Leveraging Smart meters for residential energy disaggregation

Abiodun Iwayemi, Chi Zhou
Electrical and Computer Engineering Department
Illinois Institute of Technology
Chicago, USA
{aiwayemi, zhou}@iit.edu

Abstract**—Smart meters promise to provide residents with the information required for using their appliances in the most energy efficient manner. We demonstrate how smart meter data can be used to infer individual appliance energy consumption by extracting low-frequency transient features which correspond to appliance turn-on and turn-off events. These features are a combination of a transient shape metric and the summary statistics; and are used to detect and classify appliance events. The detected events are fed into a disaggregation algorithm that is robust to missed detections. The combination of our transient features and disaggregation algorithm allows us to achieve a disaggregation accuracy of 94% on a test set of 4 major appliances without tuning to that home. This demonstrates our approaches ability to generalize to unseen homes. Our scheme uses low frequency transients, yet provides performance equivalent to more complex and costly high-frequency sampling approaches.**

Index Terms**—Disaggregation, Non-intrusive load monitoring, smart meter, energy efficiency, energy feedback.**

I. INTRODUCTION

A tremendous amount of money is being spent on smart meter deployments across the US in the hope that they will lead to reduced operational expenses for utilities and an increase in energy efficient behavior by customers. The primary benefit of smart meters to consumers is in the ability to provide real-time energy consumption and pricing data. The underlying premise is that providing consumers with this information will lead to reductions or shifts in their energy usage. The residential sector comprises 30% of electricity consumption in the US, so any reductions in residential energy usage will significantly improve power grid health and reduce greenhouse gas emission.

Unfortunately studies show that the size of energy savings is directly proportional to how detailed the energy feedback information is [1]. In [2] it is shown that detailed feedback generates the greatest savings (more than 12% of annual residential energy usage), but this level of savings are only possible if consumers are provided with real-time, per-appliance usage information combined with personalized recommendations on energy efficient behavior . It is estimated that even a 10-15% reduction in residential energy usage will result in 200 billion kWh of savings – equivalent to the output of 16 nuclear power stations [3].

We address this problem by demonstrating how smart meter data can be leveraged to provide detailed, high-quality, per appliance energy consumption data via Non-intrusive load monitoring (NILM) techniques without the need for expensive sensing equipment.

NILM utilizes signal processing and machine learning techniques to identify individual appliances from within aggregate electricity measurements of a building. This allows us to determine the electricity consumption of individual appliances without the need for metering devices on each appliance. We extend current research on NILM by proposing a smart-meter based platform that utilizes low-frequency transient waveforms extracted from smart meters to uniquely identify large loads within the home with a high degree of accuracy. Most work to date has used high-frequency transients for appliance detection and required costly high-frequency sampling equipment, but we demonstrate that high-levels of accuracy can be obtained at significantly lower sampling rates, using only data that can be extracted from a smart meter. This data can obtained via the home area networking (HAN) interface of Zigbee-equipped smart meters which enables In-home Displays (IHD's) and other power monitoring tools to interface with the smart meter to obtain near real-time data [2].

Our contributions are a system that uses low resolution ($\geq$1 Hz) smart meter data yet provides energy disaggregation with accuracy on par with costlier higher frequency approaches; a novel transient-based approach for classifying residential appliances which generalizes to unseen homes with minimal training or tuning; and a robust disaggregation algorithm that intelligently handles missed appliances events, yet still provides accurate estimates of appliance energy usage.

The rest of our paper is structured as follows: background information on NILM is provided in section II, and related work is reviewed in section III. We discuss the use of low-

978-1-4799-6415-4/14/$31.00 ©2014 IEEE

frequency signatures, and appliance disaggregation in sections IV and V respectively; and evaluate the performance of our approach in section VI. The work is concluded and summarized in section VII.

II. BACKGROUND

Insight into energy use can be provided through non-intrusive load monitoring (NILM) techniques. In NILM, the aggregate power consumption in a building $X_{1:T}$ over a time period 1 to T is disaggregated into the energy consumption N of individual appliances by means of measurements taken at a single point [4]:

$$X_{1:T} = \sum_{i=1}^{N} X_{1:T}^{(i)} \qquad (1)$$

It is non-intrusive because it does not require access to the customer premises to install measurement equipment. In figure 1 the aggregate load is broken down into a refrigerator, washer and dishwasher.

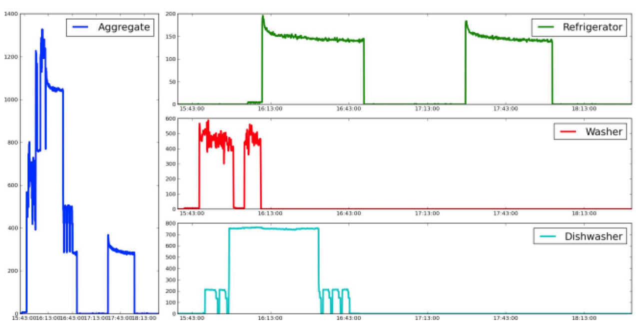

Figure 1. Appliance disaggregation

According to [5], NILM approaches fall into two main categories – low sampling frequency and high sampling frequency approaches. Low sampling frequency NILM uses sampling rates of 1s – 1hr, while high frequency approaches sample at rates ranging from 10kHz – 1 Mhz.

Low and high –frequency NILM approaches can be further subdivided into event-based and non-event based schemes [6]. Event-based approaches look for changes in the aggregate energy signal which correspond to appliance turn on/off events. These features are referred to as appliance signatures and include real and reactive power, and higher-order harmonics.

Non-event based schemes utilize Hidden Markov Models to probabilistically decompose the aggregate signal. Such methods build state models of each appliance, and tend to be more robust to noise and missed detections [7, 8]. They also permit the incorporation of additional features such as duration of appliance operation and dependencies between sets of appliances. The downside of these methods is that the complexity of the models may increase exponentially with the number of appliances in the home, which leads to scalability issues.

III. RELATED WORK

Real and reactive power are used to classify loads using low-frequency NILM [4] or high-frequency sampling combined with higher order harmonics [9]. High-frequency transients (>1 kHz) have been used for appliance detection and disaggregation [9,11] because they enable the disaggregation of overlapping appliances in the frequency domain, where appliances that appear similar in the time domain are more easily distinguished. Unfortunately this approach requires costly high-frequency sampling equipment to capture the noise signatures, transients and high-order harmonics these schemes depend on. Reactive power and higher-order harmonics are typically unavailable using smart meters, which provide only real power at sampling rates of at most 1Hz, and our approach is designed to work with such low-fidelity data while providing comparable performance. In addition, our approach is cheaper and easier to deploy than high-frequency approaches which require costly custom designed sampling hardware [9,11], and professional installation in a customer's distribution panel [9].

Shao [10] utilize appliance motifs to detect and disaggregate appliances. We adopt a similar approach but do not assume that all the appliances in the home are known, and utilize only the turn-on and turn–off transients rather than the entire shape because using the entire shape increases the disaggregation complexity significantly.

We demonstrate a low-frequency transient-based approach for automatically labeling/identifying appliances. We define a transient as any event between two steady-states, and show how these transients can be used to uniquely identify and disaggregate appliances.

IV. LOW FREQUENCY APPLIANCE FEATURES

Exploratory data analysis of appliance energy profiles show that each appliance class has a unique shape or waveform which can be used to uniquely identify it. For example, refrigerator compressor cycles are similar across different manufacturers, and we can leverage this information to identify them. This evident in figure 2, where even though each refrigerator has a unique duty cycle and average power consumption, all have the same basic shape.

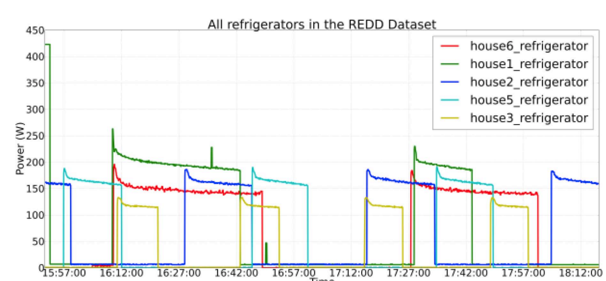

Figure 2. Refrigerator load profiles

We divide appliance events into two periods – steady-states and transients, then classify the transients into turn on and turn off events as seen in figure 3. Classification of appliances is then based on identifying the transient sequences that correspond to a particular appliance.

The use of turn on and turn off transients allows us to detect and disaggregate overlapping appliances much more easily than if we used the entire shape as a feature. It also allows us to detect appliance events using either turn-on or turn-off signatures in the event that only one of the two events was detected.

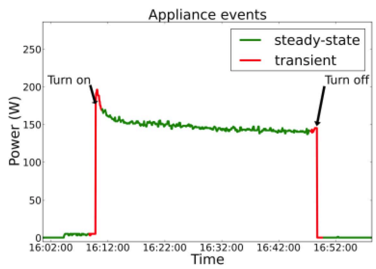

Figure 3. Appliance events

We also examined using only steady-state appliance signatures for appliance clustering but found that we obtained better clustering with transients. This is evident in figure 4, were we plotted the step change in power against the standard deviation of the steady-state event for 3 refrigerators, 1 dryer, 2 microwave ovens and 1 washer from homes 2-6 in the REDD data set [12]. The figure below clearly shows significant overlap in appliances below 1000W, which reduces disaggregation accuracy.

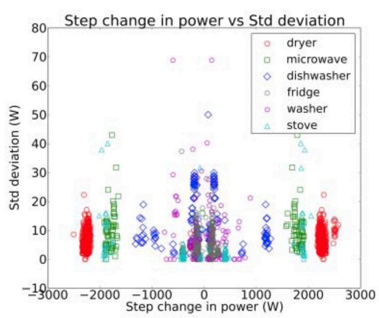

Figure 4. Steady-state signature scatterplot

The similarity between detected shape/event waveforms can measured by means of Dynamic Time Warping (DTW), an algorithm which measures the similarity between time series sequences which may differ in length [13]. DTW is a dynamic programming algorithm which determines the shortest "warping" path between two sequences and returns a metric indicating the distance between the two sequences. The smaller the DTW distance, the greater the similarity, with a DTW distance of 0 indicating that the two sequences are identical. Even though DTW is more complex to calculate than Euclidean distance, it has the advantage of allowing us to compare time series with different lengths. This is essential because different appliance events have varying lengths.

We captured the turn-on and turn-off transients for 5 appliance types – refrigerators, washers, dryers, microwave ovens and cooking ranges from the REDD Dataset[12] , examining a total of 78 appliance snapshots from 11 unique appliances. We extracted the following statistics for each snapshot – the mean power draw, DTW distance between different appliance snapshots, the skew and kurtosis.

From figure 5, it is evident that if we cluster appliances using only DTW distance and mean power consumption, we can uniquely identify each of the appliance categories. This allows us to distinguish between appliances, even those with similar power draws such as microwave and ranges/stoves.

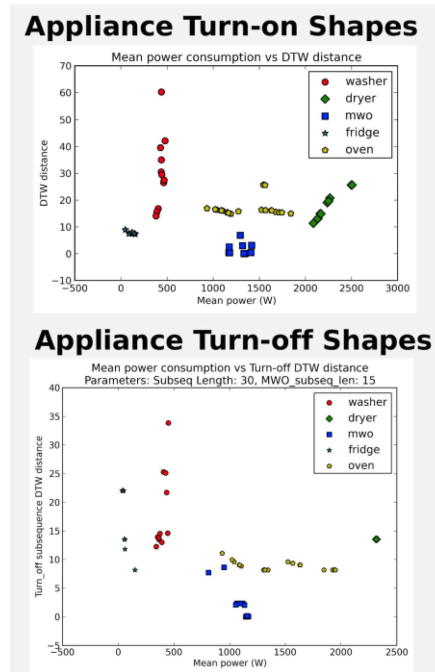

Figure 5. Appliance clustering via DTW distance and mean power consumption

V. DISAGGREGATION

The process of disaggregating whole-home data is divided into five steps which we discuss below.

A. *Event detection.*

This is the first step of any NILM algorithm. The power measurements obtained from the smart meter are analyzed to detect events which correspond to appliance state changes. We apply a median filter with a window length of 6 samples to smooth the data and reduce noise before performing event detection. We adopt an event detection scheme similar to that proposed by Hart [4], and divide power measurements into steady-state and transient periods. We use a second sliding window of 3 samples to classify time periods as steady-state and transient periods as we saw in figure 3. A steady-state is defined as any collection of sequences in which the standard deviation in a given sliding window of power measurements does not exceed a detection threshold Θ, and a transient is any event between two steady-states. This approach allows us to discriminate between a large number of appliances, while providing adequate resolution for detecting most of the significant loads in typical households.

Using the event detection approach we described earlier, we process the time series and extract all transient events to perform appliance detection. These transients correspond to appliance turn-on/off events, and allow us to distinguish between overlapping devices as long as the appliance state transitions do not occur at exactly the same time.

B. *Feature Extraction.*

Unlike [12, 14], where the step change in power is used as a feature, we calculate the statistics of each detected transient (mean, standard deviation, skew and kurtosis), include the step change in power and feed them into a classifier.

C. *Classification and Disaggregation.*

Each detected event is classified as a turn-on or turn off by means of its sign value. We then calculate the DTW distance between the detected event, and all the other events in our training set. The resulting DTW distance is combined with the transient statistics to create a feature vector. This feature vector is fed into a one-vs all k-nearest neighbor classifier or a Random Forest classifier from the Scikit-Learn Machine Learning library [15]. The use of one-vs-all classification ensures that the classifier doesn't try and force unknown appliances into one of the known appliance categories.

Disaggregation is performed by matching turn-on and turn-off events, then estimating the power consumption of between these two events. The first step is subtracting the offset or baseline from the transients in order to obtain the actual power draw of the appliance. We calculate p_{est}, the estimated power draw for the time interval between the turn-on and turn-off by averaging the last three samples of the turn-on transient and the first 3 samples of the turn-off transient and multiplying it by the time interval.

A naive approach to event matching would be to match a turn-on event with the subsequent turn-off event for that device type, but this approach is not robust to missed detections (a missed detection may result in two consecutive turn-on or turn-off events). We therefore developed a robust event matching algorithm for correctly matching turn-on and turn-off events. Our algorithm assumes that adjacent turn-on and turn-off events should be matched only if the time interval between them is reasonable. This prevents erroneous event matching due to missed events, and unrealistic appliance event durations like a microwave being used for 6 hours. The maximum duration for appliance events was calculated from data for homes 2-6 and tested on home 1.

VI. PERFORMANCE EVALUATION

Our NILM scheme was trained on Houses 2-6 and evaluated on House 1 of the REDD data set. We evaluated our detection and disaggregation scheme on an artificially generated aggregate signal shown in figure 7. This signal was obtained by summing power measurements for 4 appliances (a microwave oven, a refrigerator, a washer and dryer) over a six day period. Our transient event detector detected 329 events and fed these events into a k-Nearest Neighbor classifier using 15-nearest neighbors. The value of k was selected using cross-validation on a separate training set. The classified events were then fed into our event matching algorithm which was able to match 259 of them into pairs of turn-on and turn-off events. The 70 unpaired events were handled by our unmatched event algorithm. Without any tuning our approach correctly assigned 94.7% of the total power consumption to each appliance. This confirms the usefulness of our approach, and shows that it generalizes to unseen homes, thereby addressing an open issue with many NILM algorithms.

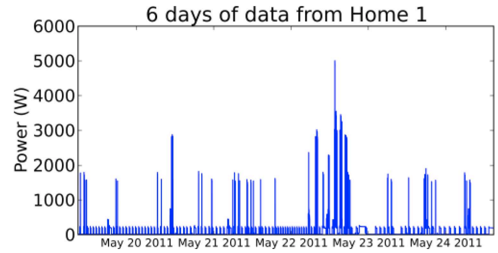

Figure 6. REDD Home 1 disaggregation test data

TABLE I. DISAGGREGATION PERFORMANCE

K-Nearest Neighbors	*Appliance*	*Actual usage (kWh)*	*Actual usage (kWh)*	*Accuracy (%)*
K=15	Washer	1.06	1.11	95.67
	Dryer	6.54	6.63	98.74
	Microwave	2.96	2.21	74.74
	Refrigerator	8.77	8.36	95.27
Summary		**19.34**	**18.30**	**94.65**

We vary the detection threshold Θ between 10 - 100W,and evaluate the effect on disaggregation accuracy. Θ is on the x-axis, and represents the threshold above which we decide an appliance event has occurred, while the y-axis is the percentage difference between the actual appliance energy consumption and what our disaggregation algorithm estimated it to be. The y-axis extends beyond 100% because we are evaluating disaggregation not detection error, and this can be several orders of magnitude greater the actual energy use. As can be seen in Figure 8(a), the optimal value of Θ is 30W, and varying the threshold does not affect all appliances in the same way. The effect of the threshold is a function of the magnitude of the steady state power draw of the appliance, and the variance the power changes. The impact of varying Θ is inversely proportional to the power draw of the appliance - the greater the average power draw, the smaller the impact of varying Θ. This is most evident in appliances like the Dryer and Microwave oven.

Varying Θ has little or no effect on disaggregation accuracy of the microwave because it quickly ramps up to its steady state, and has a relatively flat profile of 1500W so varying the threshold does not impact its detection. On the other hand, the washer is the appliance that is most affected by the varying the detection threshold. The reason why is obvious once we examine the typical cycle of the washer in house 1 in figure 8(b). We see that the washer has no steady-state, and the variance of the turn-on and turn-off events is very high, which makes it sensitive to variations in the detection threshold. Disaggregation accuracy for the refrigerator decreases as Θ approaches the average steady-state power of this device (~100W), which means that refrigerator events are more likely to be missed completely.

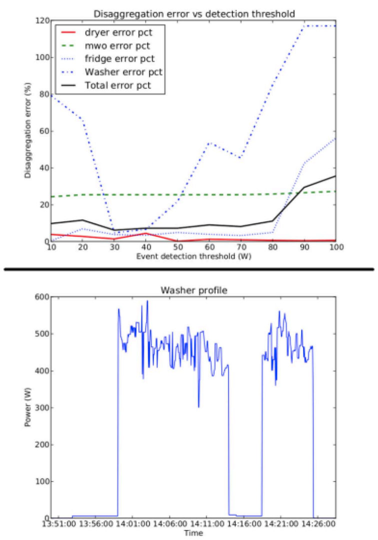

Figure 7. (a)Disaggregation error vs event detection threshold (b) REDD house 1 washer profile

A. *Choice of Classification algorithm*

The choice of classification algorithm is a choice between speed and accuracy. Our approach is geared towards online disaggregation, and using k-nearest neighbor classifier provides disaggregation accuracy of up to 95% and does so at high speed which makes it ideal for online disaggregation. Using a Random Forest classifier increases the accuracy to 99% but the classification time is ten times greater than the duration taken by the k-NN classifier. Our goal is online disaggregation on low-cost hardware so using a k-NN classifier will enable the deployment of our classifier on single-board computers such as the Raspberry Pi with low(er) overhead.

TABLE II. DISAGGREGATION ERROR COMPARISON

Appliance	***Our approach (%)***	***Hidden Markov Model approach [7](%)***
Dryer	4.73	15.13
Microwave	25.26	3.95
Refrigerator	1.26	33.99

From table 2 it is evident that our results compare favorably with the hidden Markov model approach developed in [7] but there are provisos. We cannot compare directly with Parson at al because we utilize a sampling rate of 3 seconds while they down sampled to 1 min resolution which is the sampling rate available with smart meters in the UK. Our results are better than theirs on the refrigerator and dryer, but their microwave disaggregation accuracy is significant better. We believe that tuning our generic models to the appliances in specific home will enable us to provide equivalent performance on the microwave.

VII. CONCLUSIONS

We developed a low-frequency transient-based approach for automatically identifying and disaggregating appliances from within an aggregate signal. Our approach leverages the unique shape of appliance transients, and is robust to missed turn on or turn off events. Our simulation results demonstrate the effectiveness of our scheme, and we plan to extend this work by adding the ability to tune the generic appliance models to the appliances within a given home via a semi-supervised learning approach. We are about to deploy our algorithm on Raspberry Pi's in 7 homes in the Midwest of the US and will report the performance of our algorithm in the field.

REFERENCES

[1] M. Hazas, A. Friday, and J. Scott. Look back before leaping forward: Four decades of domestic energyinquiry. IEEE Pervasive Computing, 10(1):13 –19,March 2011.

[2] K. Carrie Armel, A. Gupta, G. Shrimali, and A. Albert, "Is disaggregation the holy grail of energy efficiency? The case of electricity," *Energy Policy*, vol. 52, pp. 213–234, Jan. 2013.

[3] J. Froehlich, E. Larson, S. Gupta, G. Cohn, M. Reynolds, and S. Patel. Disaggregated end-use energy sensing for the smart grid. Pervasive Computing, IEEE, 10(1):28–39, 2011.

[4] G.W. Hart. Nonintrusive appliance load monitoring. Proceedings of the IEEE, 80(12):1870–1891, 1992.

[5] M. Zeifman and K. Roth, "Nonintrusive appliance load monitoring: Review and outlook," IEEE Transactions on Consumer Electronics, vol. 57, no. 1, pp. 76–84, Feb. 2011.

[6] A. Zoha, A. Gluhak, M. Imran, and S. Rajasegarar, "Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey," Sensors, vol. 12, no. 12, pp. 16838–16866, Dec. 2012.

[7] Oliver Parson, S Ghosh, M Weal, and A Rogers. Using hidden markov models for iterative non-intrusive appliance monitoring. In Neural Information Processing Systems, workshop on Machine Learning for Sustainability., Sierra Nevada, Spain, 2011.

[8] Hyungsul Kim, Manish Marwah, Martin F. Arlitt, Geoff Lyon, and Jiawei Han. Unsupervised disaggregation of low frequency power measurements. In SDM, pages 747–758. SIAM, 2011.

[9] Mario Berges, Ethan Goldman, H. Scott Matthews, Lucio Soibelman, and Kyle Anderson. User-centered nonintrusive electricity load monitoring for residential buildings. Journal of Computing in Civil Engineering, 25(6):471–480, May 2011

[10] Huijuan Shao, Manish Marwah, and Naren Ramakrishnan. A temporal motif mining approach to unsupervised energy disaggregation: Applications to residential and commercial buildings. In Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence, June 2013.

[11] Sidhant Gupta, Matthew S Reynolds, and Shwetak N Patel. ElectriSense: single-point sensing using EMI for electrical event detection and classification in the home. In Proceedings of the 12th ACM international conference on Ubiquitous computing, Ubicomp '10, page 139148, New York, NY, USA, 2010.

[12] J. Zico Kolter and M.J Johnson. REDD: a public data set for energy disaggregation research. In Proceedings of the SustKDD workshop on Data Mining Applications in Sustainability. 2011, San Diego, California, USA, 2011.

[13] D Berndt and James Clifford. Using dynamic time warping to find patterns in time series. In KDD workshop, volume 10, page 359370, 1994.

[14] Kyle Anderson, Adrian Ocneanu, Diego Benitez, Derrick Carlson, Anthony Rowe, and Mario Berges. BLUED: a fully labeled public dataset for event-based non-intrusive load monitoring research. In Proceedings of the 2nd KDD Workshop on Data Mining Applica- tions in Sustainability (SustKDD), Beijing, China, August 2012.

[15] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer,R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in python. Journal of Machine Learning Research, 12:28252830, 2011.