Event-based Household Energy Consumption Disaggregation of Individual Appliances

Michael Shell
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332–0250
Email: http://www.michaelshell.org/contact.html

Homer Simpson
Twentieth Century Fox
Springfield, USA
Email: homer@thesimpsons.com

James Kirk
and Montgomery Scott
Starfleet Academy
San Francisco, California 96678-2391
Telephone: (800) 555–1212
Fax: (888) 555–1212

***Abstract*—Even with advances in smart grid technology and a growing demand for cost-effective energy consumption, detailed information about energy usage is often not available for residential electricity consumers. One reason is that energy usage is typically monitored at single point by the utility, only providing information on the aggregate power consumption. In this paper, we attempt to disaggregate the energy usage data into specific appliances from single-point sensing measurements. Our method involves extracting events from the time-series data to obtain relevant features. We focus on two feature sets: low frequency and high frequency. For each set, we create an event detection algorithm for feature extraction. We present the results of our approach on a publicly available dataset.**

I. Introduction-ML

In this section, we provide the motivation and a brief description of the project. A short discussion of related work in this area, including the methodology of previous competition winners is also provided. This is followed by an overview of our approach, description of the raw data, and evaluation method.

This project follows the Belkin Energy Disaggregation Competition on kaggle.com [**?**]. The goal is to disaggregate household energy consumption into individual appliances. The dataset used for this project is publicly available on the competition website and contains voltage and current harmonics data for four households. The dataset also contains high frequency noise (EMI) data for each household.

Although the presence or absence of such EMI signatures and changes in voltage or current can indicate when a particular appliance is in use, classification can become challenging if the number of appliances in the home is large. Additionally, the signature of some appliances may drift or vary over time due to operating conditions and the mode in which they are used. Predictive modeling is then required to make an inference about the appliance class given a particular signature. The challenge is to accurately classify end-uses of energy at a fine-grained, appliance level. One application of this project is to continuously monitor real-time power consumption, broken down by electrical appliance. Consumers can then view their energy usage and cost at a detailed enough level to determine cost-effective energy saving changes to usage patterns. Based on the description above, our problem statement can be summarized as follows:

> Using the harmonic and EMI measurements as features, the objective of this project is to build a multi-class classification model to determine which appliances are active at any given time (i.e, classify the active appliances based on harmonic and EMI signatures).

A. approach/background

Our method focuses on real-time processing.

Show visualizations -training data -figures show differences domain/power domain - explain reason for looking in difference domain - test data-multiple devices, and steady state, no changes -goal is to detect events on/off -for example, constant power loads, show no change, eg. light

-explain methodology - voltage/current harmonics to P/Q to differences to distances -and why we use this approach -include block diagram -goal is to automate, this task can be done manually, but too tedious -compare to previous and to data mining methods, ours is more interpretable

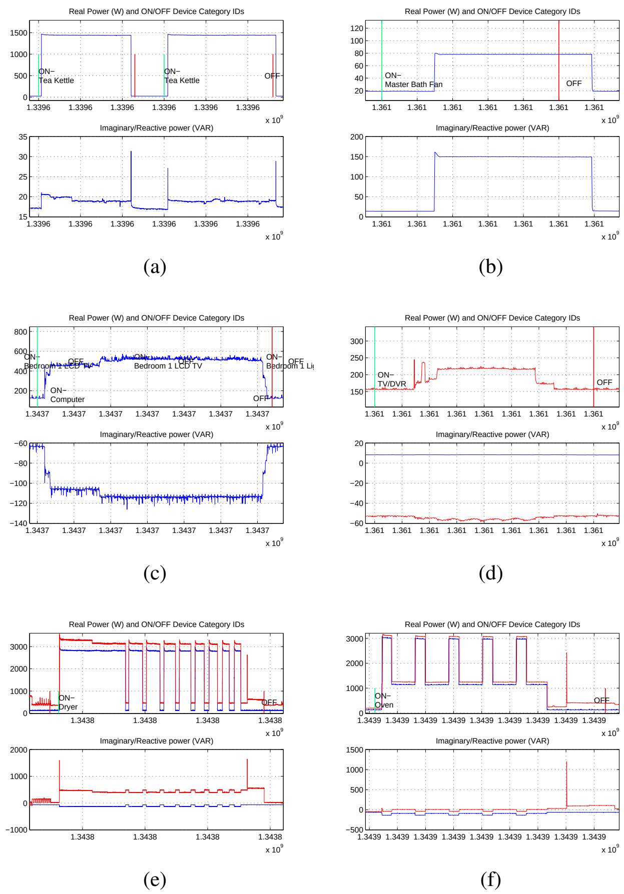

Fig. 1. Load Types (a) Tea Kettle, (b) Master Bath Fan, (c) Computer, (d) TV, (e) Dryer, (f) Oven.

II. Training - Signature Window Extraction

This section concentrates on extracting the signature windows by using the real and the reactive power differences. In particular, two ON/OFF signature windows are trained for each event of the appliances at each Home. The following sections describe training signature windows depending on the shape of the load.

A. Load Types

First, the appliances are categorized into three load types based on its shape. The types include rectangular, non-rectangular, and cyclic loads.

1) Rectangular loads: These type of loads consumes constant power when operating. The load shapes are rectangular. They include resistive loads, such as lights, heaters, and kettle, and reactive loads such as fans. A short duration window consisting of the real and the reactive power differences are used to detect these loads.
2) Non-rectangular loads: These type of loads have a longer transient characteristics when turning on. They include computers and TVs. These type of loads require longer window length than the rectangular loads.

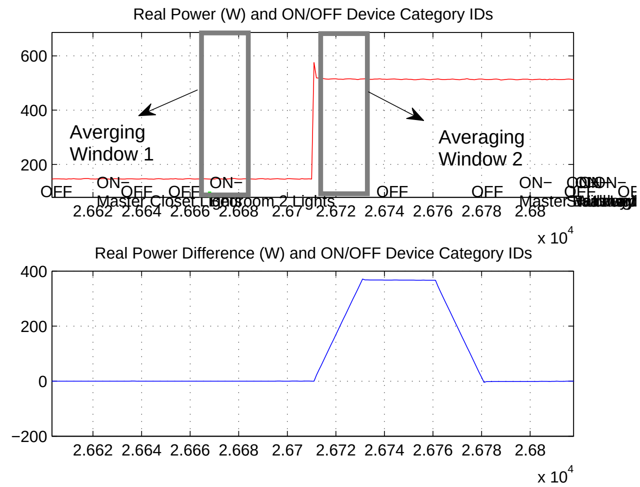

Fig. 2. Two Moving Average Windows.

3) Cyclical loads: - These type of loads change it's shape depending on the cycle of their operation. Dryer, oven, and washers are included in this category. Two approaches are suggested to detect these type of loads; using multiple windows or detecting a notable landmark during the operation.

B. Smoothed Power Differences

Now to train the windows, the real and the reactive power differences are calculated from two moving average windows. The smoothed difference $S(n)$ is defined as,

$$S(n) = \frac{1}{N}\sum_{k=0}^{N} P(n-k) - \frac{1}{N}\sum_{k=N+D}^{2N+D} P(n-k) \qquad (1)$$

where N is the window size and D is the distance between the two windows. It should be noted that if shorter window size is used, it is easier to distinguish events from other events happening at the same incidence. However, the detector becomes less robust to noise.

Using the power difference, the mismatches in the tagging info and the actual event index are corrected. A threshold for the real power difference is set to indicate whether an event occurred, (event occurred if $|S(n)| >$ threshold). Then this corrected indexes are used as the actual event index instead of the tagging info provided. On the top plot of Fig. 3, the tagging info of the master room lights do not align with the actual event incidence. The red markers in the bottom plot shows when the power difference is greater than the threshold and this index is used as the corrected tagging info. It should be noted that we discarded some appliances that consumes very small real or reactive power. For these appliances, using real and reactive power differences will rather cause more false alarms in the test data due to the effects of the noises.

¡¡¡¡¡¡¡ HEAD

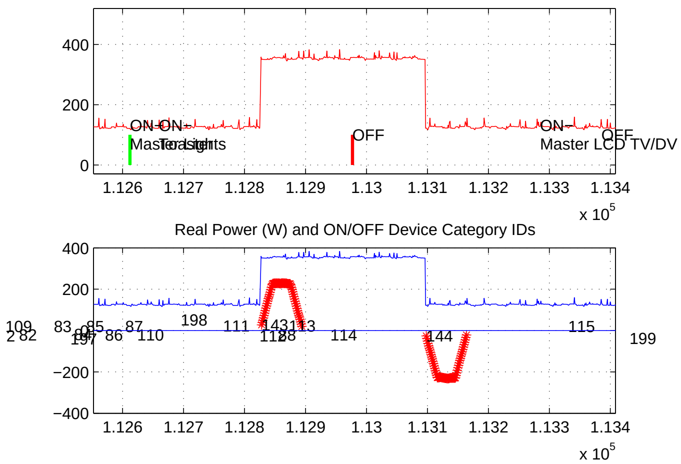

Fig. 3. Tagging Info Correction.

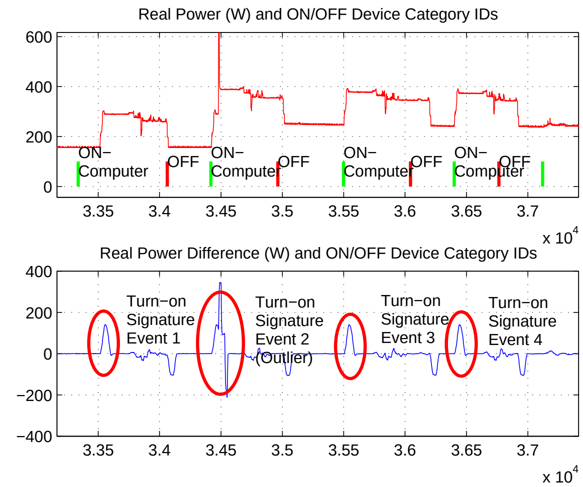

Fig. 5. Outlier.

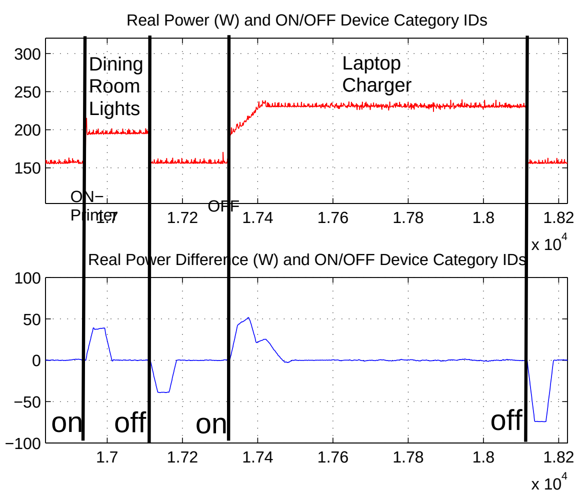

Fig. 4. Simple Distribution Feeder.

C. Signature Window Extraction

After taking the power differences, the next stop is to extract the signatures for on-events and off-events. The windows are extracted starting from the corrected tagging info. If an appliance has an longer transient characteristics (non-rectangular load shapes), longer window size is used to detect event. For example, consider the dining room light and the laptop charger profile shown in Fig. 4. Both the appliances consumes about 38 Watts at event incidence. However, laptop charger then spends additional 38 Watts to fully turn on. Therefore, longer window lengths are used to detect laptop charger than in detecting the dining room lights.

Insert windows used for cyclical loads.

Outlier detection is also held in this stage. Consider the four ON/OFF events of a computer as shown in Fig. 5. It is shown that when event 2 occurred, some background noise are introduced that affected the turn-on signature of the computer. In this case, the second event is regarded as outliers and is not used in training the windows.

=======

First a event detector is used in the real power domain to estimate the exact switching time of the event. By in the training data. This is to account for the mismatches of the provided tagging info and the exact switching index.

A typical rectangular load shape of is shown in Figure **??**. Many appliances turned out to have rectangular load shape and these appliances can be separated using real and reactive power consumption features. Figure **??** shows feature space of appliances that have rectangular load shape. The feature vector of each class are clustered together, implying that the real and reactive power features are sufficient for classifying most of the appliances.

This section concentrates on extracting the feature windows in low frequency data. In particular the real and the reactive power differences are used to extract these windows. First, a separate event detector is used to correct the event detector is used in the Training Data to correct the mismatches

The procedures for calculating the real power and the reactive power include event detection, data smoothing, and power calculation.

D. Load Types

-constant power - static - resistive (lights, heaters, kettle) - P domain and rectangular mainly - resistive + reactive (rotating like fans) - Q is a feature to help classify -dynamic - cyclical (washer dryer) - require rules, ie. landmark detection ¿¿¿¿¿¿¿ 4a1b9769f1a3f8f92b061559235daa35c859fc63

1) Subsubsection Heading Here: Subsubsection text here.

III. Cross-Validation-ML

Discuss cross validation here. this is the whole procedure for generating thresholds.

need figures as examples of P/Q domain, diff domain, distance domain, window shape (big subplot 3x3?)

IV. Test Data

- rules for on/off pairing

V. Conclusion

The conclusion goes here. All insight gained from competition. Too many appliances, noisy test data, what is the benefit of identifying small appliances?, noise necessitates clear features (large distance). also high freqency domain for hard to identify appliances. you can get 20th w/ just P/Q alone.

Acknowledgment

The authors would like to thank...