Transient Signature Detection of Appliances for Household Energy Consumption Disaggregation

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 household energy usage is monitored at no more than a single point by the utility, only providing information on the aggregate power consumption. In this paper, we attempt to disaggregate energy usage data into specific appliances from single-point sensing measurements. Our method involves extracting turn-on and turn-off signature windows from time-series real and reactive power data to obtain transient characteristics of each appliance. We focus on determining the appropriate window size for each appliance in order to capture unique signatures. We present the results of our approach on a publicly available dataset.**

I. Introduction

The objective of residential non-intrusive load monitoring (NILM) is to monitor the major loads in a home from a single-point. The alternative is to monitor each appliance individually, however, this scheme typically adds significant cost. Therefore, the challenge with NILM is to accurately disaggregate household energy consumption into the individual appliance level with data from single-point measurements.

Prior research in the area of NILM has focused on the use of aggregate power consumption patterns as features to identify what appliance is being used and how much energy it is consuming. For example, the authors in [1] discuss various approaches in NILM, including detecting changes in steady-state power measurements and characterizing them as different events. Some challenges reported by the authors include different loads not exhibiting unique signatures in the 2D feature space and the difficulty in determining steady-state features due to turn-on transient noise. Some recommended advanced techniques are to include the 3rd order harmonics as a feature and using turn-on transients for event detection. The methodology in [2] and [3] follows a similar strategy, however, the training and test data are manually generated, resulting in a clean dataset that may not be representative of real household energy usage patterns. Furthermore, the 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.

In this paper, we present a method that focuses on extracting turn-on and turn-off signature windows to obtain transient characteristics of each appliance. The transient characteristics are observed from the difference of two moving average windows from time-series real and reactive power data. We focus on determining the appropriate window length for each appliance in order to capture unique signatures. We present the results of our approach using a publicly available dataset from the Belkin Energy Disaggregation Competition on kaggle.com [2]. The dataset and methodology are described in detail in Section II. Section III describes the approach for extracting signature windows for each appliance from the training data. Section IV describes the cross-validation procedure using the training data. Section V describes application of our approach to the test data. Finally, Section VI discusses results and topics for future work.

II. Approach

In this section, we describe the dataset that was used and our methodology for detecting appliances in the test data. Because the set of appliances in each home includes multiple small devices with similar steady-state load characteristics, our methodology focuses on identifying the unique features when a device turns on and off (the transient periods). Out of approximately 37 appliances per home, only a small subset of appliances have pronounced steady state features that allow discrimination from other appliances or background noise.

A. Dataset

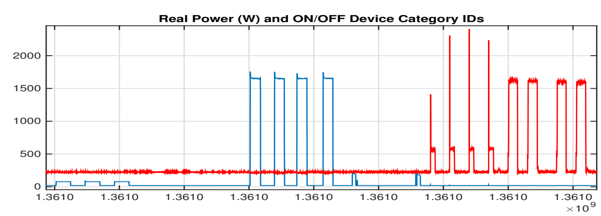
In this paper, we utilize the complete dataset from the Belkin Energy Disaggregation Competition hosted on kaggle.com, provided by Belkin Energy [2]. The dataset is publicly available on the competition website and contains single-point sensing measurements of the first five harmonics of rms voltage and current on two phases for four households (labeled H1-H4). The dataset also contains high frequency noise (EMI) data for each household, but is not required with our approach (see further discussion in Section VI). Although both training and test datasets are provided, only the training data provides labels and corresponding timestamps for each appliance. Table

I shows the number of distinct appliances in each home and the number of training and test sets provided.

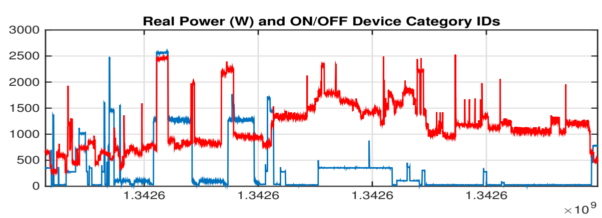
TABLE I
NUMBER OF DISTINCT APPLIANCES AND DATASETS FOR EACH HOME

Home No.	Appliances	Training	Test
H1	38	6	4
H2	37	4	4
H3	37	3	4
H4	36	2	4

Each training set consists of manually generated power consumption data for each appliance in each home. The training data is a time-serires where only one appliance is switched on at a time, although background noise can be present from untagged appliances. Tagging labels in the training data are also provided, however, these are not always precise enough to correctly identify device turn-on and turn-off times. Correction of tagging labels for training data is therefore needed and is discussed in Section III. Fig.1 shows an example of both training and test data from H2. The observations recorded in each training and test set cover approximately 24 hours.



(a) Training



(b) Test

Fig. 1. Example of (a) training and (b) test data from Home 2. In each plot, real power consumption measured from a single point is shown for both phases.

B. Methodology

Our goal is to develop an interpretable yet effective classification approach. It can be observed from Fig. 1 (b) that the test data is the aggregated load of multiple appliances. Therefore, at any given time, it is not possible to determine which appliances are operating by using only the level of power consumption. However, when an appliance is turned on or off, a step change is produced in real and reactive power. By comparing the average power consumed before and after these transient events, we can calculate averaged power differences, facilitating discrimination among appliances. For this reason, our methodology focuses on detecting unique turn-on and turn-off signatures in the power differences domain for each appliance. A flow chart of the proposed methodology is shown in Fig. 2.

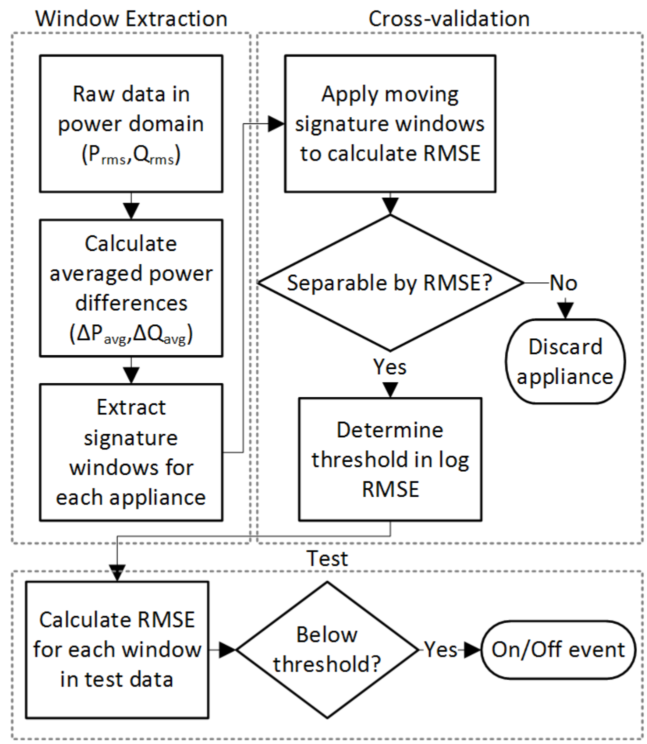


Fig. 2. Flow chart of proposed methodology. The procedure is separated into window extraction, cross-validation, and application to test data.

The first step is to extract transient signature windows for each appliance. Raw data is converted from the power domain (P and Q) into the power differences domain (ΔP_{avg} and ΔQ_{avg}) by taking the difference of two moving average windows in P and Q. After ΔP_{avg} and ΔQ_{avg} are calculated, transient signature windows are extracted for each appliance. Choosing the appropriate window length is a key component to this step and the procedure is explained in detail in Section III.

Once transient signatures are extracted for each appliance in the training data, cross-validation is performed. This step involves moving $\Delta P_{avg}(n)$ or $\Delta Q_{avg}(n)$ past the extracted signature window to calculate the root-mean-squared error (RMSE). Turn-on and turn-off events are detected when the RMSE falls below a specified threshold simultaneously for both real and reactive power differences. The ability of each signature window to detect other events of the same appliance in the training data is used to determine if the appliance is detectable or not. If an appliance cannot be detected during cross-validation, it will not be predicted in the test set. The procedure to determine the correct threshold and separability of each appliance is presented in Section IV.

Finally, for the test set, we use the cross-validated signature windows and threshold values to calculate RMSE and detect appliance turn-on and turn-off events. This process is discussed in Section VI.

III. Signature Window Extraction

This section describes how transient signature windows are extracted by using averaged real and the reactive power differences. In particular, a pair of turn-on and turn-off signature windows are trained for every event of each appliance in the training data. The following sections describe how the signature windows are extracted.

A. Load Types

The appliances are first categorized into three load types based on their load shape in P and Q. The types include rectangular, non-rectangular, and cyclic loads:

1) Rectangular loads: These type of loads consume constant power when operating and their transient period is characterized by a sharp transition from zero to rated power consumption. Their load shapes are generally rectangular and consist of primarily resistive loads, such as lights, heaters, and kettles. Some inductive loads, such as small fans, also exhibit this load shape. 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 period when turning on and include computers and TVs. Their transient period is characterized by a stepwise transition from zero to rated power consumption. Although the power consumed may vary based on how the device is operated, the turn-on signatures can be unique, allowing discrimination against other devices. These type of loads require a longer window length to capture the unique transient shape.

3) Cyclical loads: - The amount of real or reactive power consumed by these loads change depending on the cycle of their operation. Dryers, ovens, and washers are included in this category. However, because they operate on set cycles, there are only a few possible periodic load shapes they can follow. A modified approach is required to detect these type of loads: identifying a notable landmark during appliance operation.

Illustrative examples of these load types are shown in Fig. 3. The dryer for example produces several rectangular step changes over the course of its cycle, resulting in misclassifications if not accounted for. The modified approach for detecting these appliance is discussed below.

B. Averaged Power Differences

Our approach relies on identifying appliance turn-on and turn-off events by detecting unique transient signatures in real and reactive power consumption. To extract transient signatures for each appliance, the averaged power differences are calculated from two moving average windows. The averaged

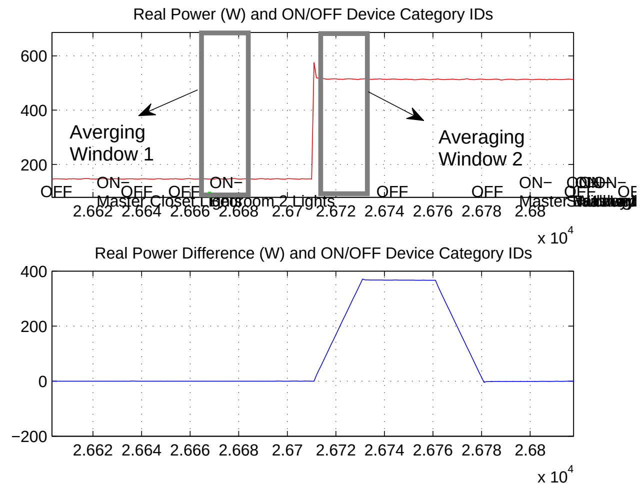


Fig. 3. Difference of two moving average windows to detect transient signatures.

differences for real power, $\Delta P_{avg}(n)$, is defined as,

$$\Delta P_{avg}(n) = \frac{1}{N}\sum_{k=0}^{N-1} P(n-k) - \frac{1}{N}\sum_{k=N+D}^{2N+D-1} P(n-k) \quad (1)$$

where N is the window size and D is the distance between the two windows. The averaged differences for reactive power, $\Delta Q_{avg}(n)$, is calculated using the same N and D. It should be noted that if a shorter window size is used, it is easier to distinguish events happening at the same incidence. However, the detector becomes less robust to noise.

Using the averaged power differences to determine true appliance event indices, tagging labels tagging labels from the raw data are first corrected. A threshold for the real power difference is set to indicate whether an event occurred, (event occurred if $|\Delta P_{avg}(n)| >$ threshold). Then the corrected indexes are used as the actual event index instead of the tagging info provided in the raw data. On the top plot of Fig. 5, the tagging labels of the master bedroom 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; this index is used as the corrected tagging label. It should be noted that some appliances that consumes very small real or reactive power are discarded as they cannot be detected. For these appliances, using real and reactive power differences will cause false detections in the test data due to the effects of the noise.

C. Signature Window Extraction

After taking the power differences, the next stop is to extract the signatures for turn-on events and turn-off events. The windows are extracted using the corrected tagging labels. If an appliance has longer transient characteristics (non-rectangular load shapes), longer window sizes are used to capture the entire unique transient period. For example, consider the dining room light and the laptop charger profile shown in Fig.

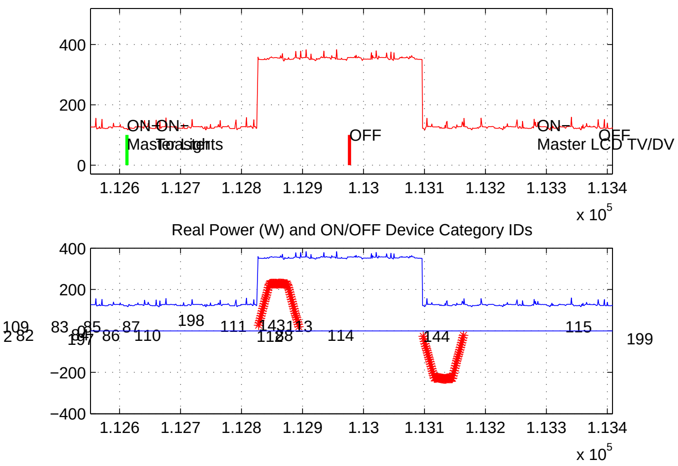


Fig. 4. Provided tagging labels are corrected in the power differences domain.

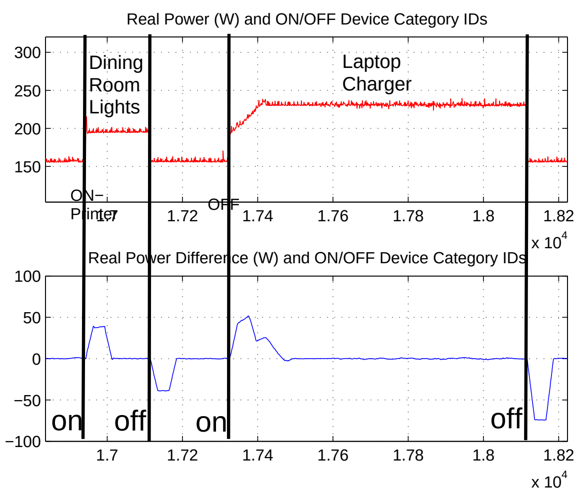


Fig. 5. Example of how choosing the appropriate window length can affect classification.

6. Both the appliances consume about 38 W at event incidence. However, the laptop charger then spends additional 38 Watts to fully turn on. Therefore, longer window lengths are needed to detect laptop charger than the dining room lights.

Cyclical loads exhibit unique characteristics that are different from the rectangular and the non-rectangular loads. These type of appliances are very difficult to be detected by using turn-on and turn-off signature windows. Therefore, for cyclical loads only, a window is trained to capture unique signatures not when the appliances are switching but during their operation. Consider the dryers in Home 2 as shown in Fig. 7. The load shape of the dryer has multiple sharp peaks during its entire operation time. Therefore, a single window can be trained to detect these sharp peaks. Then, the following rule can be made: "While the landmark window is detecting an event, the appliance is operating."

Outlier detection is also held in this stage. Consider the

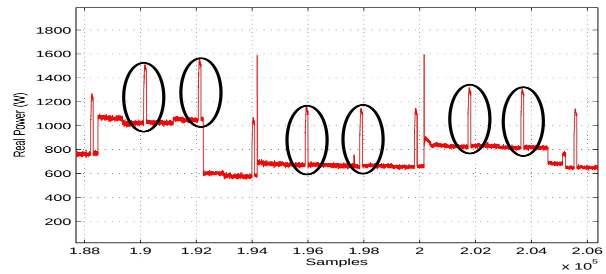


Fig. 6. Unique Characteristics of Dryers in Home 2.

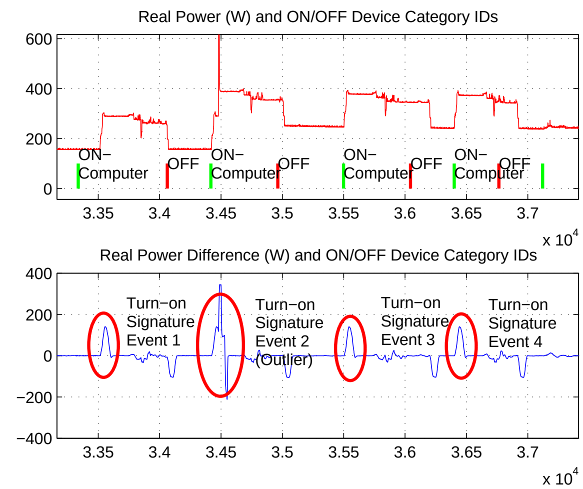


Fig. 7. Computer turn-on event outlier shown in the 2nd event.

four ON/OFF events of a computer as shown in Fig. 8. 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 an outlier and is not used in training the windows.

When the procedure is complete transient signature windows $w_{i,j}(n)$ are defined for i appliances in each home and j events in the training data.

IV. Validation

In this Section, the signature windows trained in Section III are cross-validated using the training set. The RMSE values shown in (2)-(3) are used as metrics for the evaluation.

$$\text{RMSE-P}(n_0) = \frac{1}{N}\sum_{n=1}^{N}(w_{i,j}(n) - \Delta_{avg}P(n+n_0))^2 \quad (2)$$

$$\text{RMSE-Q}(n_0) = \frac{1}{N}\sum_{n=1}^{N}(w_{i,j}(n) - \Delta_{avg}Q(n+n_0))^2 \quad (3)$$

where $w_{i,j}$ is the trained window for j-th event of an appliance with ID i.

When the signature window is scanning its target, both the RMSE-P and RMSE-Q values fall very low. On the other hand, if the window is scanning on a different time stamp, each RMSE-P and RMSE-Q value can rise, fall, or stay the

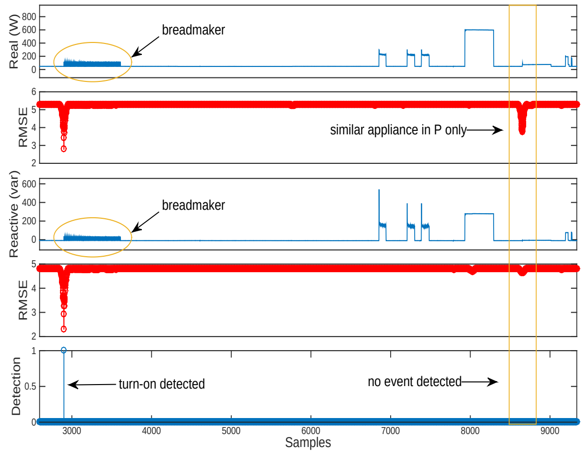


Fig. 8. Cross-validation of appliance "breadmaker" for two events in the training data.

same. It is possible that both the RMSE values drop when the window is detecting different appliance. This happens when the signature windows of the two appliances have a very similar shape and amplitude. To prevent this scenario, RMSE thresholds is set to separate these appliance from being detected. If a RMSE threshold could not be set, the window is discarded. Figure 9 shows the detection output after applying the turn-on window of a breadmaker to the training set. The window successfully detects the breadmaker turning-on with all the other events being not detected.

V. Test Data

We then apply our algorithm on the provided test data and identify appliance on and off intervals. The detections are made from the signature windows we extracted from the training set with the threshold values acquired from cross-validation. The algorithm is shown below.

for Every Time Instant n_0 **do**
  Calculate $RMSE_P(n_0)$ and $RMSE_P(n_0)$
  if $RMSE_P(n_0) < \theta_P$ and $RMSE_Q(n_0) < \theta_Q$ **then**
    Event is detected at time instant n_0
  end if
end for

Although the competition has closed, we can still submit our predictions and see its rank on the leaderboard. Our best submission would have placed 7th. However, the strategy for the test data is changed to achieve the best score. Because average Hamming distance is used, thresholds are adjusted so that only the most likely predictions, i.e., low RMSE values are submitted.

VI. 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.

A. Data Cleaning

Using this dataset, a lot of effort was put into cleaning the data. Specifically, the tagging information for appliance on off events was not precise enough to accurately extract on and off signatures windows automatically. However, if these event tags were generated correctly, event signature windows could be automatically extracted form the training data.

B. Tradeoffs

One of the advantages of this approach is that it is more interpretable as it follows intuition for detecting appliance on and off events. With the proposed approach, misclassifications can be traced back to the signature window length or detection thresholds. This is particularly true when compared to methods utilizing 'black box' predictive models such as artificial neural networks. Furthermore, the proposed approach as the ability for real time processing. We also used a simplified dataset, only needing P and Q. High frequency features could have been added to help classify smaller appliances.

However, one of the primary drawbacks is that manual identification of appropriate window lengths is requried during training.

References

[1] C. Laughman, K. Lee, R. Cox, S. Shaw, S. Leeb, L. Norford, and P. Armstrong, "Power signature analysis," *Power and Energy Magazine, IEEE*, vol. 1, pp. 56–63, Mar 2003.

[2] http://www.kaggle.com/c/belkin-energy-disaggregation-competition.