A Novel Event-based Non-intrusive Load Monitoring Algorithm

Elnaz Azizi, *Student Member, IEEE*, Mohhamd-Taghi Hamidi-Beheshti, *Member, IEEE*, and Sadegh Bolouki, *Member, IEEE*,

Abstract—Non-intrusive load monitoring (NILM), aims to infer the power profiles of appliances from the aggregated power signal via purely analytical methods. Existing NILM methods are susceptible to various issues such as the noise and transient spikes of the power signal, overshoots at the mode transition times, close consumption values by different appliances, and unavailability of a large training dataset. This paper proposes a novel event-based NILM classification algorithm mitigating these issues. The proposed algorithm (i) filters power consumption signals and accurately detects all events, (ii) extracts specific features of appliances, such as operation modes and their respective power consumption intervals, from their power consumption signals in the training dataset, and (iii) labels with high accuracy each detected event of the aggregated signal with an appliance mode transition. The algorithm is validated using REDD with the results showing its effectiveness to accurately disaggregate low frequency measured data by existing smart meters.

Index Terms—Demand-side management, clustering, event detection, non-intrusive load monitoring.

I. INTRODUCTION

Due to the unpredictable nature of both generation, caused by renewable energy resources, and consumer demand, maintaining the balance between generation and demand is one of the main challenges in smart grids [1]. Residential demand-side management programs have thus emerged as a promising set of methods to strike such balance [2]. Among them, non-intrusive load monitoring (NILM), that is the process of extracting the power consumption profile or operating pattern of each appliance from the aggregated power consumption signal of a house using purely analytical methods, has gained a great deal of attention in recent years. Practical and efficient, NILM provides consumers with an opportunity to track the energy consumption of each appliance and voluntarily change their usage patterns to save energy and reduce the cost while maintaining their comfort which also results in higher stability and efficiency of the power grid [3].

The concept of NILM was first introduced in 1992 by Hart [4]. Since then, a variety of analytical algorithms have been proposed to address the NILM problem. These algorithms employ various features and parameters such as voltage, current, and active and reactive power signals of a house. Since measuring the active power is cost-efficient, a majority of studies have focused on this feature alone [5]. NILM research based on the active power signal diverged into two main lines of study, (i) state-based algorithms that consider each

E. Azizi, M.T. Hamidi-Beheshti and S. Bolouki are with Department of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran. e-mails: {e.azizi, mbehesht, bolouki}@modares.ac.ir

appliance as a finite-state machine and disaggregate the total power signal based on the learned model of state transitions of appliances [6] and (ii) event-based algorithms, which are based on the edges or considerable variations of the signal caused by turning ON/OFF of appliances or their other mode transitions [2]. Due to the low computational complexity of event-based techniques, they have proved more popular than the state-based ones [7].

In designing an event-based algorithm, multiple challenges are involved. The first one lies in the event detection part caused by the presence of noise, spikes, uncertainties in the voltage of the grid, and overshoots in appliances' power consumption signals. The second challenge is closeness of different appliances' power consumption values which makes them somewhat indistinguishable. The third and last challenge is that high volume training datasets and ground-truth information about each appliance are scarce in practice, although a small amount of data can perhaps be collected for each residential building. Overcoming these challenges, this paper proposes an event-based NILM algorithm with competitive accuracy, which first via a novel method detects events in the power signals, then extracts specific features and information about appliances from their consumption profiles in a small training dataset, and finally utilizes them to disaggregate the aggregated power consumption signal.

A. Related Work

The most well-known state-based NILM algorithms are the Hidden Markov Model (HMM) [8] and its variants such as Factorial HMM methods [9]. The main drawback of these methods is the requirement for a large training dataset to construct and learn the model. Computational complexities of these methods also increases exponentially by adding a new appliance [10]. However, event-based NILM techniques which deal with the detected events of the aggregated signal and classify them have lower computational complexities in comparison with state-based ones [11]. Recent research of event-based NILM falls into two main categories, namely unsupervised and supervised methods [12]. Unsupervised NILM algorithms, tackling the so-called blind source problem, deal with the case where no prior information about appliances is available. In these methods, events are detected and different clustering algorithms such as subtractive clustering [13] and k-means [14] are applied to them. They detect different clusters of appliances without assigning a label to each cluster. Despite some success in the case where all appliances have only two

(ON and OFF) modes, these algorithms have been ineffective in dealing with multi-mode appliances [14], [15].

In contrast with unsupervised NILM methods, supervised algorithms such as NILM classification algorithms require prior metadata information about the number of appliances and their operation modes as well as a training dataset containing appliances' consumption profiles for a period of time. Considering modes of appliances as class labels, various classification methods such as KNN, multi-label classification [16], [17], and deep learning [18] have been utilized in this field. These methods have proved to be significantly more accurate than their unsupervised counterparts, particularly in the presence of multi-state appliances [18]. However, their main drawback is the need for an enormous training dataset that is not in general feasible to collect [19]. Therefore, extracting useful information from a small training dataset for the NILM classification problem has become a topic of great interest in the past few years [20].

B. Contributions

This paper proposes a novel event-based NILM algorithm that minimizes the ground-truth data required, performs well in analyzing real data measured by existing meters, and remains efficient and accurate even for large numbers of appliances. Major contributions of this work are detailed below.

1) Event-based algorithms are highly dependent on detection of events. Therefore, the event detection algorithm used for the NILM purpose should be accurate in the sense that it should not miss any actual event or mistake fluctuations of the signal as an event. We propose in Section III a novel Statistics-based method that filters the signal and detects events with a 100% accuracy.

2) For NILM as a classification problem, the number of operation modes of appliances and their respective power consumption values are key to assigning labels. In most of the existing literature, these modes are obtained by visually analyzing the appliances' power consumption signals in the training dataset. We introduce a clustering approach in Subsection IV-A, using in part the linkage-Ward algorithm, which automatically extracts appliances' modes and their respective consumption values. Then, in Section V, a novel classification technique, with competitive accuracy, is established for the NILM problem employing previously extracted features.

3) Existing classification algorithms consider appliances' power consumption values at each mode as their main characteristics. However, the appliance consumption pattern, its transitions between different modes, their ON duration period, and their probability of occurrence are also key information that can be used to distinguish two appliances with close power consumption values. Analyzing the training dataset in Section IV, we extract these features of appliances and utilized them for label refinement.

C. Paper Organization

The remainder of this paper is organized as follows. The terminology and problem statement are presented in Section II. In Section III, the proposed signal filtering and event detection techniques are described. The feature extraction methods are then detailed in Section IV, followed by the proposed classification method in Section V. The effectiveness and accuracy of the proposed algorithms are evaluated and compared with other algorithms using the REDD [21] in Section VI. Finally, Section VII concludes the paper.

II. TERMINOLOGY AND PROBLEM STATEMENT

In this section, we present the terminology and the event-based NILM classification problem considered in this paper.

A. Notions and Terminology

In this research, power refers to active power. *Operation modes* of an appliance refer to a fixed set of modes, including the OFF mode, in which the appliance can operate. Appliances are assumed to have two or more operation modes. The *operating mode* of an appliance is the mode in which the appliance is operating at a specific point in time. When no ambiguity results, the term *mode* is used to refer to an operation mode or operating mode. A *state* of an appliance is defined as its power consumption amount in one of its operation modes. Since this amount is assumed to vary at least slightly over time, a state is represented by a fixed closed interval within the set $\mathbb{R}$ of real numbers. One notices that there exists a state corresponding to each operation mode of an appliance.

The *aggregated power consumption signal*, or simply the *aggregated signal*, refers to the sum of power consumption signals of all appliances of a house or specific appliances of interest. The term *non-intrusive load monitoring*, or NILM, in this work is then defined as extracting the sequence of operating modes of each appliance from the aggregated signal. This NILM problem is sometimes referred to as the NILM classification problem. While deducing individual power consumption signals of appliances from the aggregated signal is also a NILM problem, that one views as a NILM regression problem, it is not considered in this work.

Given the power consumption signal of an appliance, an *event* is a change in the signal value caused by a mode transition of the appliance. Similarly, an event of the aggregated signal is a change in the signal value caused by a mode transition of any of the appliances contributing to the aggregated signal.

B. Event-based NILM Problem

The event-based NILM (classification) problem can be described as the process of assigning proper labels to events of the aggregated power consumption signal, where the set of labels consists of all appliance mode transitions. It should be noted that a training set in the form of a set or sequence of events and their corresponding labels is in general not immediately available. Instead, it has to be derived from the given individual appliances' power consumption signals over a period of time. No additional information, such as the number of modes of each appliance or their nominal power consumption values, is available. It is assumed throughout the paper that the given signals are measured in discrete time.
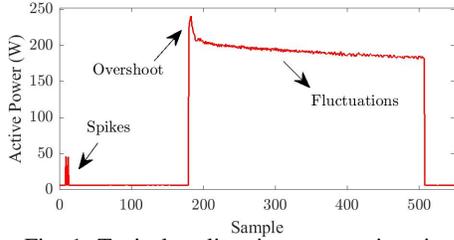
Fig. 1: Typical outliers in consumption signal

III. Signal Filtering and Event Detection

A fundamental part of event-based NILM is detecting events accurately. Event detection should be executed on individual appliances' power consumption signals in the training set as well as the aggregated signal in the test set. In a vast majority of the literature, an event is detected based on the difference between two consecutive sampled values. More precisely, if the absolute value of this difference is greater than a certain threshold, an event is considered to have occurred between the two sampling times. Existing threshold-based event detection techniques heavily rely on the threshold that is selected manually given the dataset in hand. Thus, they are not expected to perform as well on the meter's data of a different residential house. Beside this extensibility issue, there appears to exist a fundamental limit on the accuracy of threshold-based event detection techniques, which is caused by fluctuations of voltage in the power grid and noise, spikes, and the various ranges of overshoots in the signal, as shown in Fig. 1 [22], [23]. In this section, we propose a novel Statistics-based algorithm that overcomes all these challenges and achieves 100% accuracy in event detection.

While the mainstream view of an event is a *significant* value change in the signal, our view of an event is an *uncommon* value change. Thus, considering a set formed based on value changes in the signal, we search for "outliers" of the set. As it will be explained later in this section, this set consists of the min/max ratios between consecutive sampled values of the signal, subtracted from 1. Of course, careful considerations should be made with regard to transient spikes and lengthy overshoots during mode transitions, as they would also be outliers of the formed set. Detecting these spikes and overshoots and filtering them will also prove significant for getting more accurate results in the event-based NILM problem.

Our proposed event detection algorithm consists of three main steps, 1) outlier detection, 2) filtered signal construction and 3) event detection, which will be discussed in the following subsections.

A. Outlier Detection

Different fields of research have been dealing with the outlier detection problem given a dataset and different methods have been proposed to address it [24], [25]. We herein extend a Statistics-based outlier detection method to suit the NILM problem. As it is shown in Algorithm 1, the algorithm starts with calculating the min/max ratio between any two consecutive sampled values of signal $P(t)$, subtracting them from 1,

---

Algorithm 1: Proposed algorithm for outlier detection.

Step 0: Initialize the parameters, $i = 1$, $M(:) = 0$ and $M_o(:) = 0$
Step 1: Get signal $P(t)$, $t = 1, \ldots, T$
Step 2: while $t \leq T - 1$ **do**
    $S(t) = \{P(t), P(t+1)\}$
    $M(t) = 1 - \frac{min(S(t))}{max(S(t))}$
    $t = t + 1$
Step 3: Compute sd as the standard deviation of M
Step 4: Extract outliers' instances based on sd
while $t \leq T - 1$ **do**
    if $M(t) > sd$ **then**
        $M_o(i) = t$;
        $i = i + 1$;

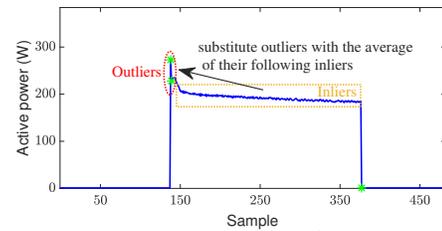Fig. 2: Construction of the filtered signal

and saving them in a vector M. Then, the standard deviation of M is computed. Finally, for any t, if $M(t)$ is grater than the calculated standard deviation, it is considered an outlier and t is saved as the outlier occurrence instance in vector M_o of outlier instances.

B. Filtered Signal Construction

Having performed outlier detection, outliers' instances are obtained, as well as instances that are not outliers, referred to as *inlier* instances. In this so-called filtering step, the aim is to flatten spikes and overshoots of the signal. To achieve this aim, the signal value at each outlier instance is substituted with the mean of the signal values at the following consecutive inlier instances, as shown in Fig. 2. You may note that, unlike for spikes and overshoots, the signal values at actual event times will not experience significant change in the filtering step.

C. Event Detection

In the final step, to detect events, Algorithm 1 is applied to the filtered signal. Now, all detected outlier instances are considered as event instances. We note that transient spikes and overshoots of the original signal have already been flattened in constructing the filtered signal, meaning that they can no longer be mistaken for events.

IV. Feature Extraction

The most common specific feature of appliances utilized in NILM algorithms is their power consumption values at their operation modes. However, since different appliances may have operation modes with close power consumption

values, an effective load disaggregation algorithm should use additional appliance features broadly called *finger prints* in [26]. In this section, we propose different methods to extract operation modes of appliances and additional features from a small training dataset.

A. Modes and States of Appliances

States of an appliances, defined as the power consumption interval corresponding to its operation modes, are its most useful features widely used for the NILM purpose. Therefore, detecting the number of modes of an appliance and their corresponding states plays a crucial role in the accuracy of NILM. As opposed to most of the existing literature that extracts an appliance's modes/states visually using its power consumption values in the training set or by using the datasheet of the appliance, we propose a novel clustering-based approach for appliance modes/states extraction from its power consumption signal in a systematic fashion.

Our approach is based on the linkage-Ward (LW) clustering algorithm [27]. The objective function of the LW algorithm is the squared sum of distances between data points and the their cluster centroids. The LW clustering algorithm first treats each data point as a cluster of its own, which means that the initial value of the objective function is 0. Then, clusters are merged together, one pair at every stage, based on the following merging policy: clusters A and B are merged if $\Delta(A, B)$ is a minimum among all pairs of clusters,

$$\Delta(A, B) = \sum_{p \in A \cup B} \|p - m_{A \cup B}\|^2$$

$$- \sum_{p \in A} \|p - m_A\|^2 - \sum_{p \in B} \|p - m_B\|^2 \qquad (1)$$

$$= \frac{n_A n_B}{n_A + n_B} \|m_A - m_B\|^2,$$

where m_A, m_B, and $m_{A \cup B}$ are centroids of clusters A, B and $A \cup B$ respectively; and n_A and n_B are the size of clusters A and B. One notices that $\Delta(A, B)$, which is non-negative, is in essence the cost of merging clusters A and B. Thus, the value of the objective function at any given stage is the total cost of all the merging up to that stage. Analyzing the increasing objective function, the *elbow method* [11] is often used to determine the optimal number of clusters, that is to determine when to stop merging. Roughly speaking, according to the elbow method, merging stops when it becomes too costly compared to the merging at the previous stage.

It can be seen from the definition of $\Delta(A, B)$ in (1) that the combination of the LW algorithm and the elbow method is susceptible to unbalanced data. More precisely, when the LW algorithm is close to reaching the optimal number of clusters, the elbow method discourages stopping where two or more small clusters exist since merging them is now not relatively "too costly" even though their centroids may be far apart. For the mode/state extraction purpose, this could be a significant issue, as infrequently occurring modes of an appliance may be lumped into one mode with a wide-ranging power consumption, which would undermine any

attempt of NILM. Therefore, one should take advantage of the LW algorithm in such a way to suit the NILM purpose by discouraging merging of clusters that are far apart. Thus, the following algorithm is suggested for mode extraction.

After filtering the power consumption signal of each appliance in the training dataset, the LW algorithm is applied to the signal's data considering K clusters, where $K \geq 10$. The cluster centroids are then computed and sorted in descending order. From this stage forward, a different merging policy, entitled as *distance-based* policy, is adopted that only depends on the distance between cluster centroids. It starts by considering the cluster with the highest centroid as the *root cluster*. If its centroid's distance to the next highest centroid is less than 15% of the root cluster's centroid, the two clusters merge and the step is repeated considering the merged cluster as the new root cluster. Otherwise, the cluster with the next highest centroid to the root cluster's centroid is considered as the new root cluster and the step is repeated. The algorithm terminates when no further merging can take place. The flowchart of the proposed mode extraction method is illustrated in Fig. 3, where centroids of the root cluster and the cluster with the next highest centroid are denoted as C_r and C_j.

B. Transition Intervals of Appliances

Consider an arbitrary mode transition of an appliance from a (relatively) high state $H = [H^{min}, H^{max}]$ to a (relatively) low state $L = [L^{min}, L^{max}]$. Then, the interval of this transition is defined as

$$[H^{min} - L^{max}, H^{max} - L^{min}]. \qquad (2)$$

C. Transition Participation Indices

Two appliances with some overlapping transition intervals can be difficult to separate in the aggregated signal. The *participation index* of their transitions defined as (3), obtained from each appliance's usage pattern in the training set, may help separate these appliances. Given a transition T of an appliance, its participation index P_{par} is defined as

$$P_{par} = \frac{1}{N_{days}} \sum_{day=1}^{N_{days}} \frac{\text{number of } T \text{ in } day}{\text{number of events in } day}. \qquad (3)$$

where N_{days} is the total number of days of training dataset in which the transition T happened. In other words, this parameter measures the daily average of contribution of a specific transition of an appliance in events of the total signal.

D. Additional Features of Appliances

Analyzing the power consumption signals of appliances in the training dataset, one observes that some of the appliances exhibit very specific behavioral patterns. Some of these features, that can further help improve the accuracy of load disaggregation, are listed below.

- As shown in Fig. 4, the dishwasher exhibits the same pattern when it is ON. As this pattern is complex, using it may hurt
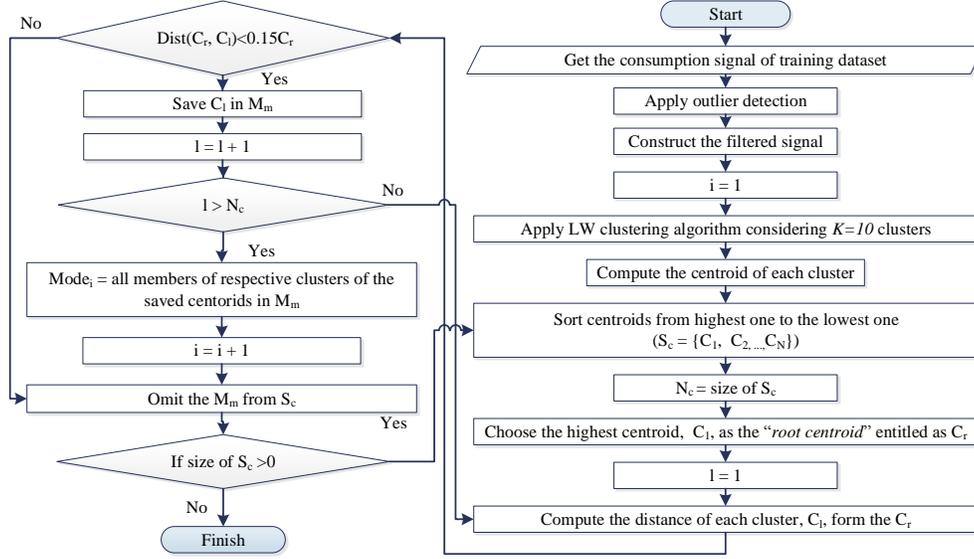
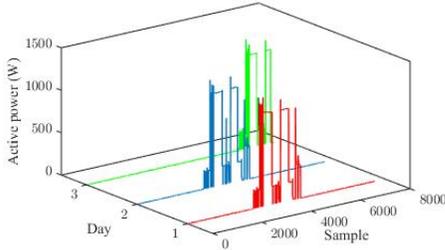Fig. 3: Flowchart for the proposed mode extraction algorithm

Fig. 4: Daily usage pattern of the dishwasher

efficiency of NILM. However, one notices that in any given day, the dishwasher operates in either all or non of its modes. This simple characteristic will prove significant for NILM.

- One observes that not all possible mode transitions of multi-mode appliances can ever occur.
- Some appliances appear to have unique mode transition overshoots in their signals.
- The time period between two consecutive ON samples (with OFF samples in between) of different appliances are significantly different.

V. CLASSIFICATION

In this section, a novel classification algorithm is proposed to address the NILM problem, that is to determine each event in the aggregated power consumption signal corresponds to which appliance mode transition. To label a given event, the classifier first obtains all transition intervals containing the value of that event. As a simple example, consider two ON/OFF appliances 1 and 2, whose ON modes correspond to states [890,1000] and [970,1050], respectively. Given an event with value 980 in the aggregated test signal, the classifier first associates this event with OFF-to-ON transition of appliance 1

as well as that of appliance 2. Afterwards, taking other specific features of each appliance into considering along with the test signal's behavior near the event, a single label is assigned to the event.

Defining N_e and N_T as the total number of events in the test signal and the number of all possible appliance mode transitions, respectively, let L_e be an $N_T \times N_e$ binary-valued matrix, with columns corresponding to events and rows corresponding to mode transitions, which represents predicted labels for events of the test signal. Obviously, an element 1 of L_e indicates that its row's corresponding transition is the predicted label of its column's event. The proposed classification algorithm is detailed in 4 steps below.

Step 1: Given an event of the test signal and a mode transition of an appliance, if the event value is within the transition interval, the respective element of L_e is labeled 1. Otherwise, it is labeled 0. It should be clear that since transition intervals may overlap, some events may be assigned multiple labels in this step. We also point out that some events may remain unlabeled, which means that matrix L_e may have some all-zero columns. Each of these unlabeled events is then labeled the transition whose interval is closest to the event value. We note that the distance between a value and an interval is calculated as the minimum distance between the value and any point within the interval.

Step 2: Analyzing the daily aggregated signal, it can be observed that in a majority of time samples, appliances are in their OFF modes. With that in mind, one should focus on parts of the aggregated signal where at least one appliance is ON. In particular, given the aggregated signal, one obtains its *cycles*, where each cycle starts with an event succeeding an all-OFF sample and ends with the nearest event preceding an all-OFF sample. Fig. 5 illustrates how cycles of an aggregated signal are derived. Over each cycle, the labels assigned to the
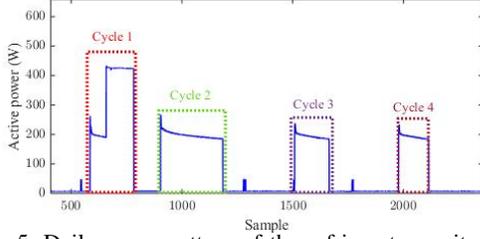
Fig. 5: Daily usage pattern of the refrigerator vs its cycles

events should be *compatible*.

To clarify what is meant by labels compatibility over a cycle, one thinks of an undirected graph in which nodes represent all mode vectors θ, where each element θ_a of θ is a mode of appliance a. Two nodes are then connected by an edge if their corresponding vectors differ in exactly one element. It should be clear that an edge represents a single appliance mode transition, while noting that two different edges may correspond to the same transition. Now, a sequence of labels over a cycle are said to be *compatible* if starting from the all-OFF node of the graph, one can walk according to the labels in the sequence and terminate at the all-OFF node. In other words, labels/transitions over a cycle are compatible if they form a cycle in the graph constructed above.

Using the compatibility condition described, the multiple labels assigned to some events can be narrowed down as some labels are deemed inadmissible. More precisely, a label within a cycle is removed if it is not part of any compatible sequence of labels over the cycle. As an example, in cycle 1 of Fig. 5, after Step 1, assume that the first and second events are assigned single labels, that are mode transitions of appliance 1; while the third event is assigned two labels, one a mode transition of appliance 1 and one a mode transition of appliance 2. Then, in Step 2, the mode transition of appliance 2 is ruled out as a label of the third event since it is not part of any compatible sequence of labels from Step 1 over cycle 1. In other words, appliance 2 cannot possibly be ON when the third event occurs.

Step 3: After Step 2's cycle-based label refinement, each event may still be assigned multiple labels. Considering extracted specific features of some appliances as described in IV-D, some labels are removed from the multi-labeled events.

Step 4: Finally, based on participation indices of appliance transitions explained in IV-C, the most probable label is chosen for the multi-labeled events.

VI. SIMULATION STUDY

In this section, the accuracy and effectiveness of our algorithms proposed in Sections III–V for the NILM purpose are evaluated by applying them to a low-frequency dataset, gathering which is practical as it can be done using existing smart meters [28]. The dataset considered consists of 28 days of power consumption data for seven appliances of house 1 in the REDD [21]. These appliances are listed as oven (OV), microwave (MW), kitchen outlets (KO), bathroom GFI (BGFI), washer/drier (W/D) with a high consumption state, refrigerator (RFG), and dishwasher (DW). The first five listed
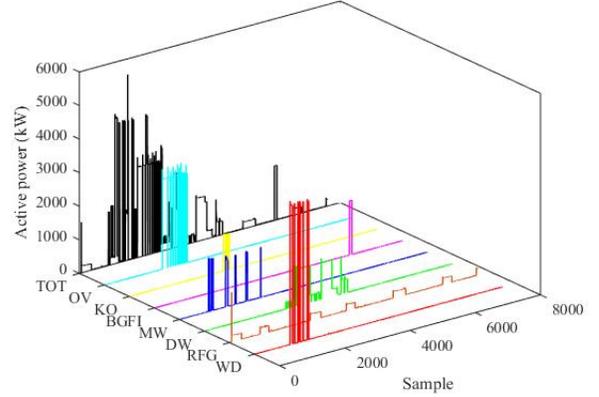

Fig. 6: Total consumption vs the consumption of each appliance

appliance only have ON and OFF modes, while the last two have more than two operation modes. Fig. 6 shows the power consumption signals of all seven appliances in a day of the dataset. To measure the accuracy of the proposed classification, the $F_{measure}$ is used [29],

$$F_{measure} = \frac{2 \times TRP \times RC}{TRP + RC}, \tag{4}$$

$$TPR = \frac{TP}{TP + FN}, \ \ RC = \frac{FP}{FP + TN}, \tag{5}$$

where TRP and RC show the precision and recall; TP is true positive, FP is false positive, TN is true negative, and FN is false negative.

In the following subsections, first the proposed filtering end event detection method are applied to training and test dataset. Then, based on the filtered signal of each appliance, their specific features are extracted. Finally, considering these features the proposed classification technique is utilized to disaggregate the test signal.

A. Signal Filtering and Event Detection

The filtering and event-detection method of Section III is applied to individual appliances' power consumption signals in the training dataset as well as the aggregated signal in the test dataset. As an example, Fig. 7 illustrates the outliers of the signal, the overshoots and spikes in the signal and the constructed filtered signal of the dishwasher for a period of time, respectively.

B. Feature Extraction

Having individual appliances' signals filtered, their features are extracted via methods of Section IV as discussed below.

1) States of Each Appliance: Applying the LW-based clustering method of Subsection IV-A to the filtered signal of each appliance, its modes and their corresponding states are obtained. One recalls that the merging policy is that of the LW algorithm until 10 clusters are obtained, then changes to that of distance-based policy. As an example, Table I shows the 10 obtained clusters for modes/states of the dishwasher before the merging policy changes from the LW method, and Table II shows the final obtained states of all appliances.
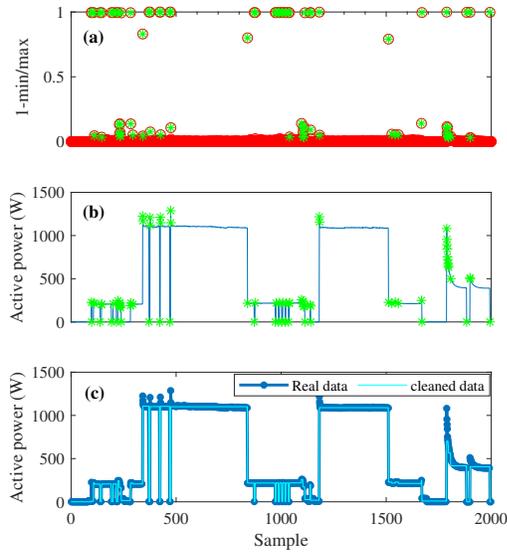
Fig. 7: Filtering the dishwasher's signal: (a) inliers vs. outliers in 1-min/max ratios of consecutive samples, (b) detected outliers in the power consumption signal, (c) the orignal signal vs. the filtered signal

TABLE I: Minimum, maximum, and centroid of the 10 LW-obtained clusters for dishwasher's modes/states.

	min	max	centroid
	154	183	168
	198	208	205
	210	231	219
	236	261	236
Active power (W)	398	420	449
	416	496	489
	500	598	589
	643	737	680
	1078	1110	1099
	1115	1247	1173

2) Transition Intervals of Appliances: Intervals of probable mode transitions of all appliances are now derived using (2). When a mode transition is from or to the OFF mode of an appliance, it simply coincides with the state corresponding to the other appliance mode involved in that transition. Considering probable transitions of multi-mode appliances, non-trivial transition intervals [817-1049] and [155-240] should be considered for DW and RFG, respectively.

3) Transition Participation Index: One recalls that transition participation index is only used to separate appliance tran-

TABLE II: States corresponding to non-OFF modes of appliances

App	States
DW	157-183,198-261,398-496,643-737,1078-1247
RFG	153-183,185-260,415-425
MW	1439-1598
BGFI	1580-1620
KO	1064-1087
W/D	2641-3000
OV	4138-4157

TABLE III: Transition participation index

	Group	Corresponding app	P_{par}
Overlapping transitions group	1	BGFI	0.11
		MW	0.20
	2	KO	0.17
		DW	0.73

TABLE IV: $F_{measure}$ for different methods

App	[20]	ML-KNN [30]	Proposed method
DW	0.59	-	0.96
RFG	1	0.88	0.81
MW	0.40	0.76	0.82
BGFI	0.97	0.88	0.91
KO	1	0.84	0.86
WD	0.92	0.94	1
OV	-	0.86	1
LIGHT	-	0.76	-
AVERAGE	0.81	0.80	0.90

sition with overlapping intervals. Table III shows participation index for different groups of the overlapped appliances.

C. Classification

To apply the proposed classification method, events of the filtered test signal are detected. Then, the classification-based algorithm is applied to detected events based on the transition intervals of appliances. Finally, after cycle-based label refinement, to refine the remained multiple labeled events and choose the most probable label for each event, the following specific features of appliances are considered:

1) Based on the power consumption signal of the dishwasher, one observes that during every ON cycle, all its modes occur. Thus, if the transition with interval [643-737] of the dishwasher, which happens to not overlap with any other transition interval, is not detected, the dishwasher can be assumed OFF during that whole day. Consequently, all candidate labels/transitions involving a non-OFF mode of the dishwasher in that day can be ruled out.

2) Three modes for the refrigerator have been detected, namely the OFF mode, the lower ON mode, and the higher ON mode. It is observed in the training dataset that the transition from the OFF mode to the higher ON mode never happens. Thus, all candidate labels of such can be removed.

3) Overshoot values appearing in the refrigerator's signal are higher than 500 W as opposed to those of the dishwasher. This can be used to separate transitions of the refrigerator and dishwasher with overlapping intervals.

Applying aforementioned features, for each remaining event with multiple labels, the participation index is calculated for overlapped appliances separately. The appliance which has close P_{par} to the calculated participation index of appliances in training dataset, is assigned to events. Table IV shows the high accuracy of our classification method for each appliance in comparison with the results of [20] and [30]. Keeping in mind that a higher number of appliances should diminishes the accuracy of NILM, one notes that the number of appliances considered in this work and [20] are seven and

six. On the other hand, considering multi-mode appliances such as dishwasher increases the complexity of disaggregation. However, the accuracy of our proposed method in which we have considered dishwasher is higher than [30] which did not considered dishwasher in appliances' set.

VII. Concluding Remarks and Future work

In this paper, we have proposed a novel classification method to address the NILM problem given a small dataset. The proposed algorithm has three main phases: 1) filtering training and test signals and accurately detect their events using a statistics-based method, 2) extracting features of appliances, most notably their modes and states via a clustering approach that in part uses the LW clustering method, 3) proposing a classification algorithm labeling events of the aggregated test signal with mode transitions of appliances, where various features and techniques are utilized to enhance its accuracy. The proposed event detection and modes/states extraction methods have been done in a systematic fashion in such a way to perform well for any set of power consumption data. The proposed filtering method, feature extraction techniques, and event-based NILM classification algorithm have been validated using the REDD. Results show 100% accuracy in event detection. Juxtaposing the results of our classification algorithm with two recently introduced event-based NILM methods indicate a relatively high accuracy of our algorithm.

Reconstructing the power consumption signals of appliances, which can be cast as a regression problem as opposed to the classification problem we considered in this work, is one of the main challenges of event-based NILM problems. We aim to modify our algorithms, mainly the one in Section V, to address the reconstruction problem. Moreover, due to the lack of a training dataset for each residential building, we wish to move a step further to use transfer learning to circumvent the training phase of the proposed method, with the practical assumption that nominal values for appliances' power consumption are given.

References

[1] M. S. H. Nizami, M. J. Hossain, and E. Fernandez, "Multiagent-based transactive energy management systems for residential buildings with distributed energy resources," *IEEE Trans. Industrial Informatics*, vol. 16, no. 3, pp. 1836–1847, 2020.

[2] M. Lu and Z. Li, "A hybrid event detection approach for non-intrusive load monitoring," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 528–540, 2019.

[3] T. Ji, L. Liu, T. Wang, W. Lin, M. Li, and Q. Wu, "Non-intrusive load monitoring using additive factorial approximate maximum a posteriori based on iterative fuzzy *c*-means," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6667–6677, 2019.

[4] G. W. Hart, "Nonintrusive appliance load monitoring," *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992.

[5] V. Rathore and S. K. Jain, "Non intrusive load monitoring and load disaggregation using transient data analysis," in *2018 Conference on Information and Communication Technology (CICT)*. IEEE, 2018, pp. 1–5.

[6] B. Zhao, K. He, L. Stankovic, and V. Stankovic, "Improving event-based non-intrusive load monitoring using graph signal processing," *IEEE Access*, vol. 6, pp. 53 944–53 959, 2018.

[7] Q. Liu, K. M. Kamoto, X. Liu, M. Sun, and N. Linge, "Low-complexity non-intrusive load monitoring using unsupervised learning and generalized appliance models," *IEEE Trans. Consumer Electronics*, vol. 65, no. 1, pp. 28–37, 2019.

[8] S. Singh and A. Majumdar, "Deep sparse coding for non–intrusive load monitoring," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4669–4678, 2017.

[9] R. Bonfigli, E. Principi, M. Fagiani, M. Severini, S. Squartini, and F. Piazza, "Non-intrusive load monitoring by using active and reactive power in additive factorial hidden markov models," *Applied Energy*, vol. 208, pp. 1590–1607, 2017.

[10] K. He, D. Jakovetic, B. Zhao, V. Stankovic, L. Stankovic, and S. Cheng, "A generic optimisation-based approach for improving non-intrusive load monitoring," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6472–6480, 2019.

[11] E. Azizi, A. M. Shotorbani, M.-T. Hamidi-Beheshti, B. Mohammadi-Ivatloo, and S. Bolouki, "Residential household non-intrusive load monitoring via smart event-based optimization," *IEEE Trans. Consumer Electronics*, vol. 66, no. 3, pp. 233–241, 2020.

[12] B. Zhao, M. Ye, L. Stankovic, and V. Stankovic, "Non-intrusive load disaggregation solutions for very low-rate smart meter data," *Applied Energy*, vol. 268, p. 114949, 2020.

[13] N. Henao, K. Agbossou, S. Kelouwani, Y. Dubé, and M. Fournier, "Approach in nonintrusive type i load monitoring using subtractive clustering," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 812–821, 2015.

[14] W. Kong, Z. Y. Dong, J. Ma, D. J. Hill, J. Zhao, and F. Luo, "An extensible approach for non-intrusive load disaggregation with smart meter data," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 3362–3372, 2016.

[15] C. Dinesh, S. Makonin, and I. V. Bajić, "Residential power forecasting using load identification and graph spectral clustering," *IEEE Trans. Circuits and Systems II: Express Briefs*, vol. 66, no. 11, pp. 1900–1904, 2019.

[16] S. M. Tabatabaei, S. Dick, and W. Xu, "Toward non-intrusive load monitoring via multi-label classification," *IEEE Trans. Smart Grid*, vol. 8, no. 1, pp. 26–40, 2017.

[17] L. Massidda, M. Marrocu, and S. Manca, "Non-intrusive load disaggregation by convolutional neural network and multilabel classification," *Applied Sciences*, vol. 10, no. 4, p. 1454, 2020.

[18] V. Singhal, J. Maggu, and A. Majumdar, "Simultaneous detection of multiple appliances from smart-meter measurements via multi-label consistent deep dictionary learning and deep transform learning," *IEEE Trans. Smart Grid*, 2018.

[19] C. C. Yang, C. S. Soh, and V. V. Yap, "A systematic approach in load disaggregation utilizing a multi-stage classification algorithm for consumer electrical appliances classification," *Frontiers in Energy*, vol. 13, no. 2, pp. 386–398, 2019.

[20] S. Dash, R. Sodhi, and B. Sodhi, "An appliance load disaggregation scheme using automatic state detection enabled enhanced integer-programming," *IEEE Trans. Industrial Informatics*, 2020.

[21] J. Z. Kolter and M. J. Johnson, "Redd: A public data set for energy disaggregation research," in *Workshop on Data Mining Applications in Sustainability (SIGKDD), San Diego, CA*, vol. 25, no. Citeseer, 2011, pp. 59–62.

[22] N. Henao, K. Agbossou, S. Kelouwani, Y. Dubé, and M. Fournier, "Approach in nonintrusive type i load monitoring using subtractive clustering," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 812–821, 2017.

[23] T. Lu, Z. Xu, and B. Huang, "An event-based nonintrusive load monitoring approach: Using the simplified viterbi algorithm," *IEEE Pervasive Computing*, vol. 16, no. 4, pp. 54–61, 2017.

[24] S. Kandanaarachchi, M. A. Muñoz, R. J. Hyndman, and K. Smith-Miles, "On normalization and algorithm selection for unsupervised outlier detection," *Data Mining and Knowledge Discovery*, vol. 34, no. 2, pp. 309–354, 2020.

[25] Y. Chakhchoukh, S. Liu, M. Sugiyama, and H. Ishii, "Statistical outlier detection for diagnosis of cyber attacks in power state estimation," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*. IEEE, 2016, pp. 1–5.

[26] M. Ma, W. Lin, J. Zhang, P. Wang, Y. Zhou, and X. Liang, "Toward energy-awareness smart building: Discover the fingerprint of your electrical appliances," *IEEE Trans. Industrial Informatics*, vol. 14, no. 4, pp. 1458–1468, 2017.

[27] E. Alpaydin, *Introduction to machine learning*. MIT press, 2020.

[28] A. Elnaz, S. Amin M, B. Mohammad TH, I. Behnam M, and B. Sadegh, "Residential household non-intrusive load monitoring via smart event-based optimization," *IEEE Trans. Consumer Electronics*, vol. 10, no. 4, pp. 4615–4627, 2020.

[29] Z. Xu, W. Chen, and Q. Wang, "A new non-intrusive load monitoring algorithm based on event matching," *IEEE Access*, vol. 7, pp. 55 966–55 973, 2019.

[30] D. Li and S. Dick, "Residential household non-intrusive load monitoring via graph-based multi-label semi-supervised learning," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4615–4627, 2018.