NOVEL TECHNIQUES FOR THE DETECTION OF *ON* AND *OFF* STATES OF APPLIANCES FOR POWER ESTIMATION IN NON-INTRUSIVE LOAD MONITORING

*Suman Giri[1], Po-Hsiang Lai[2], Mario Bergés[1]

[1]Department of Civil and Environmental Engineering,
5000 Forbes Avenue,
Carnegie Mellon University,
Pittsburgh PA, USA 15213
(Corresponding author: sgiri@andrew.cmu.edu)

[2]Samsung Telecommunications America,
1301 E Lookout Drive,
Richardson, TX, USA 75082

ABSTRACT

Non-Intrusive Load Monitoring (NILM) is a method of extracting appliance-level power consumption information from aggregate circuit-level data with the goal of giving users feedback regarding their energy consumption so they can take control of their consumption habits. In this paper, we present a novel algorithm for classification of *on* and *off* states of appliances. We compare the performance of our algorithm in *on* state detection with a pervious paper that evaluated the same dataset and show that it performs up to 13% better. We also present the results of a case study where we collected data for different modes of a cooktop, microwave and dishwasher and used our algorithms to perform power estimation. The error on ten different setups in the test bed ranges from 1% to 32%. We discuss our results and lay out ideas for future work.

KEYWORDS

Non-Intrusive Load Monitoring, Energy Disaggregation, Energy Efficiency, Machine Learning, Power Estimation, Normalization.

INTRODUCTION

Non-Intrusive Load Monitoring (NILM) is a technique that aims to provide energy end users with feedback regarding their appliance-level energy consumption habits without having to monitor each appliance individually. This idea has been around since the early 1980s when George Hart and his colleagues filed it as a US patent (Hart, Kern, & Schweppe, 1989). They proposed to achieve this by monitoring the main circuit that feeds current and voltage into a house, and based on the changes in electrical power signatures there, establish an estimate of what and when each appliance goes *on* and *off*. Then the end user is given an energy report that details what appliance in the house consumed what portion of their total energy use. Research has shown that feedback can motivate energy savings of up to 20% (Darby, 2006). This is important because electricity, on its own, constitutes 41% of total annual energy consumption in the US, and 67% of that is produced from fossil fuels (United States Energy Information Administration, 2009). Hence, savings due to NILM based feedback methods, in terms of both monetary units and impact on environment, quickly add up to large amounts when nationwide or even global figures are considered- validating its importance and the need for its adoption.

The idea of energy monitoring at the main circuit level has already penetrated the mainstream commercial market with devices like The Energy Detective[1], Envi[2], Efergy[3] etc. Over the years, various improvements have been made to Hart's original idea by integrating machine learning and signal processing techniques (Zeifman & Roth, 2011). Pattern recognition methods (Farinaccio & Zmeureanu, 1999), the use of higher order harmonics (Laughman et al., 2003), and features from the raw current waveforms (Ito et al., 2004) are among the few techniques implemented with some level of success. Researchers have also used fundamental decomposition methods such as principal component analysis (PCA) on power signatures (Kao, Cho, Lee, Toyomura, & Yamazaki, 2009). Algorithms like support vector machines (SVM) (Kao et al., 2009) and neural networks (NN) (Prudenzi, 2002) have also been tried. Recently, unsupervised methods mostly based on graphical models such as Bayesian networks (Lin, Lee,

---

[1] http://www.theenergydetective.com/
[2] http://www.currentcost.net/
[3] http://www.efergy.us/

Hsun, & Jih, 2010), Hidden Markov Models (Kim, Marwah, Arlitt, Lyon, & Han, 2011) and others have gained traction. Electric noise in the voltage line generated by mechanical switches (Patel, Robertson, Kientz, Reynolds, & Abowd, 2007) and electromagnetic interference (EMI) noise resulting from switch mode power-supplies (SMPS) (Gupta, Reynolds, & Patel, 2010) have also been studied as possible features for device classification.

To date, no NILM solution suitable for all types of household appliances is available. Zieffman and Roth (2011) note in their review of NILM methods that most of research effort in this area has been focused on signature exploration (also known as feature extraction). As a consequence, other aspects of NILM like power estimation and classification algorithm development have not been studied in detail. Even among the features that get proposed, due to lack of a standard dataset and accuracy metrics, there is no sense of how well the features generalize to other sets of appliances. Also, very few of the available solutions have been developed with the practicality of implementation in mind.

To fill this void, this paper will leverage a simple and easy-to-implement algorithm to obtain state change information in appliances. The contributions of this paper are twofold. First, we present a novel feature extraction technique that processes *on* transients (spikes observed when an appliance turns *on*) to extract the meaningful features. We evaluate the performance of our feature extraction method on the dataset used by Bergés et al. and compare their results to ours (Berges, Goldman, Matthews, Soibelman, & Anderson, 2011). Secondly, we also present a novel algorithm for *off* state detection that relies on steady state (state after the transient when the device is operating) current information. To evaluate the performance of our *off* state detection method, we present results of power estimation (done by using *off* state detection) on a case study where we collected data for 8 different modes of three appliances.

A FRAMEWORK FOR POWER ESTIMATION

The ultimate goal of NILM algorithms is to estimate the power consumed by all the devices being monitored. To achieve this, we created a framework that comprises of four distinct steps typically involved in a NILM setup. In this section, we summarize our methods and proposed algorithms for all of these steps.

Event Detection

Event Detection involves identifying when an event of interest (device turning *on* or *off,* for instance) occurs in the circuit. The algorithm we propose for this is based on step changes in real power. The step changes are monitored every sample. If they are above (or below) a certain threshold (50 Watts, for instance), they are labeled as an event of interest. After this we check to see if the average energy for two seconds after the event is at least 50% of the event. If so, the event is labeled as *on.* In the case of *off* events, we check if the average energy for two seconds before the event is at least 80% of the event. The numbers are chosen empirically given the nature of typical *on* and *off* transients, where *on* transients usually start with a few samples long spike followed by a dip, while *off* transients just fall off steeply. Figure 1 shows the results of our event detection algorithm on a combined operation of washer and microwave.

Feature Extraction

Once event detection is done, necessary features from the event should be extracted in

order to identify what device caused it. Our algorithm extracts a three second long transient for each event. Based on the direction of the event (positive or negative), it is labeled as *on* or *off*. The *on* transients are then *normalized* with the goal of highlighting the features that are characteristic to it. Most *on* transients of appliances consist of a similar large energy portion (roughly rectangular) on top of which small appliance and mode related variations reside. The goal of our *normalization* is to reduce the common shape portion and enhance the small variations for classification. The effects of signature normalization are illustrated in Figure 2.
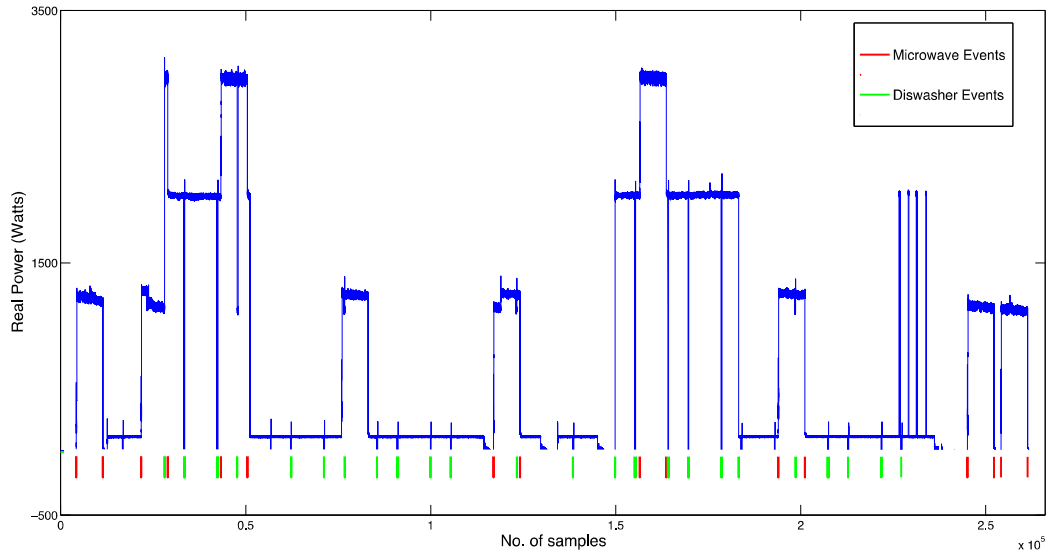

Figure 1- Event detection on aggregate signal of microwave and washer. The red marks indicate a microwave event while the green marks indicate a washer event.

The *normalization* is done in the following steps:
1. Smoothing: The resulting transient from the event detection is smoothened by a moving average of 10 milliseconds. This is achieved by convolving it with a 5 milliseconds long unit rectangular pulse. Then, the 5 milliseconds portions at the beginning and at the end of the resulting signal are removed, as these portions mostly contain artifacts from the smoothing.
2. Threshold selection: For each smoothened transient, the root mean squares (RMS) of the real power values are sorted. The values at the lowest 5 percent are then selected as a threshold.
3. Nonnegative threshold normalization: The resulting transient is normalized point-wise minus this threshold, and any resulting negative value is changed to 0.
4. Normalization by maximum: The resulting transient from step 3 is then normalized by its maximum RMS value. The *normalized* transients are then taken as features for *on* events.

For *off* events, feature extraction is different because *off* transients look the same for all devices. So, current signatures are used to extract the *off* event features. For any *off* event detected, we check for the current signatures right before and after the event. The current waveforms are then aligned according to their phase relative to the voltage. This was done by finding the zero-crossing (point where the sinusoidal voltage waveform is zero) of the voltage waveform before and after the event and using the difference to shift and align the current waveform accordingly. The aligned waveforms are then subtracted to find the current waveform of the appliance that went *off*. Then, one period of that waveform is extracted and used as feature for classification of that appliance. Figure 3 details this process.

Device Classification

Once the necessary *on* and *off* features are extracted, a nearest neighbor algorithm is implemented for classification. The algorithm was chosen for its performance and efficiency, keeping in mind the feasibility of using it online and in a practical setting. The algorithm works by calculating the distance (Euclidean, in this case) between the extracted test transient and all the data in the training repository. It then classifies the test sample as whichever label it is closest to. The training data is collected by extracting aforementioned features *(on* transients and current waveforms) from similar devices in a standalone setting.
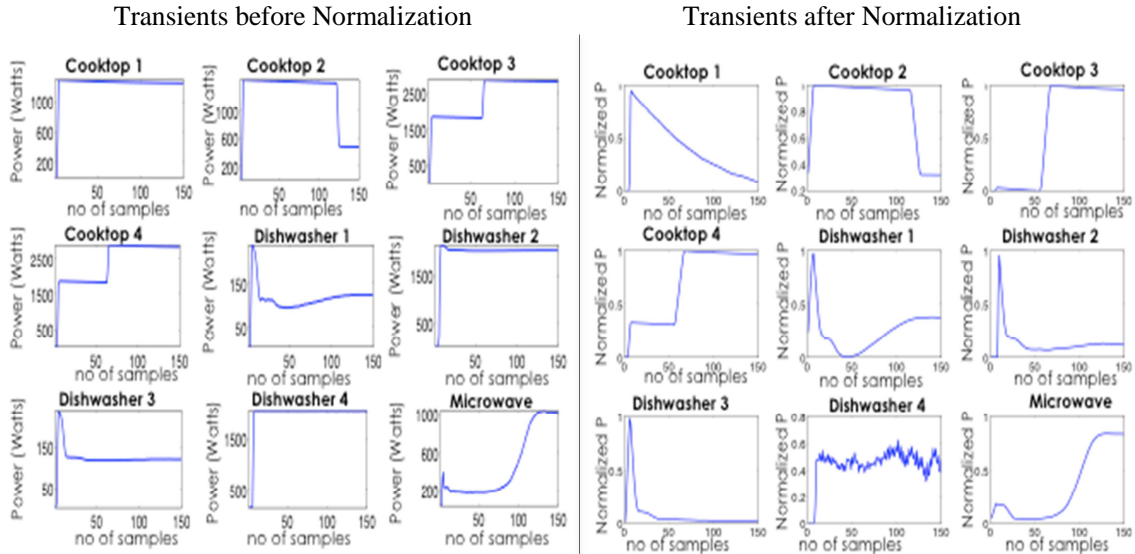

Figure 2- Effects of *normalization* on extracted transients. The left plots show the transients before normalization, while the ones on the right show the same transients after normalization. For instance, the transient labeled Cooktop1 on the right is the transient of mode 1 of Cooktop from the left after *normalization*, and so on. Normalized P stands for Normalized Power.

The same is done for steady state current signatures of each mode. We wait a few seconds after the state change of a device for the transient to become steady and take that as the steady state. This is done to identify what device went *off*. Figure 4 shows results of device classification on the power signal shown in Figure 3.

Power Estimation

Our power estimation algorithm leverages step changes in power as well as steady state current information to perform power estimation. We track all the devices that are *on* at any given instance based on the results of our classifier. Anytime there is an *off* event, we first check to see if the *off* drop is within 10% magnitude of any of the *on* transients at that instance. If such an *on* is found then those *on* and *off* transients are grouped together, and power is calculated based on an extrapolation of their average value between those time periods. If such an *on* is not found (for various reasons including a very sharp spike in the *on* transient), we extract a period of the current waveform of the device that went *off*, and run it against the nearest neighbor classifier.
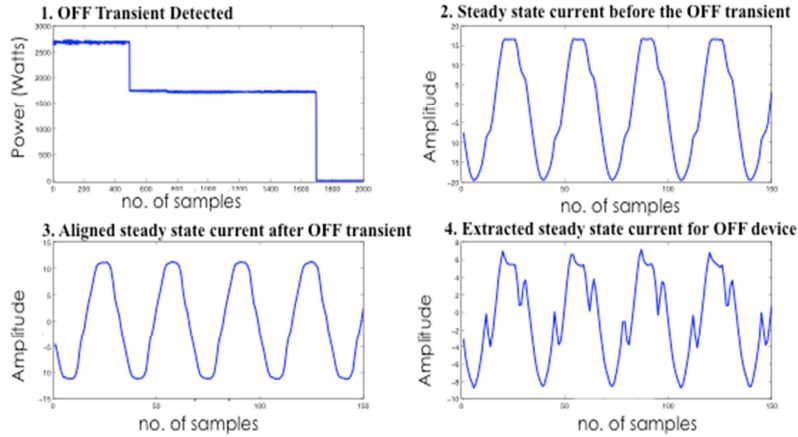
Figure 3- Steps involved in extracting the current waveform of the device that changed state from *on* to *off* after an *off* transient is detected.
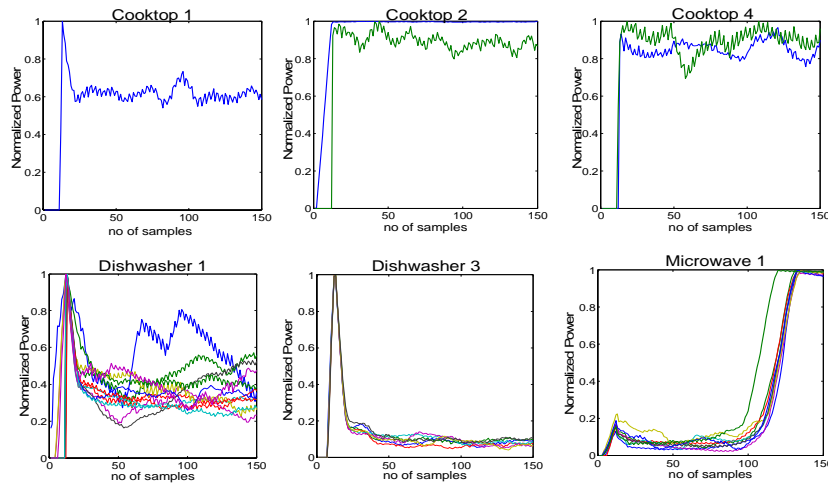

Figure 4- All the transients (extracted after event detection in Figure 4) that were labeled as one of the eight possible categories after Device Classification. Dishwasher 1, Dishwasher 3, etc. refer to different modes of the dishwasher. The transients in a particular category are overlaid on top of each other.

EVALUATION

Evaluation of *on* transient feature extraction technique

This section summarizes the performance of our algorithms on real-world data. To evaluate the robustness of our *on* transient feature extraction algorithm, we tested it on the dataset described by Bergés et al. (2011). Of the three datasets presented in the paper, we explored the dataset that includes data from a house (with 17 appliances and 176 *on* events) and the dataset that includes data from an apartment (with 14 appliances and 35 *on* events). These datasets included current and voltage readings from the main circuit panel sampled at a frequency of 15 KHz using a split core current transformer and a 1/100 voltage attenuator similar to the one described in (Anderson et al., 2012). We tested our feature extraction algorithm with theirs by using the same classifier (nearest neighbor) on the extracted features. Table 1 summarizes the

results. Our features (which are only dependent on real power *on* transients) outperformed features described in the paper (which included regression coefficients from both real and reactive power). As a benchmark, we also tested the performance of nearest neighbor algorithm on *non-normalized on* transients.

Table 1- Comparison of the performance of our feature extraction techniques for *on* transients with feature extraction techniques described by Bergés et al. (2011). The datasets were divided into training and testing so that there was at least one training sample and one testing sample for all events.

	# of events		Accuracy of nearest neighbor classifier		
	Train set	*Test set*	*No normalization*	*After normalization*	*Regression coefficients*
Dataset 1 (Whole-House)	90	86	76.4%	82.25%	76.4%
Dataset 2 (Apartment)	17	18	52.2%	65.56%	56.67%
Dataset 1 and Dataset 2	107	104	56.1%	65.4%	58.4%

Evaluation of *off* event feature extraction technique

The dataset used in the previous section does not contain information about steady state current and hence cannot be used for evaluation of our *off* event feature extraction technique. So, we conducted a case study with three devices– microwave, cooktop (electric stove) and washer (dishwasher) to test our framework of power estimation as a whole. The devices were selected because of the following reasons: (i) the power consumption is roughly the same for a microwave and cooktop which forces our algorithm to look for other features than just power differences; (ii) the transients for certain washer modes are almost the same as certain cooktop modes which forces us to look for other characteristic features; (iii) the microwave has a slow moving transient, while the washer has multiple modes, which allows us to tackle the problem of power estimation in multiple modes and under different conditions; and (iv) these are the major electricity consuming appliances in a regular household kitchen.

The data collection setup included an analog-to-digital converter that sampled both voltage and current at 2 kHz, measured on a power strip to which all of the above appliances were connected. A 1/100 voltage attenuator and a split-core current transformer, similar to the one used in (Anderson et al., 2012) were used to obtain the analog signals. Data was also collected for cases when two appliances were operating at the same time. We trained our algorithms on individual appliances and built a small repository of signatures. Then we tested the aggregate signals (with multiple appliances) on our training set. Figure 2 shows some of the transients for each of the appliances in our case study and Figure 1 shows the aggregate signatures for a case when two appliances were operating.

A training repository was built with 504 *on* transient signatures for all possible modes of the three devices. There were eight distinct modes for the three devices. We averaged all the transient signatures in a particular mode to get a median signature. This way there were eight signatures, one for each mode, in our repository. This significantly reduced the complexity of the nearest neighbor algorithm without sacrificing the performance.

Table 2- Percentage error in power estimation and number of events (*on* and *off*) in the test sample. The training was done on 504 *on* transients collected on individual appliances. Microwave and Cooktop-1 refers to one trail of aggregate signature where both the appliances were operating simultaneously, and so on.

Devices in test sample	Number of events	Error in Power estimation
Dishwasher	68	12%
Microwave	4	6%
Microwave and Cooktop- 1	8	4%
Microwave and Cooktop- 2	8	1%
Microwave and Cooktop- 3	8	5%
Microwave and Cooktop- 4	8	1%
Microwave and Dishwasher- 1	78	16%
Microwave and Dishwasher- 2	90	9%
Microwave and Dishwasher- 3	90	7%
Microwave and Dishwasher- 4	96	32%

DISCUSSION

The feature extraction technique that we propose for *on* transients performs better than more complex methods that utilize additional information as was shown in the previous section. To test for its robustness, we compared its performance with *non-normalized* transients in the data from our case study, and saw accuracy improvements of up to 40% upon *normalization*. The low accuracy values in Table 1 should not be taken at face value because event detection, centering of extracted transients and ground truth labeling were not done using our algorithms. In a controlled setting where had control over the detection, centering and training, our feature extraction method had accuracies of up to 98%.

The results from Table 2 reflect the effectiveness our *off* transient detection and power estimation algorithms. The power estimation results for Microwave and Dishwasher-4 are not on par with the rest of the results because some of the dishwasher modes were classified as cooktops modes-as their power consumption and transient shapes both look similar. Features like reactive power or higher order harmonics specific to a dishwasher could be added to the model to correct for this in expenses of extra complexity.

The error in power estimation was calculated by comparing the total power consumed by the devices (Ground Truth) with the power consumed by devices after classification. This model of error estimation may be argued to be a lower bound in conveying the total picture, as over and under power estimate errors at event level may cancel out resulting in smaller total error. The only way to test for this is to verify the model over a large sample of data for consistency. Yet, for the test samples in the case study, values were largely consistent and the larger errors were borne out of misclassification between appliance types. We leave development of a better metric for evaluating power estimation error and expansion of the case study to include more appliance types as future work.

CONCLUSION

In this paper we outlined a brief history of Non-Intrusive Load Monitoring (NILM) techniques and presented a framework for estimating power at appliance level. As our contributions, we developed an algorithm that utilizes novel *normalization* techniques for *on* transient detection, and compared its performance against a standard dataset. We found that it

outperformed *non-normalized* data by up to 13%. We also presented a method that looks at steady state current signatures for *off* event classification to perform power estimation. We reported on the results of using our algorithm on the data from our case study. The average error on power estimation was around 9%.

References:

Anderson, K., Ocneanu, A., Benitez, D., Carlson, D., Rowe, A., & Berges, M. (2012). BLUED: A Fully Labeled Public Dataset for {Event-Based} {Non-Intrusive} Load Monitoring Research. Presented at the Proceedings of the 2nd {KDD} Workshop on Data Mining Applications in Sustainability SustKDD.

Berges, M., Goldman, E., Matthews, H., Soibelman, L., & Anderson, K. (2011). User-Centered Nonintrusive Electricity Load Monitoring for Residential Buildings. *Journal of Computing in Civil Engineering*, *25*(6), 471–480. doi:10.1061/(ASCE)CP.1943-5487.0000108

Darby, S. (2006). The effectiveness of Feedback on energy Consumption. *A Review for DEFRA of the Literature on Metering, Billing and direct Displays, April*.

Farinaccio, L., & Zmeureanu, R. (1999). Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses. *Energy and Buildings*, *30*(3), 245–259. doi:10.1016/S0378-7788(99)00007-9

Gupta, S., Reynolds, M. S., & Patel, S. N. (2010). ElectriSense: single-point sensing using EMI for electrical event detection and classification in the home. *Proceedings of the 12th ACM international conference on Ubiquitous computing*, Ubicomp '10 (pp. 139–148). New York, NY

Hart, G. W., Kern, E. C., Jr., & Schweppe, F. C. (1989, August 15). Non-intrusive appliance monitor apparatus. Retrieved from http://www.google.com/patents/US4858141

Ito, M., Uda, R., Ichimura, S., Tago, K., Hoshi, T., & Matsushita, Y. (2004). A method of appliance detection based on features of power waveform. *Applications and the Internet, 2004. Proceedings. 2004 International Symposium on* (pp. 291–294).

Kao, T., Cho, H., Lee, D., Toyomura, T., & Yamazaki, T. (2009). Appliance recognition from electric current signals for information-energy integrated network in home environments. Presented at the International Conference on Smart Homes and Health Telematics, Tours, France.

Kim, H., Marwah, M., Arlitt, M., Lyon, G., & Han, J. (2011). Unsupervised Disaggregation of Low Frequency Power Measurements. Presented at the SIAM International Conference on Data Mining (SDM 11), Mesa, Arizona.

Laughman, C., Kwangduk Lee, Cox, R., Shaw, S., Leeb, S., Norford, L., & Armstrong, P. (2003). Power signature analysis. *Power and Energy Magazine, IEEE*, *1*(2), 56–63. doi:10.1109/MPAE.2003.1192027

Lin, G., Lee, S., Hsun, J. Y., & Jih, W. (2010). Applying power meters for appliance recognition on the electric panel. Presented at the Industrial Electronics and Applications (ICIEA), 2010 the 5th IEEE Conference on. doi:10.1109/ICIEA.2010.5515385

Patel, S., Robertson, T., Kientz, J., Reynolds, M., & Abowd, G. (2007). At the Flick of a Switch: Detecting and Classifying Unique Electrical Events on the Residential Power Line. *UbiComp 2007: Ubiquitous Computing* (pp. 271–288).

United States Energy Information Administration. (2009). *Annual Energy Review*. Washington, D.C. Retrieved from http://www.eia.gov/totalenergy/data/annual/archive/038409.pdf

Zeifman, M., & Roth, K. (2011). Nonintrusive appliance load monitoring: Review and outlook. *2011 IEEE International Conference on Consumer Electronics (ICCE)* (pp. 239 –240). Presented at the 2011 IEEE International Conference on Consumer Electronics (ICCE). doi:10.1109/ICCE.2011.5722560