Sustainable Cities and Society 43 (2018) 487–494

Contents lists available at ScienceDirect

Sustainable Cities and Society

journal homepage: www.elsevier.com/locate/scs

Energy disaggregation of overlapping home appliances consumptions using a cluster splitting approach

Misbah Aiad*, Peng Hin Lee

School of Electrical and Electronic Engineering, Nanyang Technological University, 50 Nanyang Avenue, 639798, Singapore

ARTICLE INFO

Keywords:
Non-intrusive load monitoring (NILM)
Energy disaggregation
Energy savings
Cluster splitting
Normality tests

ABSTRACT

Non-intrusive load monitoring (NILM) is a set of techniques that aims to decompose the aggregate energy consumptions of a household into the energy consumed by the respective individual appliances. When some of the home appliances have power consumptions levels that overlap with each other, it becomes a challenging problem to disaggregate the energy consumed by each of these appliances. In this work, we present an approach that split the clusters of the overlapping energy consumptions into the respective energy consumed by the individual appliances. The proposed approach involves firstly to analyze the cohesion of devices clusters to determine if a cluster should be split into two clusters. The proposed cluster splitting approach was tested on cases of overlapping devices clusters from six real houses available from the REDD public data sets. The results showed that the performance of the proposed approach depends on the degree of overlapping of the devices clusters, on whether the clusters are tight or loose and on the sizes of the clusters. The proposed approach can be applied to a clustering-based load disaggregation method as a subsequent step to deal with situations of overlapping appliances consumptions, so as to improve the overall energy disaggregation accuracy.

1. Introduction

Energy savings and energy management in the residential sector are of increasing interests in the fields of energy systems and smart grids. In the broad scope of energy saving, there is a potential to achieve energy saving by replacing existing appliances with energy-efficient appliances. In addition, it is essential that the energy consumers be knowledgeable to adopt behaviors that lead to reduce their energy consumptions. In general, households' energy smart energy meters provide information about the total aggregate consumption of a household. Studies conducted have shown that energy savings of between 5%–20% are achievable when the household occupants have information on the detailed breakdown of the energy consumed by each individual device in their homes (Henao, Agbossou, Kelouwani, Dubé, & Fournier, 2017). This detailed information were found influencing to occupants to use appliances (especially those of high power consumptions) less often (Henao et al., 2017).

Non-intrusive load monitoring (NILM) is a set of techniques that aims to decompose the total aggregate energy consumptions of a household into the energy consumed by the respective individual appliances that are present in the household (Aiad & Lee, 2016a; Zeifman & Roth, 2011; Zoha, Gluhak, Imran, & Rajasegara, 2012). The NILM problem can be defined as follows

Given the total measured power consumption $p(t)$ at any time instant t, we can write

$$p(t) = \sum_{j=1}^{N} p_j(t) \tag{1}$$

The NILM problem is to estimate each $p_j(t)$, the individual power consumed by the device j at the time instant t, where $j = 1, 2, \dots N$, $t = 1, 2, \dots T$ and N is the total number of the operating appliances within the load disaggregation period T (Zoha et al., 2012). In NILM, the total aggregate measurements from smart meters are used by the disaggregation processes to detect the individual devices and infer their operations. The main objective of an energy disaggregation approach is often to achieve a high disaggregation accuracy for individual appliances. However, there are many challenges in tackling the NILM problem. In this paper, we focus on a specific challenge when two or more energy consumptions clusters overlap and it is necessary to split them into smaller distinct clusters to represent the respective individual appliances. The proposed approach can be applied to a clustering-based load disaggregation method as a subsequent step to deal with situations of overlapping appliances consumptions, so as to improve the overall energy disaggregation accuracy. The proposed clusters splitting method in this paper is not a clustering method nor a NILM approach by itself,

* Corresponding author.
E-mail address: misb0001@e.ntu.edu.sg (M. Aiad).

https://doi.org/10.1016/j.scs.2018.08.020
Received 11 January 2018; Received in revised form 1 July 2018; Accepted 16 August 2018
Available online 15 September 2018

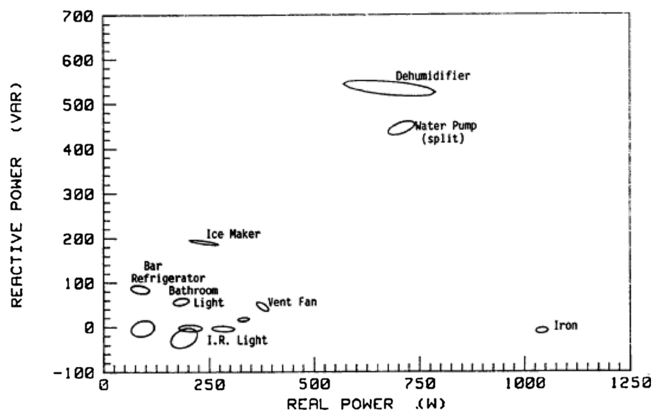

Fig. 1. Features space (real versus reactive power). From (Hart, 1992). © 1992 IEEE.

but an addition action to be used in *clustering-based* NILM approaches to deal with cases of overlapping clusters and decide if they should be split The proposed approach refines the outcomes of the home devices clusters by investigating which clusters that may be needed to be split into smaller clusters and then splitting them when necessary.

2. Background and related work

The concept of NILM was first proposed by Hart in 1992 (Hart, 1992) where the proposed disaggregation process consisted of two steps. The first step was to capture the significant features whenever a transition in the aggregate signal is detected when a device switches state from the ON state or the OFF state. The primary features that were measured are the change in the real and the reactive power consumptions. The second step is to group these features, also called distinct signatures or attributes into distinct clusters of the different appliances. It was observed that higher power consuming devices are usually easier to be classified since they are placed far from other devices in the features space. There are may be overlapping devices clusters, especially those lower power consuming devices as shown in Fig. 1 (Hart, 1992). The cases of overlapping appliances clusters the basic challenge that we aim to tackle with the proposed clusters cohesion test in this work.

Some work on NILM tried to obtain information from the aggregate measurements that may help in the improvement of the disaggregation accuracy (see, for example, (Aiad & Lee, 2016b). Others have also obtained information, without the need of additional sensors to be installed, such as the time of use, duration of operations of the devices and outdoor weather information to aid in the disaggregation process. In general, extracting useful information and features from a data stream falls under the broad scope of data mining techniques as discussed in (Yu, Haghighat, & Fung, 2016). Besides the extracting of additional features, improving the disaggregation and identification algorithms is also essential to enhance the overall disaggregation accuracy. Hidden Markov models (HMMs) and their extensions were found functional in the representations and identifications of household appliances. Kim, Marwah, Arlitt, Lyon, and Han (2010) tested four extensions of the HMMs for better representations and disaggregations. They also used non-electrical features such as the distributions of the ON/OFF durations and the inter-dependency between appliances. (Kolter and Jaakkola (2012) expressed a new approximate inference method based on the factorial HMM (FHMM). They used two complementary models, namely, the additive and the difference FHMM. The additive FHMM captures the total aggregate output signal while the difference FHMM encodes the signal differences between subsequent levels when a transition is detected in the aggregate total signal. Parson, Ghosh, Weal, and Roger (2014) proposed an approach in which a one-time supervised learning process with an already labeled data set was used to create general probabilistic models of appliances. Thereafter, the general models are tuned to previously unseen appliances in an unsupervised manner. Gabaldón, Molina, Marín-Parra, Valero-Verdu, and Alvarez (2017) proposed a non-intrusive approach based on the Hilbert transform to identify the most suitable loads that can be used to validate customers' effective responses.

In the domain of machine learning and artificial intelligence, overlapping of clusters has been a challenge in many real-life applications including overlapping appliances clusters in NILM. Wagstaff, Cardie and Schroedl (2001) proposed constraint-based clustering which modifies the *K-means* method (Bishop, 2006) by considering a set of constraints to be satisfied. The *K-means* clustering is an iterative method of clustering based on achieving the minimum distances of the elements to their cluster center. Lu and Leen (2007) used a mean field approximation to produce the *Penalized Probabilistic Clustering* (*PPC*) to handle increasing complexity in large data sets. In the PPC, a set of constraints are defined based on prior information. These constraints define item-to-item relations, for example, whether they must or must not ultimately belong to a specific cluster. Instead of using item-to-item constraints as in *PPC*, the *Class-Level Penalized Probabilistic Clustering* (*CPPC*) was proposed by defining cluster-to-cluster constraints (Preston, Brodley, Khardon, Sulla-Menashe, & Friedl, 2010). These constraints provide the probabilities of how likely two clusters should be merged, split or kept unchanged. As the number of clusters is often significantly less than the total number of items to be classified, *CPPC* provides a noticeable reduction in complexity over the *PPC* (Preston et al., 2010). The hierarchical divisive clustering is a top-to-bottom clustering that repeatedly partition a present cluster into two smaller clusters till reaching some stopping criteria. Common approaches in considering such splitting decisions Some of the common factors used to decide on whether to split or not to split are size priority, cluster cohesion tests and dissimilarities between the possible sub-clusters (Ding & He, 2002).

Clustering methods used in load disaggregation are usually efficient in identifying non-overlapping and higher power consuming devices (Zeifman & Roth, 2011). In (Farinaccio & Zmeureanu, 1999), the standard clustering approach was extended by filtering and smoothing mechanisms to deal with power variations. Nonetheless, this method requires excessive training and it was applicable mainly to high power loads (Zoha et al., 2012).

3. Proposed approach to split devices clusters

We propose an approach that utilizes some prior information to check if a cluster should be split and perform the splitting, once a decision is made. We are not proposing a new clustering method to tackle the basic NILM problem of energy disaggregation. Such clustering approaches to the NILM problem in energy disaggregation can be found in, for example, (Farinaccio & Zmeureanu, 1999; Hart, 1992). From the outcomes of such clustering approaches to the NILM problem or as Fig. 1 has shown, it may result that there are overlapping appliances clusters and the proposed work in this paper seeks to address this. The proposed approach in clusters splitting aims to obtain the individual appliance clusters from the overlapping clusters to enhance the overall disaggregation of the energy consumed by all the respective individual appliances. The ultimate objective is to improve the overall disaggregation accuracy by investigating cases of devices clusters that may have originated from two different home devices. An important prior assumption here is the normality of the data distributions within each cluster. In the context of NILM, each cluster represents the power consumptions of an individual household appliance. This assumption applies mainly to ON/OFF devices with consistent power consumption and to finite state machines (FSMs). In our approach, we have avoided splitting methods that are based on size priority because the devices clusters sizes are usually unpredictable. The sizes of the devices clusters essentially depend on how long the device will be in operation, i.e. the