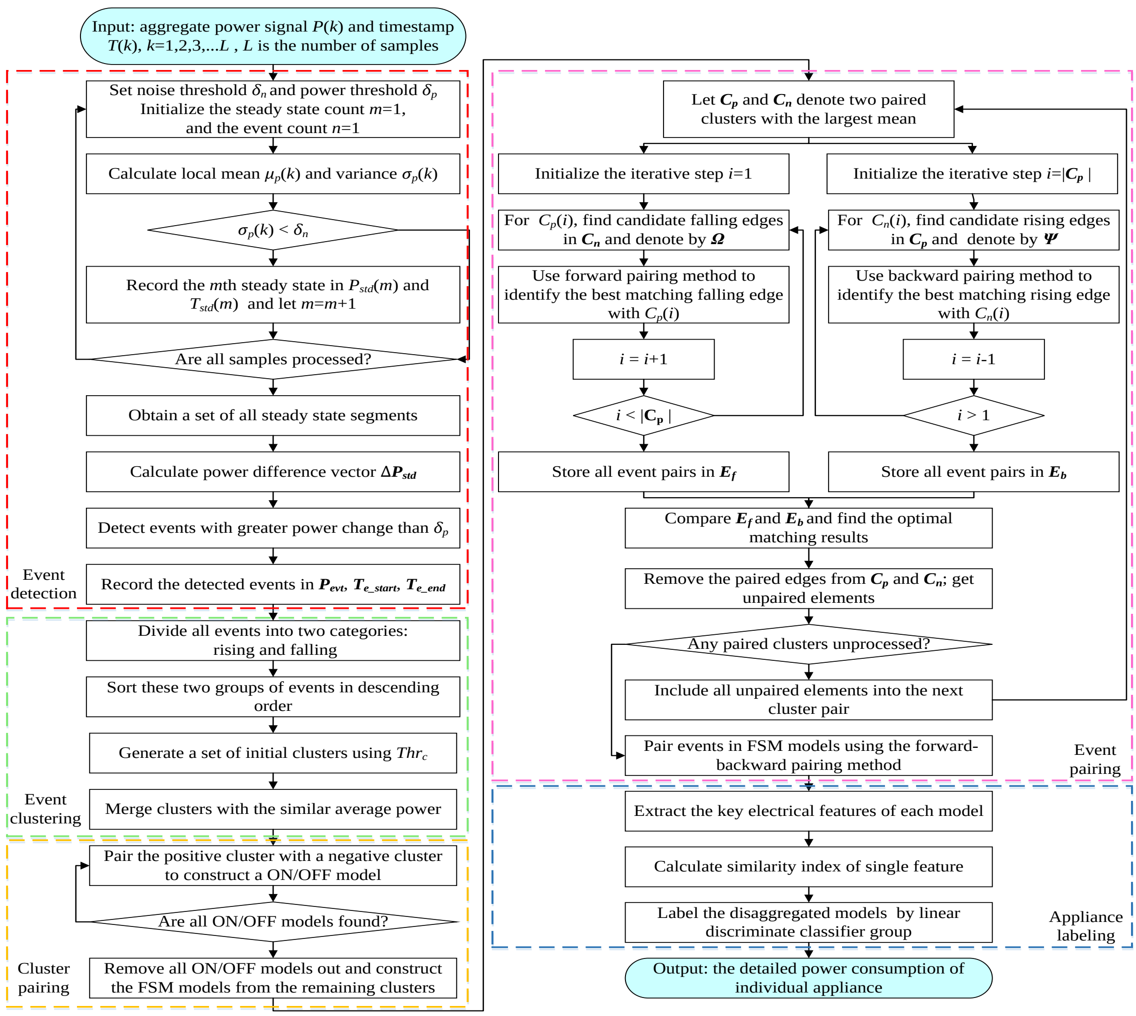

Figure 2. Flowchart of the proposed algorithm

3.3. Steady-state Segment Based Event Detection

One important characteristic of event-based load disaggregation is to detect the significant rising or falling edge in active power, and record the power value and occurrence time of events. A steady-state segment based event detection method is presented. It has two parameters, one is the noise threshold d_n, the other is the power threshold d_p. The schematic diagram of the event detection is indicated in Figure.3.

Each appliance can be represented by two states: 1) steady state, including both ON state and OFF state; and 2) transition state, i.e., the process of switching ON/OFF or changing operation state of multi-state appliance. As long as the steady-state segments have been identified, the duration of the power-on or power-off processes for different appliances can be determined adaptively. Based on the switch continuity principle proposed in [3], we can suppose that load switching is successive, i.e., only one state transition can occur within the sampling time interval. Thus, it is feasible to use the step power variation in aggregated signal as the discriminant feature of event occurrence.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 June 2019 doi:10.20944/preprints201906.0169.v1

Peer-reviewed version available at *Appl. Sci.* **2019**, *9*, 3558; doi:10.3390/app9173558

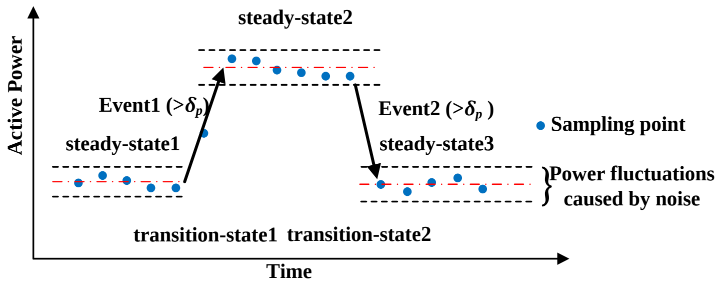

Figure 3. Schematic diagram of the event detection

On the one hand, the power grid noise exists always in the actual electric environment. On the other hand, the time of state transition can be longer than the sampling interval. Consequently, it is necessary to design an appropriate event extraction method, which is robust to the possible variation of power amplitude due to sampling and noise. We adopt the method of calculating local mean and variance to find steady state. Assuming $P(k)$, k=1,2,3,…,L is a given aggregate power signal and $T(k)$, k=1,2,3,…,L is the corresponding timestamp, two quantities need to be calculated by equation(3) and (4), the former is the local average power and the latter is the local variance.

$$\mu_P(k) = \tfrac{1}{3}\sum_{i=-1}^{1} P(k+i) \tag{3}$$

$$\sigma_P(k) = \tfrac{1}{3}\sum_{i=-1}^{1} (P(k+i) - \mu_P(k))^2 \tag{4}$$

Let δ_n^2 denote the noise variation in power grid. If $\sigma_p(k)<\delta_n$, $P(k)$ is considered to be in a steady state and then two new variables $P_{std}(m)$ and $T_{std}(m)$ are added to record the mth steady state, i.e.,

$$P_{std}(m) = \mu_P(k) \tag{5}$$

$$T_{std}(m) = k \tag{6}$$

After all the steady state segments have been identified, the power difference between two consecutive steady states will be calculated as follows:

$$P_{std}(m) = P_{std}(m) - P_{std}(m-1) \tag{7}$$

Given the power threshold δ_p, it is set by users and denotes the events they are interested. For example, if users want to find events with power change greater than 100 W, they will set δ_p=100. If abs($\triangle P_{std}(m)$)> δ_p, it indicates that a new event is detected and the power value and timestamp of this event will be registered as

$$P_{evt}(n) = \Delta P_{std}(m) \tag{8}$$

$$T_{e\_start}(n) = T(T_{std}(m-1)) \tag{9}$$

$$T_{e\_end}(n) = T(T_{std}(m)) \tag{10}$$

where n represents the nth event. $P_{evt}(n)$ stands for the power value of the nth event, $T_{e\_start}(n)$ and $T_{e\_end}(n)$ stands for the start time and end time of the event respectively.

3.4. Event Clustering

The collection of the registered events $P_{evt}(n)$, n=1,2,…,N_e is the basis of the event clustering. N_e denotes the number of events detected. Since we assume that each state of appliances has a unique value and only one appliance may have state transition in one sampling interval, it is reasonable to gather events with similar value into one cluster. Ideally, each cluster represents one kind of state transition of appliance.

Peer-reviewed version available at Appl. Sci. 2019, 9, 3558; doi:10.3390/app9173558

In this paper we design a clustering algorithm without prior knowledge, which can adaptively determine the number of clusters. There are two steps in this algorithm.

Step 1) Separate the rising and falling edges of the event candidates into two collections of $\boldsymbol{P_{evt\_up}}$ and $\boldsymbol{P_{evt\_down}}$. Then the rising and falling edges are arranged in descending order according to the absolute value of power respectively. Set a cluster threshold Thr_c. When the difference between two consecutive rising or falling edges is greater than Thr_c, a new cluster is generated. Here we set Thr_c smaller so that the clustering results will be more detailed. However, it is also easy to separate some events with power fluctuation but belonging to one appliance into different clusters. In order to solve this problem, we merge some clusters with the similar average power. The detailed process is illustrated in step 2.

Step 2) Calculate the mean power of each cluster, and gain the mean power difference between two adjacent clusters. If the difference is less than a certain value, it can be considered that these two adjacent clusters belong to the same appliance state. Thus, we will merge them together and the new cluster candidates will be formed.

3.5. Building Appliance Candidate Model

After clustering the events, we get some "positive" clusters containing rising edges and "negative" clusters composed of falling edges. Next, the pairing method is designed to automatically generate appliance candidate models.

Most of the existing NILM algorithms only consider the single-state appliances, but in practice, the multi-state appliances are very common. These appliances cannot be described by ON/OFF model, so it is necessary to establish an appropriate model for them. The finite state machine (FSM) [1] is a typical model for these appliances. The sum of power changes in any cycle of state transition is zero, which can be called Zero Loop-Sum Constraint (ZLSC) [1]. Besides, different operating states in an FSM model have different power levels, i.e., Uniqueness Constraint (UC). The two constraints make it possible to construct individual FSM from streams of events.

In the following, the method of generating ON/OFF or FSM models is introduced. It includes two main steps, i.e., cluster pairing and event pairing.

Step 1) To construct ON/OFF models for the single-state appliance candidates, all we need to do is pairing the "positive" cluster and "negative" cluster with similar absolute average power. However, the construction of FSM models is somewhat complex, which should take advantage of special algebraic properties of events in a complete transition cycle, i.e., ZLSC and UC. In order to reduce the complexity of cluster pairing, the ON/OFF models are built firstly. After all positive-negative cluster pairs have been found, we remove them from the total clusters and search for FSM models from the remaining clusters.

Step 2) Following the cluster pairing, some cluster pair candidates for single-state or finite-state appliances will be generated. Then, it is essential to further match the events in each cluster pair. That is to say, looking for the complete state transition processes of each appliance candidate. For example, we need to match each rising edge in the positive cluster with a falling edge in the paired negative cluster, exploiting power difference and time intervals between two events as pairing features. Next we will focus on the event pairing process of ON/OFF models, which is also applicable to the FSM model. In order to increase pairing accuracy, a specialized forward-backward pairing procedure is designed.

Let $\boldsymbol{C_p}$ and $\boldsymbol{C_n}$ denote two paired clusters, where $|\boldsymbol{C_p}|$ and $|\boldsymbol{C_n}|$ denotes their cardinality.

(a) Forward Pairing

The goal of forward pairing is, for each $C_p(i) \in \boldsymbol{C_p}$, match an optimal falling edge among all elements in Cn according to the order from i=1 to $|\boldsymbol{C_p}|$-1. Normally, the ON and OFF events appear alternately, that is, using time stamps to sort events that belong to the same appliance will get an ON/OFF/ON/OFF… sequence. Thus, the falling edge paired with $C_p(i)$ must occur after $C_p(i)$ and before $C_p(i+1)$. The subset of $\boldsymbol{C_n}$ that satisfy the above condition is considered as a set of candidates, denoted by $\boldsymbol{\Omega}$. Let $|\boldsymbol{\Omega}|$ represent the element number of $\boldsymbol{\Omega}$. The values of $|\boldsymbol{\Omega}|$ can be divided into three cases. Different pairing processes are designed for these cases. Define two vectors $\boldsymbol{M_p}$ and $\boldsymbol{M_t}$ to record the power difference and time intervals between paired events.

Case 1: When $|\boldsymbol{\Omega}|=1$, the absolute power difference and time interval between $C_p(i)$ and the only element in $\boldsymbol{\Omega}$ are calculated and denoted by Ω_p and Ω_t respectively. Then the probability of pairing $C_p(i)$ and the element in $\boldsymbol{\Omega}$ can be defined as:

$$c_i = \frac{\Omega_p m_p + \Omega_t m_t}{\sqrt{(\Omega_p{}^2 + \Omega_t{}^2)(m_p{}^2 + m_t{}^2)}} \tag{11}$$

Where, m_p stands for the mean value of the elements in $\boldsymbol{M_p}$ and m_t denotes the median value of the elements in $\boldsymbol{M_t}$.

If c_i is larger than a given threshold, the only element in $\boldsymbol{\Omega}$ can be considered as the paired falling edge for $C_p(i)$, otherwise they are not match. Then record Ω_p and Ω_t between paired events in vector $\boldsymbol{M_p}$ and $\boldsymbol{M_t}$.

Case 2: When $|\boldsymbol{\Omega}|>1$, the absolute power differences and time intervals between $C_p(i)$ and each candidate in $\boldsymbol{\Omega}$ are calculated and denoted by Ω_p and Ω_t respectively. Then the probability of pairing $C_p(i)$ and the jth candidate in $\boldsymbol{\Omega}$ can be defined as:

$$c_i(j) = \frac{\Omega_p(j) m_p + \Omega_t(j) m_t}{\sqrt{(\Omega_p(j)^2 + \Omega_t(j)^2)(m_p{}^2 + m_t{}^2)}} \tag{12}$$

Find the falling edge $\boldsymbol{C_n}$ corresponding to the maximum value in vector ci. If the maximum value is larger than the given threshold, $\boldsymbol{C_n}$ can be judged to the paired falling edge for $C_p(i)$. Then record the power difference and time intervals between paired events by vector $\boldsymbol{M_p}$ and $\boldsymbol{M_t}$.

Case 3: When $|\boldsymbol{\Omega}|=0$, we think there is no appropriate element in $\boldsymbol{C_n}$ pairing with $C_p(i)$. However, it is not applicable to a special situation, i.e., $C_p(i+1)$ is wrongly cluster. In this case, the accuracy of the results obtained by forward pairing is lower, so backward pairing is proposed.

When the forward pairing is complete, all event pairs are stored in matrix $\boldsymbol{E_f}$.

(b) Backward Pairing

The goal of backward pairing is, for each $C_n(i) \in \boldsymbol{C_n}$, match an optimal rising edge among all elements of $\boldsymbol{C_p}$ according to the order from $i=|\boldsymbol{C_n}|$ to 2. According to the analysis in forward pairing, the rising edge paired with $C_n(i)$ must occur before $C_n(i)$ and after $C_n(i\text{-}1)$. The subset of $\boldsymbol{C_p}$ that satisfies the above condition is considered as a set of candidates, denoted by $\boldsymbol{\Psi}$. Let $|\boldsymbol{\Psi}|$ represent the element number of $\boldsymbol{\Psi}$. The specific realization process is basically the same with the former pairing. However, it is worth noting, that when $|\boldsymbol{\Psi}|=0$ the accuracy of the results obtained by backward pairing is lower, which needs to be analysed with forward pairing results. When the forward pairing is complete, all event pairs are stored in matrix $\boldsymbol{E_b}$.

Finally, comparing $\boldsymbol{E_f}$ and $\boldsymbol{E_b}$, we can find the optimal matching results. The event pairs that appear in both $\boldsymbol{E_f}$ and $\boldsymbol{E_b}$ are thought to be matched correctly. If there are some event pairs in $\boldsymbol{E_f}$ and $\boldsymbol{E_b}$ that have the same rising edges but the corresponding falling edges are different, then the pairing results in $\boldsymbol{E_f}$ are considered to be optimal. The essence of this situation is that there are multiple falling edges between two successive rising edges, leading to inaccurate results of backward pairing. Moreover, if there are some event pairs in $\boldsymbol{E_f}$ and $\boldsymbol{E_b}$ that have the same falling edges but the corresponding rising edges are different, then the pairing results in $\boldsymbol{E_b}$ are considered to be optimal. The essence of this situation is that there are multiple rising edges between two successive falling edges, leading to inaccurate results of forward pairing.

3.6. Appliance Identification Based on Multi-feature Integrated Classification

With the aforementioned process, the raw smart meter recordings are decomposed to a set of appliance candidate models and each model carries unique information that corresponds to an appliance footprint. Next, the key features are extracted to label each candidate model combined with an existing feature library for the particular house.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 June 2019 doi:10.20944/preprints201906.0169.v1

Peer-reviewed version available at Appl. Sci. 2019, 9, 3558; doi:10.3390/app9173558

3.6.1. Similarity Index of Single Feature

Intrinsic features are determined by the internal structure of appliances, which are not affected by the user's behaviour habits, and usually expressed as fixed values without some small fluctuations caused by the grid noise. The similarity indices of these features can be quantified as

$$S_{(\cdot)}(v) = \exp\left(-k\frac{|v - v_{mean}|}{H(v - v_{mean})}\right) \tag{13}$$

$$H(\Delta v) = \begin{cases} |v_{\max} - v_{\text{mean}}|, & \Delta v \geq 0 \\ |v_{\min} - v_{\text{mean}}|, & \Delta v < 0 \end{cases} \tag{14}$$

where $(\cdot)$ denotes the intrinsic feature of detection. v represents the detected value of certain feature and v_{mean} is the average value of certain feature recorded in the feature library. Considering the noise in appliances or power grid, there are some small fluctuations in these intrinsic features, so v_{max} and v_{min} are used to represent the limits of up and down fluctuation respectively. $H(\cdot)$ is a piecewise function. k is a calibration parameter to ensure that when the detected value v exceeds v_{max} and v_{min}, its similarity index is almost 0. In this paper, k=1 is used.

Statistical features are expressed as a range rather than a fixed value. The similarity calculation of statistical features is defined as

$$S_{(\cdot)}(x) = \begin{cases} 1, x \in R_{(\cdot)} \\ 0, x \notin R_{(\cdot)} \end{cases} \tag{15}$$

where $(\cdot)$ denotes the statistical feature of detection. x is the statistical value of specific feature. $R(\bullet)$ stands for the range of possible values for a certain feature.

3.6.2. Appliance Recognition Based on Linear Discriminant Classifier Group

The task of this section is to label each appliance candidate model based on similarity indices. In order to synthetically consider the effects of various features on appliance identification, a linear discriminant classifier is designed for each appliance based on the overall similarity and all of them constitute a linear discriminant classifier group. Overall similarity is calculated by weighted composition. The features of marking different appliances are inconsistent. Therefore, the particular weight vector needs to be set for each linear discriminant classifier separately. It is firstly estimated by observing the identifiability of different features. For instance, knowing refrigerator has specific ON-duration and OFF-duration, these two features will be emphasized. But they are insignificant in lights identification, which can range from a few minutes to a whole day. Generally, the intrinsic features are important than statistical features since statistical features are easily influenced by external environment. For multi-state appliances, the sequence feature is particularly important. After the weight vectors are predefined, it is necessary to adjust their values exploiting the test data in different time and environment, so as to ensure recognition accuracy.

The detailed process of labelling appliance candidate models is described below. At first, the intrinsic and statistical features of each model are extracted. Then each unlabeled model will be classified by the linear discriminant classifier group in this particular house. The classification result of the jth classifier is shown as

$$d(\mathbf{S}_j) = \boldsymbol{\omega}_j^T \mathbf{S}_j - \delta_j \tag{16}$$

where ω_j is the weight vector of the jth classifier, $\mathbf{S}_j$ includes feature similarity indices between the unlabelled model and the jth classifier, and δ_j represents the judgment threshold of the jth classifier. If d(S_j)>=0, the unlabelled model is determined as the appliance corresponding to the jth classifier; Otherwise not.

4. Experiment and result analysis

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 June 2019 doi:10.20944/preprints201906.0169.v1

Peer-reviewed version available at Appl. Sci. 2019, 9, 3558; doi:10.3390/app9173558

In order to verify the effectiveness of the proposed algorithm, simulation experiments are carried out using the low-frequency database: REDD [31]. It is one of commonly used database in the field of NILM, so testing our method with it allows better comparison with other approaches.

4.1. Performance Metrics

In order to evaluate the performance of separated signals and compare existing implemented algorithms, some indices are needed to calculate the performance quantitatively. Since an event-based NILM algorithm is designed in this paper, it is essential to measure the accuracy of this method in predicting what appliance is running in each state. Classification accuracy indices, such as precision, recall, and F-measure, are well suitable for this task. Precision denotes the positive predictive values, i.e., the correct proportion of samples identified as appliance c. Recall represents the true positive rate, i.e., the proportion of samples belonging to appliance c that are recognized correctly. F-measure is harmonic mean of precision and recall. These typical classification metrics can be formulated as follows:

$$P_c = TP_c / (TP_c + FP_c) \tag{17}$$

$$R_c = TP_c / (TP_c + FN_c) \tag{18}$$

$$F_c = 2 \times (P_c \times R_c) / (P_c + R_c) \tag{19}$$

where, the subscript c is used to mark different appliances or states. TP_c indicates true positive, i.e., the correct judgment that appliance c was ON, FP_c represents the false positive, i.e., judged appliance c was ON but actually was OFF, FN_c denotes false negative, that is, appliance c was ON but was wrongly judged as OFF. Note that these indices (TP_c, FP_c, FN_c) are accumulations over a certain experimental time period.

On the other hand, it is important to feedback the detailed power consumption of each appliance to users, so the accuracy based on power estimated also needs to be considered. The frequently used measures to compare the estimated power with the actual power consumption include disaggregation accuracy (*DA*), disaggregation error (*DE*) and percentage of contribution in energy consumption (*PCEC*). The first two metrics provide a global comparison between the estimated power and the ground truth, while *PCEC* is used to calculate the contribution of each appliance in total power consumption. In this paper, *DA* and *PCEC* are exploited to evaluate the ability of different algorithms for reconstructing power profiles.

$$DA = 1 - \frac{\sum_{k=1}^{L} \sum_{n=1}^{N} |\hat{p}_n(k) - p_n(k)|}{2\sum_{k=1}^{L} P(k)} \tag{20}$$

where L is the number of disaggregated readings, N denotes the number of appliances in the house, $\hat{p}_n(k)$ represents the estimated power consumption of appliance n at the kth sample, $p_n(k)$ is the actual power consumed at the kth sample for appliance n, and $P(k)$ stands for the aggregate power measured at the kth sample.

$$PCEC^n = \sum_{k=1}^{L} \hat{p}_n(k) / \sum_{k=1}^{L} P(k) \tag{21}$$

where $PCEC^n$ represents the contribution of appliance n to total power consumption.

Furthermore, energy usage profile pie graph and comparison curve between disaggregated signals and their ground truth are other means to verify the performance of different algorithms.

4.2. Case Study 1

Several appliances are selected from House 2, namely refrigerator, microwave and dishwasher. Refrigerator and microwave are used frequently and have high power consumption in this house. Dishwasher is a typical multi-state load with adjustable potential. Since the recognition process

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 June 2019 doi:10.20944/preprints201906.0169.v1

Peer-reviewed version available at *Appl. Sci.* **2019**, *9*, 3558; doi:10.3390/app9173558

depends on some statistical features, the aggregate power signals in one week are selected for the experiment. Figure.4 shows the results from aggregate signal to event detection and clustering.

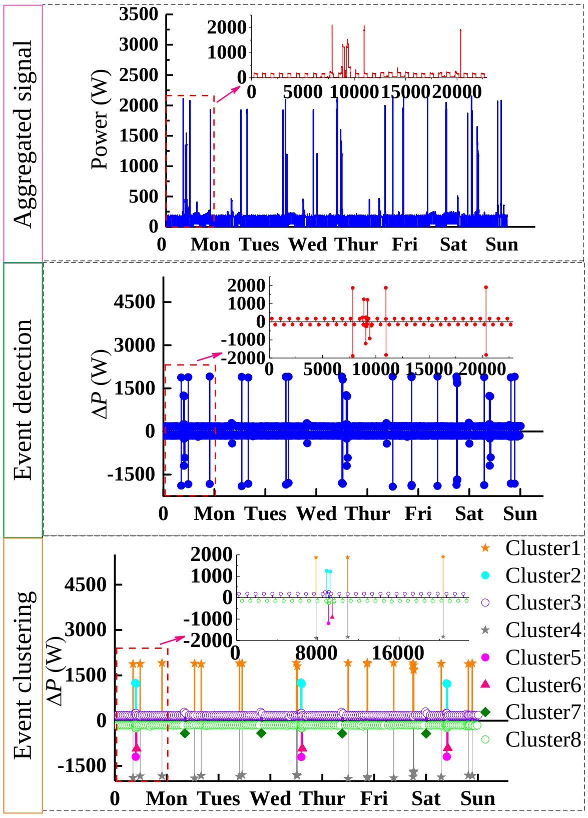

Figure 4. Results of event detection and clustering for House 2

Figure.4 illustrates the results of event detection and clustering in one week. In order to show the process more clearly, the results in Monday are expanded and reported in Figure.4. Using the steady-state segment based event detection method, the significant changes from aggregate power signal can be detected accurately. After processing all events with the proposed clustering method, eight different clusters are formed, including three "positive" clusters and five "negative" clusters. It can be seen that the elements in each cluster have similar power value.

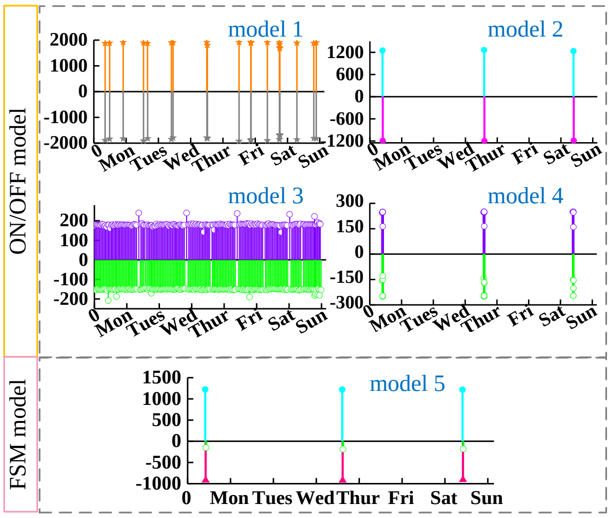

Figure 5. Results of event pairing for House 2

Next, each positive cluster is matched to the magnitude-wise closest negative cluster with the cluster pairing method. Then for each cluster pair, the forward-backward pairing approach is adopted to find all matching events and reject the unmatched events. Repeat the above process

Peer-reviewed version available at Appl. Sci. 2019, 9, 3558; doi:10.3390/app9173558

through the remaining unmatched events, until no events can be matched. After each loop, all positive-negative cluster pairs that have completed the matching process will be stored as the ON/OFF models, passed to the following appliance identification step. When all ON/OFF models have been found, we remove them from the set of events and attempt to establish FSM models from the remaining events. Figure.5 presents the results of building appliance candidate models. Four ON/OFF models and one FSM model are established, and their corresponding time profile features are shown in Figure.5. It should be emphasized that although the power values of model 3 and model 4 are close, they can still be separated due to the large difference in the ON duration.

Finally, the linear discriminant classifier group of house2 is used to label each candidate model and the results of each classifier are reported in Table 2. It can be seen that the classifier group of house 2 consists of seven appliances, namely kitchen outlets, stove, microwave, high-power state of dishwasher, low-power state of dishwasher, multi-state of dishwasher and refrigerator. They are typical loads with high power consumption or adjustable potential in house2, which are interested by users and registered in feature database. Note that different modes of dishwasher with different operation cycles are treated as separated appliances. Therefore, in the final labelling stage, the models that are identified as appliance 4, 5 and 6 are all labelled as a dishwasher. From the signs of classifier value, model 1 is determined as appliance 3 while model 2 is appliance 4 since their corresponding values are greater than 0. In addition, model 3 and 4 represent appliance 7 and 5, respectively. The disaggregated FSM model 5 corresponds to appliance 6. If there is no positive value for a model, it means that the model is caused by an unregistered appliance in the feature database, which may be a new or low power consumption appliance not interested by users.

Table 2 Identification results of separated candidate models.

Separated model	Appliance Classifiers						
	App1	App2	App3	App4	App5	App6	App7
Mod 1	-0.85	-0.79	0.04	-0.85	-0.70	0	-0.79
Mod 2	-0.62	-0.74	-0.38	0.06	-0.56	0	-0.70
Mod 3	-0.47	-0.77	-0.73	-0.77	-0.65	0	0.03
Mod 4	-0.34	-0.55	-0.45	-0.60	0.10	0	-0.58
Mod 5	—	—	—	—	—	0.02	—

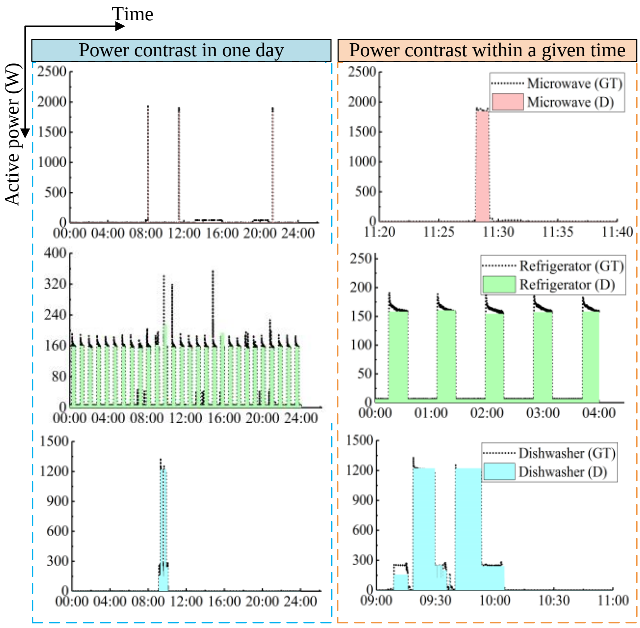

Figure 6. Ground truth (GT) and estimated appliances' power consumption (D) for House 2

For a more detailed comparison between the disaggregated models and corresponding appliances, the power consumption of each model is reconstructed. The essence is to translate a predicted label into predictive power consumption. The real and reconstructed power profiles of each appliance are illustrated in Figure.6. The three figures on the left shows the power comparison in one day. It can be seen that the algorithm can accurately identify the operation state of each appliance. The power profiles within an adequate time interval is displayed on the right side. For the single-state load such as microwave, the power signal can be estimated quite accurately.

To comprehensively explore the suitability of designed algorithm in solving the signal reconstruction problem and estimate the percentage of contribution of each appliance in the whole home energy consumption, the *PCEC* values are shown in a schematic pie plot. The idea is to compare how closely the energy contribution of an appliance estimated by our algorithm matches with the ground truth. The results of House 1, 2, 6 during one week are present in Figure.7. Figure.7 illustrates that some appliances included in the feature database are OFF during the whole period and the proposed algorithm can detect this pattern accurately. Likewise, there is not any miss identification of an appliance being OFF shown in the disaggregation result. The *PCEC* values estimated by our method are closer to the ground truth, which further confirms the ability of the proposed algorithm in signal reconstruction. For the three houses, the average absolute differences between the results and the actually measured values are 3.97%, 0.30% and 0.73%, respectively.

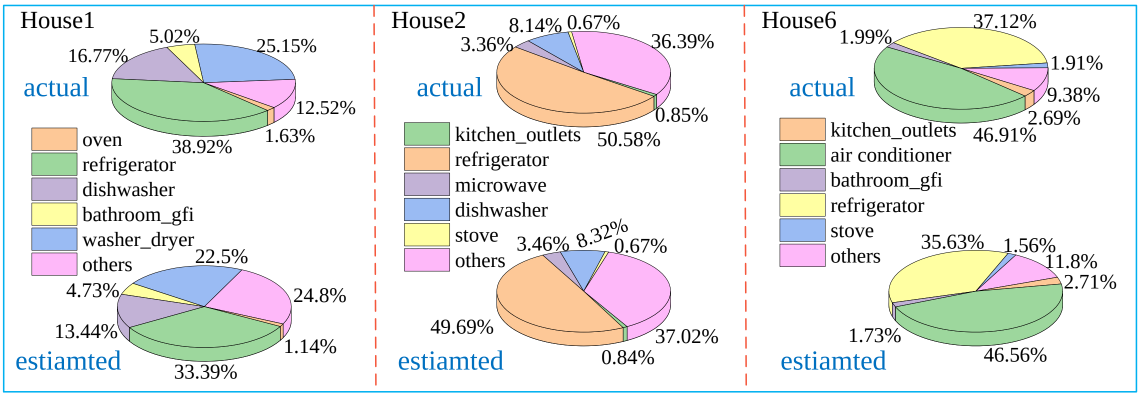

Figure 7. Comparison of actual and estimated PCEC values for House 1, 2, 6

4.3. Case Study 2

In this section, the main purpose is to compare our results with some existing classification based NILM algorithms. Precision, recall and F-measure are selected as the base measure.

In order to verify the classification accuracy, the results of the proposed algorithm and unsupervised GSP-based approach [29] on house 1, 2 and 6 are represented in Figure.8. It demonstrates that the proposed method provides more accurate classification than GSP-based method for all houses.

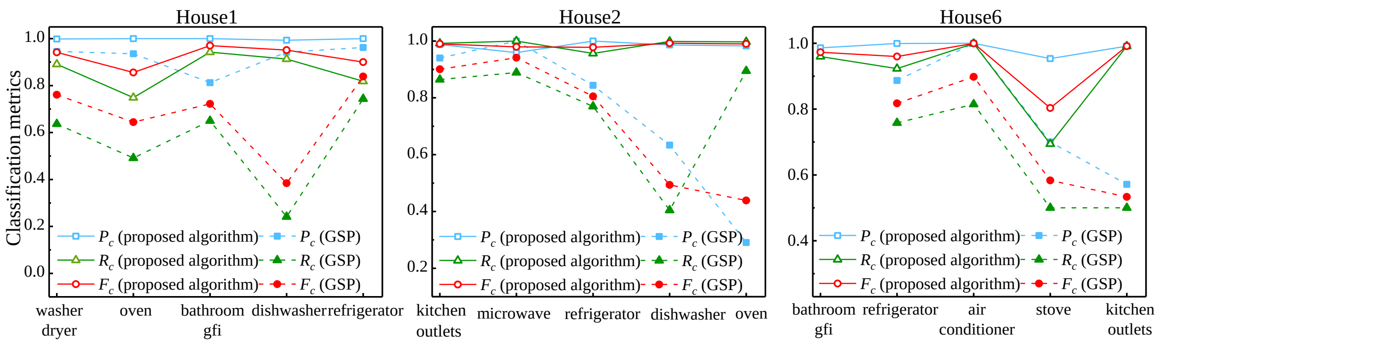

Figure 8. Classification performance comparison with GSP-based method for House 1, 2, 6

Furthmore, the performance of proposed MFIC algorithm is compared with the state-of-the-art NILM approaches used for low sampling rate and power signals. The MFIC results, F_U, are compared with those of the combined k-means/SVM classification [30], F_S, the HMM- based method, F_H, and the Decision Tree (DT) approach, F_{DT}, as reported in [17]. The results are shown in Table 3.

Peer-reviewed version available at Appl. Sci. 2019, 9, 3558; doi:10.3390/app9173558

Table 3 Comparison of four low-rate NILM algorithms using REDD database.

Appliance	F_S	F_H	F_{DT}	F_U
Air conditioner	—	0.12	0.89	99.9
Washer dryer	75.36	0	0.88	94.2
Dishwasher	35.97	0.04	0.32	99.2
Oven	79.13	—	—	85.6
refrigerator	94.35	0.90	0.97	97.8
Microwave	25.91	0.47	0.97	97.9
Stove	44.4	0.21	0.33	98.9
House	F_S	F_H	F_{DT}	F_U
House 1	77.52	77.06	78.09	92.39
House 2	82.17	82.38	81.41	98.58
House 6	95.58	72.76	75.94	94.53

On the one hand, we study the variation of performance with respect to different appliances, mainly including some controllable or high-power loads in REDD database. It can be seen that the MFIC algorithm achieves the best disaggregation in terms of F-measure for all appliances. HMM yields significantly worse results, but it usually performs well identifying refrigerator because the continuity and singleness (i.e., no other devices operates, especially at night) of its operation bring sufficient data for training. The results of k-means/SVM and DT are relatively good but worse than those of MFIC. On the other hand, the average results for three REDD houses are compared. Note that the results of combined k-means/SVM classification and HMM are shown in [30], and the results of DT are reported in [32]. It can be seen that the proposed MFIC algorithm has consistently high performance across all three houses and outperforms other algorithms in both House 1 and House 2. The combined k-means/SVM method shows a higher accuracy for House 6.

Moreover, we also use the disaggregation accuracy metric to compare the performance of MFIC algorithm with the Bayesian HMM based technique [21], segmented integer quadratic constraint programming (SIQCP) based algorithm [22], sparse coding (SC), and discriminating sparse coding (discSC) [24].The comparison results are given in Table 4. Obviously, the MFIC algorithm improves significantly compared with the SC-based method from previous work and is slightly superior to the Bayesian HSMM model and SIQCP solver, with the disaggregation accuracy increased by 11.1%, 7.1% respectively.

Table 4 Disaggregation accuracy comparison with other methods.

Algorithm	Disaggregation accuracy (*DA*)			
	House1	House2	House6	Mean
MFIC	90.3%	92.0%	95.5%	92.6%
Bayesian HSMM	82.1%	84.8%	77.7%	81.5%
SIQCP	78.4%	86.4%	91.6%	85.5%
SC	57.2%	65.4%	58.1	60.2%
discSC	58.1%	68.3%	53.9%	60.1%
DSC(Greedy)	60.8%	71.1%	61.7%	64.5%
DSC(Exact)	64.3%	74.9%	64.2%	67.8%

5. Conclusion

In this paper, a MFIC load disaggregation technique is presented, where the only input is the time-stamped power readings from the smart meter. In order to meet the load monitoring requirements of demand response and energy conservation, the load space to be monitored is narrowed down to some controllable and high-power consumption loads. The systems uses steady-state segmentation based method to detect events and combines multi-dimensional features both electrical and non-electrical features to improve the accuracy of load identification. As the experiment results demonstrated, the linear discriminant classifier group gives excellent classification performance and correctly identifies the devices present in the load space, establishing the applicability of the MFIC algorithm. Compared with existing NILM approaches, the separation accuracy of the MFIC algorithm is significantly improved due to the reduced load space. Meanwhile, for some loads with similar power, our algorithm can still correctly separate them from aggregate signals.

Author Contributions: C.S. and J.Y. conceived and designed the experiments; J.Y. and Y.W. performed the simulation and analyzed the results; and J.Y and Y.W. wrote the paper.

Funding: This research was funded by [the Fundamental Research Funds for the Central Universities of China] grant number [2017MS006].

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

References

1. Jung, S.; Yoon, Y.T., "Optimal Operating Schedule for Energy Storage System: Focusing on Efficient Energy Management for Microgrid, " ***Processes***, vol. 7, no. 2, pp. 80, Dec. 2018, 10.3390/pr7020080.
2. Lin, Y.-H.; Hu, Y.-C, "Electrical Energy Management Based on a Hybrid Artificial Neural Network-Particle Swarm Optimization-Integrated Two-Stage Non-Intrusive Load Monitoring Process in Smart Homes, " ***Processes***, vol. 6, no. 12, pp. 236, Nov. 2018, 10.3390/pr6120236.
3. Hart, G.W., "Nonintrusive appliance load monitoring, " ***Proc. IEEE***, vol. 80, no. 12, pp. 1870-1891, Dec. 1992, 10.1109/5.192069.
4. Dinesh, C., Nettasinghe, B.W., Godaliyadda, R.I., *et al.*, "Residential appliance identification based on spectral information of low frequency smart meter measurements," ***IEEE Trans. Smart Grid***, vol. 7, no. 6, pp. 2781-2792, Nov. 2016，10.1109/TSG.2015.2484258.
5. Lin, Y.-H.; Hu, Y.-C, "Electrical Energy Management Based on a Hybrid Artificial Neural Network-Particle Swarm Optimization-Integrated Two-Stage Non-Intrusive Load Monitoring Process in Smart Homes, " ***Processes***, vol. 6, no. 12, pp. 236, Nov. 2018, 10.3390/pr6120236.
6. Wu X, Gao Y, Jiao D, "Multi-Label Classification Based on Random Forest Algorithm for Non-Intrusive Load Monitoring System," ***Processes***, vol. 7, no. 6, pp. 337, Jun. 2019, 10.3390/pr7060337.
7. Liang, J., Ng, S.K., Kendall, G., *et al.*, "Load signature study—part I: basic concept, structure, and methodology," ***IEEE Trans. Power Delivery***, vol. 25, no. 2, pp. 551–560, Apr. 2010, 10.1109/TPWRD.2009.2033799.
8. Hassan, T., Javed, F., Arshad, N., "An empirical investigation of V-I trajectory based load signatures for non-intrusive load monitoring," ***IEEE Trans. Smart Grid***, vol. 5, no. 2, pp. 870-878, Mar. 2014, 10.1109/tsg.2013.2271282.
9. Du, L., He, D., Harley, R.G., *et al.*, "Electric load classification by binary voltage–current trajectory mapping," ***IEEE Trans. Smart Grid***, vol. 7, no. 1, pp. 358-365, Jan. 2016, 10.1109/TSG.2015.2442225.
10. Wang, L., Chen, X., Wang, G., "Non-intrusive load monitoring algorithm based on features of V-I trajectory," ***Electr. Power Syst. Res***, 157, pp. 134–144, Apr. 2018, 10.1016/j.epsr.2017.12.012.
11. Huang, T.D., Wang, W.S., Lian, K.L., "A new power signature for nonintrusive appliance load monitoring," ***IEEE Trans. Smart Grid***, vol. 6, no. 4, pp. 1994-1995, Jul. 2015, 10.1109/tsg.2015.2415456.
12. Tabatabaei, S.M., Dick, S., Xu, W., "Towards non-intrusive load monitoring via multi-label classification," ***IEEE Trans. Smart Grid***, vol. 8, no. 1, pp. 26-40, Jan. 2017, 10.1109/TSG.2016.2584581.
13. Gillis, J.M., Alshareef, S.M., Morsi, W.G., "Nonintrusive load monitoring using wavelet design and machine learning," ***IEEE Trans. Smart grid***, vol. 7, no. 1, pp. 320-328, Jan. 2016, 10.1109/TSG.2015.2428706.
14. Le, T.-T.-H.; Kim, H., "Non-Intrusive Load Monitoring Based on Novel Transient Signal in Household Appliances with Low Sampling Rate," ***Energies***, vol. 11, no. 12, pp. 3409, Dec. 2018, 10.3390/en11123409.
15. Guo, Z., Wang, Z.J., Kashani, A., "Home appliance load modeling from aggregated smart meter data," ***IEEE Trans. Power Syst.***, vol. 30, no. 1, pp. 254-262, Jan. 2015, 10.1109/TPWRS.2014.2327041.

16. Anderson, K.D., Berges, M.E., Ocneanu, A., *et al.*, "Event detection for Non-intrusive load monitoring," ***in IEEE Conf. Ind. Electr. Society***, Oct. 2012, 10.1109/IECON.2012.6389367.

17. He, K., Stankovic, L., Liao, J., *et al.*, "Power-spectrum-based wavelet transform for nonintrusive demand monitoring and load identification," ***IEEE Tran. Smart Grid***, vol. 9, no. 3, pp. 1739-1747, May. 2018, 10.1109/TSG.2016.2598872.

18. Makonin, S., Popowich, F., Bajić, I.V., *et al.*, "Exploiting HMM sparsity to perform online real-time nonintrusive load monitoring," ***IEEE Trans. Smart Grid***, vol. 7, no. 6, pp. 2575-2585, Nov. 2016, 10.1109/TSG.2015.2494592.

19. Aiad, M., Lee, P.H., "Non-intrusive load disaggregation with adaptive estimations of devices main power effects and two-way interactions," ***Energy Buildings***, pp. 131-139, Oct. 2016, 10.1016/j.enbuild.2016.08.050.

20. Kim, H., Marwah, M., Arlitt, M.F. *et al.*, " Unsupervised disaggregation of low frequency power measurements," ***in Eleventh Siam Int. Con. Data Mining***, pp. 747-758, Apr. 2011, 10.1137/1.9781611972818.64.

21. Johnson, M.J., Willsky, A.S., "Bayesian nonparametric hidden semi-Markov models," ***J. Mach. Learn. Res.***, vol. 14, no. 1, pp. 673, Feb. 2013, 10.1109/MCS.2012.2225913.

22. Kong, W., Dong, Z.Y., Ma, J., *et al.*, " An extensible approach for non-intrusive load disaggregation with smart meter data," ***IEEE Trans. Smart Grid***, vol. 9, no. 4, pp. 3362-3372, Jul. 2018, 10.1109/TSG.2016.2631238.

23. Kolter, J.Z., Batra, S., Ng, A.Y., "Energy disaggregation via discriminative sparse coding," ***in 23rd Int. Conf. Neural Inform. Process. Syst. Curran Associates***, pp. 1153-1161, Jan. 2010, http://dx.doi.org/.

24. Singh, S., Majumdar, A., " Deep sparse coding for non-intrusive load monitoring," ***IEEE Trans. Smart Grid***, vol. 9, no. 5, pp. 4669-4678, Sept. 2018, 10.1109/tsg.2017.2666220.

25. Khandelwal, T., Rajwanshi, K., Bharadwaj, P., *et. al.*, "Exploiting appliance state constraints to improve appliance state detection," ***in ACM Int. Conf. Future Energy Syst.***, pp. 111-120, May. 2017, 10.1145/3077839.3077859.

26. Basu, K., Debusschere, V., Bacha, S., *et al.*, "Nonintrusive load monitoring: a temporal multi-label classification approach," ***IEEE Trans. Ind. Inform.***, vol. 11, no. 1, pp. 262-270, Feb. 2015, 10.1109/TII.2014.2361288.

27. Yang, C.C., Soh, C.S., Yap, V.V., "A systematic approach in appliance disaggregation using k-nearest neighbours and naive Bayes classifiers for energy efficiency," ***Energy Effic.***, vol. 11, no. 1, pp. 239-259, Aug. 2017, 10.1007/s12053-017-9561-0.

28. Henao, N., Agbossou, K., Kelouwani, S., *et al.*, "Approach in nonintrusive type I load monitoring using subtractive clustering," ***IEEE Trans. Smart Grid***, vol. 8, no. 2, pp. 812-821, Mar. 2017, 10.1109/TSG.2015.2462719.

29. Zhao, B., Stankovic, L., Stankovic, V., "On a training-less solution for non-intrusive appliance load monitoring using graph signal processing," ***IEEE Access***, vol. 4, pp. 1784–1799, Apr. 2016, 10.1109/access.2016.2557460.

30. Altrabalsi, H., Liao, J., Stankovic, L., *et al.*, "A low-complexity energy disaggregation method: Performance and robustness," ***IEEE Symposium Compu. Intell. Appl. Smart Grid***, Jan. 2015, 10.1109/CIASG.2014.7011569.

31. Kolter, J.Z., Johnson, M.J., " REDD: A public data set for energy disaggregation research," ***in Proc. 2011 SustKDD Workshop Data Mining Appl. Sustainability***, 2011, 10.1109/TIA.2013.2283318.

32. Liao, J., Elafoudi, G., Stankovic, L., *et al.*, "Non-intrusive appliance load monitoring using low-resolution smart meter data," ***in IEEE Int. Conf. Smart Grid Commun***, pp. 535-540, Nov. 2014, 10.1109/SmartGridComm.2014.7007702.