Automatic Recognition of Major End-Uses in Disaggregation of Home Energy Display Data

Michael ZEIFMAN, Senior Member, IEEE, Kurt Roth, and Johannes STEFAN
Fraunhofer Center for Sustainable Energy Systems, Cambridge, USA

Abstract—Disaggregation can make actionable the information provided by home energy displays or smart meters. However, the known disaggregation methods require training to match disaggregated data to actual appliances. We propose a statistical approach for automatic matching between the disaggregated data and major end-uses.

I. INTRODUCTION

Home energy displays (HEDs) usually provide whole house, near real-time electricity consumption information [1]. The primary goal of HEDs is to help save energy for the homeowner. Since the whole-house information is not actionable, data disaggregation, or nonintrusive appliance load monitoring (NIALM), can be implemented to provide appliance-specific information [2]. However, available NIALM methods suitable to the HED data (i.e., real power sampled at 1 Hz) require significant occupant efforts to train the algorithms to recognize the actual appliances [2], [3]. In this paper, we propose an approach for automatic recognition of the major end-uses in the disaggregated data. We implement power- and time features for statistical characterization of the end-uses and apply a simple Bayes classifier to matching between the disaggregated patterns and appliance classes.

II. METHODOLOGY

A. NIALM Method

To disaggregate the low-frequency HED data, NIALM algorithms usually use stepwise power changes as the main feature. These stepwise power changes occur when, e.g., the on-off appliances are turned on or off. Since the distributions of the power draw of appliances often overlap, the accuracy of traditional NIALM methods is usually low [2]. We have recently developed a new NIALM method that implements time features, duration of time on and duration of time off, to better separate the overlapped in power draw appliances [3]. The method is capable of finding on-off appliances or their combinations and tracking them in time, but it yet cannot match the found appliances to the actual household appliances. Figure 1 illustrates this problem.

B. Major End-Uses

Among various appliances and appliance groups, the following end-uses account for more than 80% of average electricity household consumption [4]: (1) space cooling systems, (2) space heating systems, (3) domestic hot water, (4) lighting, (5) refrigerators, (6) electric clothes dryers, and (7) consumer electronics. All of these end-uses, but dimming lights that are not considered in this work, can be approximated as on-off appliances. In this work, we concentrate on these seven end-uses. For the consumer electronics, we consider televisions and desktop computers with monitors as these devices account for more than 85% of consumer electronics energy consumption [4].

Figure 1. Power consumption of an appliance disaggregated by our NIALM algorithm. Which actual appliance does this pattern correspond to? (clothes dryer). Note that this particular appliance is a combination of two on-off devices, motor and heating element.

C. Statistical Characterization by Prior Knowledge

Prior data exist for the power draw information of the seven major end-uses we selected. For example, Ref. [5] lists average lighting power draw, number of bulbs, number of switches, and average duration of time on for different room types (e.g., living room, kitchen), building establishments (e.g., multifamily) and bulb types (e.g., incandescent, CFL or halogen). These data can be used in designing statistical distribution models.

In this work, we use the maximum entropy concept [6] to select a statistical model that is most suitable to the available prior knowledge. This concept yields a Beta distribution for the case of the known range and mean value [7]. However, the underlying random variables, the duration of time on and the change of power, are not statistically independent for the lighting loads. This dependence is based on the room type. For each room type, nonetheless, these variables can be assumed to be independent so that the joint probability density function (PDF) for each room is a product of the marginal PDFs. For the entire household lighting, the joint distribution function will be a mixture of the joint PDFs corresponding to the room types.

Advantage of the discrete wattage values of the bulbs on the market can be taken to better characterize the distribution of the power draw. Assuming that the fluctuations of the power draw around the face value can be characterized by a normal
distribution, we arrive at a convolution of normal and beta distributions for the power draw of lighting.

Figure 2 shows a marginal PDF of power draw for lighting loads in a living room of a single detached U.S. home. The characteristic peaks on the PDF are due to the discrete wattage of the bulbs.

The other six end-uses are easier to characterize, because the underlying prior knowledge is not as granulated as the knowledge on the lighting loads. We get the joint Beta distributions for the power and durations of time on/off using the average values from the literature. Note that the power off distribution of, e.g., the clothes dryers is a mixture of two Beta distributions, with one component corresponding to the cycling load of the heating element (see Figure 1) and the other component corresponding to the time between drying cycles.

D. Statistical Characterization by Fine Features

To further improve the statistical characterization, we performed signature collection for the seven end-uses. Typical signatures for lighting and televisions are shown in Figures 3 and 4.

It is seen in the Figures that the lighting load is very steady, and that the televisions are characterized by either significant fluctuations of the power draw due to the variability of brightness and sound (CRT) or by typical power drops at channel changes (LCD). These fine features can be characterized mathematically and be used in conjunction with the general prior knowledge (see Section II.C).

III. TESTING

We have designed a naïve Bayes classifier [6] on the basis of the statistical models for the seven end-uses considered in Section II. For all other possible loads, we use a single uniform joint distribution of power/time. We then conducted testing in residences using the TED home energy displays [8] for both aggregated power and submetering. The testing was performed during a cooling season so that only six end-uses were present at the test homes.

The disaggregation was performed by our set of algorithms [3]. The obtained disaggregated data were matched to the six end-uses or to the class “else.” For the most challenging end-uses, lighting and consumer electronics, the classification accuracy in terms of the F-measure [3] was 65% and 70% respectively if we used only the prior-knowledge characterization. The use of the fine features dramatically increased the accuracy, to 92% for the lighting and to 90% for the consumer electronics.

IV. CONCLUSION

A statistical classification scheme, capable of automatic recognition of disaggregated HED data is proposed and preliminary tested. The scheme implements prior knowledge on the major electric residential end-uses along with their fine features to get reasonable classification accuracy. More comprehensive tests of the scheme and development of more complex appliance models are currently underway.

REFERENCES