Disaggregation of Household Energy Consumption into 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.

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

II. Training - Feature Window Extraction

This section concentrates on extracting the feature windows in low frequency data. In particular the real and the reactive power differences at the fundamental frequency (50 or 60 Hz) are used to characterize the on/off signatures of a given appliance.

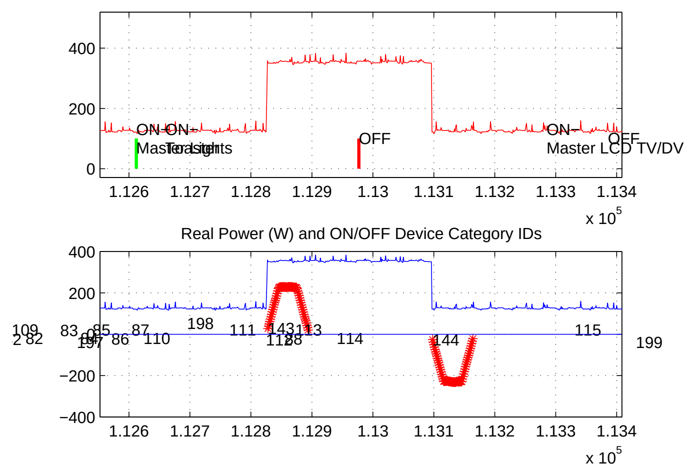

Fig. 1. Simple Distribution Feeder.

A. *Event Detection Using Power Consumption Difference*

The exact time-index of the switching on/off time of the appliances should be detected in order to calculate the power consumption of each appliance. The event detector is applied to both training and testing data due to the following reasons.

1) Training Data: the tagging info of the appliances is provided. However, they do not represent the exact on and off time of the appliances. In addition, there are multiple mislabels such as the on-label and the off-label indicating the same time index. The event detector should be programmed separately to account for these offsets and mislabels.
2) Test Data: the tagging info is not provided. The event detector should be able to capture when an appliance is switched on and off.

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.

B. 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

C. Window making

-because the loads are as described above, we are taking this approach, using different windows (different lengths) for different loads to detect ON/OFF events - these are the most effective approach vs. continuous approach (ours is event based)

-devices have unique features during on off mainly

-outliers - problems with training data, mention issues w/ kaggle ground truth

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...

References

[1] H. Kopka and P. W. Daly, *A Guide to LaTeX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.