ElectriSense: Single-Point Sensing Using EMI for Electrical Event Detection and Classification in the Home

Sidhant Gupta[1], Matthew S. Reynolds[2], Shwetak N. Patel[1]

[1]Computer Science & Engineering,
Electrical Engineering,
UbiComp Lab
University of Washington, Seattle, WA 98195
{sidhant, shwetak}@uw.edu

[2]Electrical and Computer Engineering
Duke University
Durham, NC, 27708
matt.reynolds@duke.edu

ABSTRACT

This paper presents ElectriSense, a new solution for automatically detecting and classifying the use of electronic devices in a home from a single point of sensing. ElectriSense relies on the fact that most modern consumer electronics and fluorescent lighting employ switch mode power supplies (SMPS) to achieve high efficiency. These power supplies continuously generate high frequency electromagnetic interference (EMI) during operation that propagates throughout a home's power wiring. We show both analytically and by in-home experimentation that EMI signals are stable and predictable based on the device's switching frequency characteristics. Unlike past transient noise-based solutions, this new approach provides the ability for EMI signatures to be applicable across homes while still being able to differentiate between similar devices in a home. We have evaluated our solution in seven homes, including one six-month deployment. Our results show that ElectriSense can identify and classify the usage of individual devices with a mean accuracy of 93.82%.

Author Keywords

Infrastructure-mediated sensing, activity sensing, activity recognition, energy monitoring

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms

Algorithms, Experimentation, Measurement.

INTRODUCTION

Low-cost and easy-to-install methods to sense and model human activity in the home have long been a central focus of UbiComp research. Sensing disaggregated electricity usage in the home (i.e., at the individual source level) has emerged as one particularly promising area for activity-inference research because of how often it reveals the resident's current activity (e.g., stove usage implies cooking). In previous work, Patel et al. [15] introduced a single-point sensing method for automatically identifying the usage of resistive and inductive electrical loads. In this work, we focus on sensing a different class of devices, namely switched mode power supplies (SMPS), which are used in most modern consumer electronics as well as in fluorescent lighting.

Devices that rely on SMPS have become increasingly prevalent because of their higher efficiency, smaller size, and lower cost compared to traditional power supplies. Manufacturers are increasingly employing SMPS in their products to meet minimum efficiency requirements (e.g., the Department of Energy's Energy Star program [4]). For example, in one of our deployment sites, all lights and most appliances used SMPS. In contrast to Patel et al.'s approach, which sensed voltage transients from the activation/deactivation of resistive and inductive electrical loads, we examine the *continuous* EMI signals generated by *all* SMPS based devices due to their reliance on asynchronous high frequency oscillators for operation.

In particular, our system, called ElectriSense, senses the electromagnetic interference (EMI) created by SMPS oscillators. We have found through experimentation that the EMI generated by SMPS has a highly repeatable frequency-domain signature that can be sensed and identified throughout a typical home during device operation. Perhaps more importantly, we have found that these signatures are largely specific to a device's circuit design and maintain consistent properties across homes. This is in contrast to much of the work in infrastructure-mediated sensing [3,6,15,16,17], which involve sensing techniques that require *per-home* calibration. Moreover, because we rely on a *continuous* noise signature, we can identify devices that do not generate transients such as those with "soft switches" and transient suppressors. Finally, given that Patel et al. [15] and ElectriSense actually sense different classes of devices, they are complementary rather than competing techniques.

Through experimental trials in seven homes and one six-month long deployment we show that ElectriSense can correctly identify and classify SMPS electrical events, reasonably well down to the individual source level (*e.g.*, a particular TV, Laptop Charger, or a CFL Lamp). We also show that ElectriSense is robust to simultaneous device

UbiComp 2010, September 26–29, 2010, Copenhagen, Denmark.

activation and can reliably detect overlapping device activation events (one aspect of other infrastructure-mediated sensing approaches that is not often evaluated, e.g., [6,15]). Finally, we show that our approach works irrespective of a device's plug-in location in the home and also that we can distinguish between two devices of the same model and brand.

RELATED WORK

Distributed Sensing

Detecting and classifying electrical appliance activation has been done in the distributed model wherein each device has a dedicated sensor, which looks for device state change [14,18]. Device level sensing is conceptually straightforward but requires time consuming and expensive installation and maintenance. Indirect sensing techniques have also been used where researchers placed microphones, accelerometers and video cameras throughout a home to infer activity [2,5]. Such techniques are effective but require tedious installation and maintenance and may also raise privacy concerns in a home setting. A more recent technique for electrical event detection is to indirectly listen to the activation of switches and motors through microphones distributed through a living space [13].

Single-Point Sensing

Pioneering work from the 1980s in single point sensing of electrical events focused on the design of new in-line metering techniques for monitoring whole-house appliance usage [1,8,9]. Hart et al.'s approach used a current sensor installed by an electrician in-line with a home's power meter to monitor voltage and current waveforms for the incoming power to the home. Because of the limitations of low cost digital signal processors in the 1980s and 1990s, this work considered only line frequency at 60Hz and low order harmonics up to a few kHz. They relied on step change in power, the active and reactive power components of each appliance and the time of day statistics to detect and classify appliance usage. Additionally, most consumer electronic devices at that time did not employ SMPS, because of the immature state of SMPS technology and the absence of low cost, single chip SMPS implementations.

Infrastructure-Mediated Sensing

Recent advances in infrastructure-mediated sensing have provided an alternative approach to inferring in-home activity through the detection of events affecting the home's utility infrastructure [3,6,15,16,17]. Our approach falls within this class of sensing systems, but provide a significant advancement over prior work.

Previous IMS approaches for detecting electrical appliance activation from an ordinary outlet (without requiring in-line installation) by Patel et.al. [15] leveraged transients generated by mechanically switched incandescent, heating, and motor loads to detect and classify electrical events (continuous noise was only briefly mentioned). This work showed that transient noise is generated whenever a mechanical switch is turned on or off due to tiny arcs generated inside the switch, which excites the step response of the home's electrical wiring transfer function. Capturing and analyzing transient noise is a computationally expensive process, and requires that the system continuously capture and analyze every transient noise event for reliable detection. There is no way to coherently integrate these transient events to improve detection likelihood as they occur infrequently and are relatively weak due to their broadband distribution of signal energy.

Transients are an extrinsic property of mechanical switched loads, making their characteristics unpredictable and hence requiring supervised training for each physical device. Thus, signatures learned from one device cannot be applied to another even if they are similar. This also poses a challenge for identifying mobile devices, as the transients generated by a particular device change as the device is moved from one location to another. In contrast, our approach leverages signals that manifest from an engineered process, specifically the particular circuit design and the individual components a device uses, thus making the signal learned from one device applicable to other similar devices within and across homes. This eases the training process by allowing crowd-sourcing techniques to be applied.

Another key motivation for our work is the fact that most modern consumer electronic appliances are moving towards having a "soft switch", i.e. unlike a mechanical switch they make use of a software driven push button that cycles the power to the appliance electronically (for example via the infrared remote control of a flat-panel TV). In such devices, the indirect activation of the device by software driven electronic switch minimizes the transient generated at the moment of activation. We observed several devices, such as LCD monitors and DVD players that did not generate any detectable transients. Fortunately, these software driven devices are nearly always SMPS-based.

THEORY OF OPERATION

The electrical noise present on a power line when a device is operational is called conducted electro-magnetic interference (EMI), which can be classified into two types: transient and continuous. Transient noise is characterized by the short duration for which it can be observed, generally few tens of nanoseconds to a few milliseconds. Continuous noise on the other hand can be observed for as long as the device is operational. Both transient and continuous noise can either be concentrated within a narrow frequency band or spread over a wider bandwidth (also called broadband noise). A compact fluorescent light bulb (CFL) is an example of a device that generates continuous noise, which is conducted over the power line due to its physical contact with the power line. Since a home's electrical distribution system is interconnected in parallel at the home's circuit breaker panel, conducted EMI propagates widely from a given device throughout the electrical infrastructure.

Continuous noise is usually intrinsic to the device's operation and internal electronics. Appliances like grinders, fans and hair dryers that make use of a motor create voltage noise

synchronous to the frequency of AC power (60 Hz in the USA) and its harmonics (120Hz, 180Hz, etc.) due to the continuous making and breaking of electrical contact by the motor bushes. In contrast, modern SMPS based electronic devices generate noise that is synchronous to their power supply's internal oscillator.

In contrast to traditional linear power regulation, a SMPS does not dissipate excess power as heat, but instead stores energy in an inductance and switches this stored energy in from the line and out to the load as required, thus wasting much less energy. The key to a SMPS's smaller size and efficiency is its use of a power transistor to switch the stored energy at a high frequency, also known as the *switching frequency*. The switching frequency is much higher than the 60Hz AC line frequency because at higher frequencies the inductors or transformers required are much smaller [7]. Typical SMPS operate at tens to hundreds of kHz. The switching waveform is adjusted to match the power requirements of the appliance it is powering.

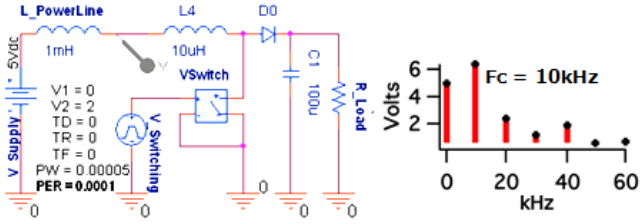

Figure 1: (Left) Circuit model of a SMPS with placement of the voltage probe. (Right) Frequency domain analysis at the voltage from probe showing EMI at 10 kHz.

A CFL's power supply employs the same fundamental switching mechanism to generate high voltages necessary to power the lamp. The switching action, which is the cornerstone of a SMPS's operating principle, generates a large amount of EMI centered in frequency around the switching frequency. This phenomenon can be understood by modeling a simple DC-DC SMPS circuit that uses the same fundamental switching topology (See Figure 1).

The large inductor *L_PowerLine* mimics the power line inductance. The SMPS is plugged into the power line. To measure the conducted EMI, we place a voltage probe *V* on the power line, which is analogous to having the single sensor plugged into the power line with a SMPS based device operational somewhere else. The switching frequency *fc* for the model is governed by the PER (period) parameter of the *V_Switching* component. We arbitrarily set it to 10 kHz. Figure 1 shows a frequency domain plot of the noise at probe, which clearly shows that the power supply emits EMI, which is conducted over to the power line and is most prominent at the switching frequency *fc* (10 kHz here) and its harmonics. This is the same behavior that we observe when a SMPS based appliance is turned on in the home.

In the US, the Federal Communications Commission (FCC) sets rules (47CFR part 15/18 Consumer Emission Limits) for any device that connects to the power line, which dictates the maximum amount of EMI a device can conduct back onto the power line. This limit is 66 dBuV for frequency range between 150 kHz to 500 kHz, which is nearly -40dBm across a 50 ohm load. The ElectriSense data acquisition system is sensitive enough to capture noise from -100 dBm to -10 dBm across a frequency range of 36kHz – 500kHz.

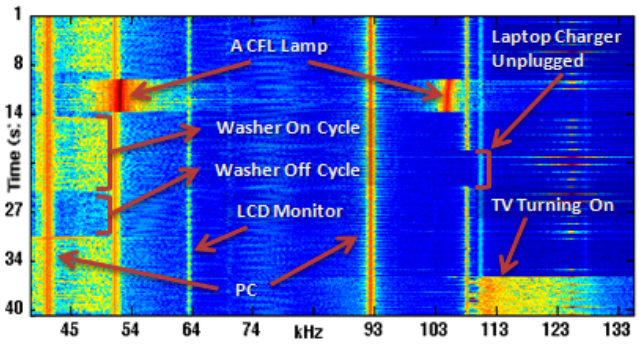

Figure 2: Frequency spectrogram showing device actuation in a home.

Figure 2 shows a frequency domain waterfall plot showing appliances being turned on and off. As is evident from the graph, when the device is turned on we see a narrowband continuous noise signature that lasts for the duration of the device's operation. Also note that the noise center is strongest in intensity and then extends to lower and higher frequencies with decaying intensity, which can loosely be modeled with a Gaussian function having its mean at the switching frequency. This behavior can be attributed to the error tolerance of the components that make up the switching circuit core, as well as the characteristics of the power supply's load. If all the components were ideal, we would see a single narrow signal peak at the switching frequency. The error tolerance of SMPS components also allows for distinction between otherwise identical devices, such as a variety of units of the same model of CFL bulb. Finally, the power line itself can be thought of as a transfer function (difference in the inductance between the sensing source and the appliance) and provide additional discrimination among multiple similar devices. We show this experimentally later in this paper.

Dimmers also produce continuous noise due to the triggering of their internal triac switches, which can be used to detect and identify incandescent loads they control. In contrast to the narrowband noise produced by SMPS, a dimmer produces broadband noise spanning hundreds of kHz, which could be modeled as a Gaussian having very large variance. A detailed treatment of dimmers and differentiating between identical devices is presented later.

SYSTEM DESCRIPTION AND OVERVIEW

Our prototype system consists of a single custom Power Line Interface (PLI) plug-in module that can be plugged into any ordinary electrical outlet in the home (Figure 3). The output of the plug-in module is connected to a high speed data acquisition system based on the Universal Software Radio Peripheral (USRP) that digitizes the analog signal from the plug-in module and streams it over a USB connection to a data collection and analysis PC running GNU Radio, which in real-time samples and conditions the incoming signal. The ElectriSense algorithms then watch for an event and extract features that are used to identify and classify the device causing the event.

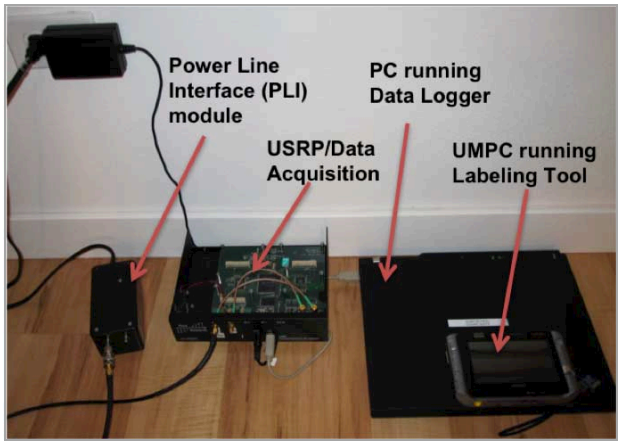

Figure 3: Our prototype system consists of a single plug-in module, acquisition hardware and the supporting software

Though we tested our system on a 120V, 60 Hz electrical infrastructure, our approach can easily be applied to electrical infrastructure utilizing different frequency and voltage rating with little change to the hardware and no change to the software. For homes that have split phase wiring (i.e. two 120 V branches that are 180-degrees out of phase), the coupling between the two phases allows us to monitor at a single location and capture events on both. The exception to this and further discussion is presented later in the paper.

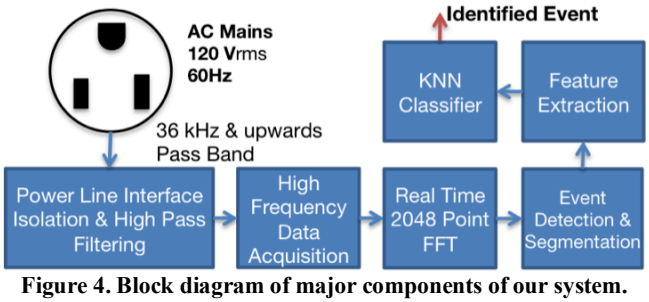

Figure 4. Block diagram of major components of our system.

IMPLEMENTATION DETAILS

In this section, we describe the various components of our prototype system (Figure 4) and present in detail the hardware design and software algorithms used for event detection, feature extraction and classification.

Hardware

To capture and analyze the electrical noise on the power line, we built a custom power line interface (PLI) module for our analog frontend. It is necessary for this module to filter out the AC line frequency (60 Hz in the U.S.) so that the spectrum analyzer or any analog-to-digital device is not overloaded by the strong 60Hz frequency component. The PLI consists of a high pass filter, which has an essentially flat frequency response from 50 kHz to 30 Mhz. The 3 dB corner is at 36.7 kHz giving us a wide enough band to look at the complete range of any conducted EMI (see Figure 5 for the schematic). The filter design also includes a 10 dB attenuator so that a constant 50-ohm load is presented at the input of the data acquisition hardware, irrespective of the signal frequency or the AC line conditions. For safety and isolation from the line voltage, high voltage capacitors are required. It should be noted that the polarity shown should be strictly followed, i.e., the line and neutral lines should not be connected in reverse and the isolation capacitors should be of AC-line rated polyester film type for safety.

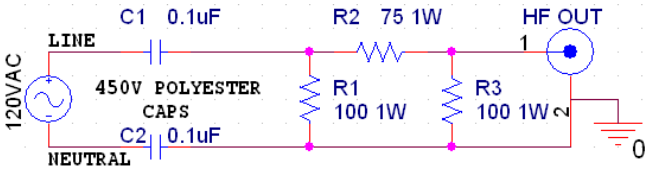

Figure 5: Schematic of the plug-in power line interface module.

The filtered signal is then fed into the USRP, which acts as a general purpose analog to digital converter sampling at a rate of 1 MHz set through the software. The digitized signal is streamed from the USRP to our processing software over USB.

Software

The incoming time domain signal stream from the USRP is buffered as 2048-point vectors and FFTs of these are computed to obtain the frequency domain signal. The 2048 points are spread equally over the spectral width of 500 kHz, which yields a resolution of 244 Hz per FFT bin. The FFT vector or frequency vector is computed 244 times per second, which is then fed into our event detection and extraction software.

We found that most SMPS devices generate noise peaks that are 8 dB to 60 dB above the baseline. The baseline noise in a home varies unpredictably between -90 dBm to -70 dBm across the entire spectrum with a period of few Hertz. Since the variability of the baseline noise is high, we must average the incoming frequency vector over time to obtain a stable baseline. We use a sliding window average with a window size of 25. Using a window that is too small results in an increase in the false positives whereas a large window size increases the shortest duration between near simultaneous events that is needed for the system to detect them as separate events.

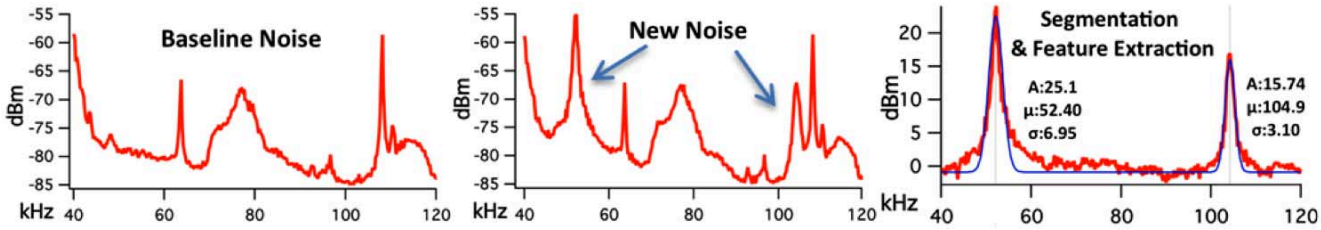

Figure 6: (Left) Background noise observed on a particular power line. (Center) A new device is turned on, producing EMI that introduces new signals to the power line. (Right) After background subtraction the new signal features are extracted. The resulting Gaussian fit and its features amplitude (A), mean (μ) and variance (σ) are also shown.

When the system first starts, it computes an average of 25 frequency vectors and stores it as the baseline noise signature (Figure 6 Left) Thereafter, a new window is calculated every 25 frequency vectors, and a difference vector with the baseline noise signature is computed. When a device is actuated and new noise in the frequency domain is introduced to the power line (Figure 6 Center), the difference vector reflects this change, thus segmenting the event (Figure 6 Right).

The event detection algorithm scans the difference vector to find any values that are greater than a predefined threshold. It should be noted that this is a global threshold that is set once and works across different homes. In our deployments, we found that 8 dB above the noise baseline was a sufficient power threshold. Since the vectors in a window are averaged, if the window only partially overlaps with an event, the detection algorithm may still correctly detect it, but the difference vector will reflect a smaller magnitude. To mitigate this problem, when an event is detected, a new difference vector is calculated using the next window cycle. After the detection of an event, the baseline noise vector is updated appropriately so as to reflect the new noise floor of the power line.

Our feature extraction algorithm finds peaks above the threshold using the difference vector and fits a Gaussian function to extract the mean, amplitude and variance parameters (Figure 6 Right). The change in amplitude can be positive or negative depending on whether the device is turned on or off. These noise signatures are always the inverse of each other for opposite state transitions.

A feature vector for the suspected event is then created using the parameters of the center frequency, which is generally the global maximum frequency component. Other peaks may also be present as harmonics. We use K-Nearest Neighbor (KNN) with K=1 and a Euclidean distance metric with inverse weighting, which is well suited for this kind of low-dimensionality data. These parameters were derived experimentally using a cross validation test over our entire dataset from seven different homes.

IN HOME DEPLOYMENTS AND EVALUATION

To validate our learning approach, we conducted experiments in seven different homes. We collected data from one house for a longer period, spanning six months and from multiple homes for a shorter period (spanning a single day). This allowed us to show the general applicability of our approach to a diverse set of homes as well as the long-term temporal stability of our sensing solution. Table 1 shows the summary of the homes used in our evaluation.

ID	Style/Built/ Remodeled	Size/Floors	No. of Test Devices	No. of Events
H1	Apartment/1985/ NA	750 sq. ft./ 1 flr.	10	135
H2	Single Family/2003/NA	3000 sq. ft./ 2 flrs.	15	203
H3	Single Family/1974/2009	1200 sq. ft./ 2 flrs. + basement	13	170
H4	Apartment/1910/ NA	450 sq. ft./ 1 flr.	7	108
H5	Single Family/1960/NA	1700 sq. ft./ 1 flr.	13	198
H6	Single Family/1926/2003	2800 sq. ft./ 2 flrs. + basement	20	404
H7 *	Apartment/2009/ NA	657 sq. ft./ 1 flr.	16	1358

Table 1: A summary of the homes showing the style, size, number of appliances we tested and the number of events (* Long-term 6-month deployment).

Data Collection Procedure

Our system was packaged such that it could be rapidly setup in a home. The laptop, data acquisition hardware and a wireless router were pre-configured, connected properly, and setup on a rolling cart. For each home, we picked at random an available electrical outlet that had two sockets and plugged the PLI module into one and used the other to power our laptop and the USRP. Our laptop and the USRP of course generate their own EMI but this noise is subtracted from the baseline. After the installation, we made a note of every appliance, electronic device, and light fixture that incorporated a switching power supply. This included incandescent lights that were driven by a dimmer switch in addition to any light fixtures with CFL bulbs. For dimmers, we only collected events at 0% and 100% dim levels, because of the challenge of accurately and repeatedly setting intermediate dim levels. More analysis on dimmers is presented later. The collected labels were then fed into our ground truth labeling software. We then went through the home in two phases.

In the first phase, we actuated each appliance on and off five to six times individually to ensure that we captured an

isolated signature for each appliance. Every time our system detected an event (i.e., when we turned on a device), the features were automatically extracted and sent to the labeling software's queue running on an ultra-mobile Sony Vaio UX computer (UMPC). As events were being captured, we could label those events using the UMPC. This process was carried out for all of the devices in the home.

For our second phase of data collection in each home, our goal was to collect data under a naturalistic setting allowing for overlapping and simultaneous events. To do so, (1) we asked the homeowner to perform certain activities (on an average three activities per home involving 2-4 devices each), such as watch TV, prepare a meal, etc. and (2) we actuated random appliances and/or logical groups of appliances such as turning on the DVD player, digital cable box, a gaming console and then a TV to simulate a *TV watching* experience through a universal remote. Events generated from these tasks were manually labeled.

In addition to the devices already found in each of the test homes, we collected data from a laptop power adaptor, two CFL lamps and a camera charger that we took to each one of the deployment sites. To simulate a naturalistic use of each device, we moved them around in the home and plugged them into any available outlet. This allowed us to analyze the stability of these noise signatures across different homes.

The data collected by the labeling software was time stamped, labeled and stored in an XML database. Since we collect more features than required by our current approach, the XML database allows us to easily filter and parse the data with help of an XQuery to generate output data in a format that can be directly imported and processed by the Weka machine learning toolkit.

Long-Term Deployment

To test the temporal stability and long-term feasibility of our approach, we deployed our prototype system at one of the homes for over six months. Events were collected and labeled throughout this entire period manually using a slight variant of the labeling tool that was used in other homes. Every time an event was detected, the logging tool would put the extracted features in a queue and generate an audible beep. The home occupants were instructed to label the events with the electrical appliance they actuated using the labeling tool.

The labeling tool was designed to have a highly streamlined interface, requiring only the selection of device labels from an onscreen selection list. If the user did not label an event in the queue for more than six minutes, it was labeled as unknown and purged from the queue. This feature allowed the home occupants to handle any erroneously detected events and ignore events when the occupants were unable to attend to the labeling tool promptly. Though a convenient feature, this also meant that we lost labels on actual events that the occupant missed or decided to ignore. However, the intent of this experiment was to gather as many ground truth labels as possible for an extended period of time.

RESULTS AND ANALYSIS OF IN-HOME EXPERIMENTS

In this section, we detail the datasets collected and present the results of our approach.

Data Collected

We collected a total of 2576 (Table 1) electrical events from seven homes. The largest number of events came from lighting, which tended to be either CFL- or dimmer-based. Most of the other detectable devices were common consumer electronic devices, such as LCD or LED TVs, gaming consoles, PCs, power adaptors, etc.

Appliances such as dryers and electric stoves did not appear to generate events in some of our test homes. Generally such devices are large resistive loads and hence do not emit any high frequency noise. We did observe continuous noise events from a washer in H7. Out of all our test homes only H7 had a modern Energy Star compliant high efficiency washer that generated continuous noise, which was from its electronically controlled DC motor powered through a SMPS. Also, for most large appliances the use of a SMPS is considered negligible compared to the overall power consumption for the appliance.

Classifying Specific Devices within the Homes

To determine the classification accuracy of electrical device actuations in a home, we evaluate our classification approach using two different procedures. In the first, we evaluate the performance of our KNN-based classifier using a 10-fold cross validation for each home. In the second evaluation, we use a minimal training set (a single training example for each device of interest) in order to simulate a more practical and real-world situation. Table 3 shows the overall accuracy for classifying devices within each of the seven homes. Using 10-fold cross validation, we observed an overall average accuracy of 91.75%.

Upon analyzing the confusion matrices for each of the homes, we found that for both H5 and H6, there was confusion between some of the lighting (see Figure 7 for the confusion matrix). From our data, we observed that our classifier correctly identified similar light fixtures (i.e, the same model and brand) that were located in different rooms. However, H5 and H6 had rooms where the same models of fluorescent light fixtures were installed spatially near each other (1-2 feet apart), which produced very similar noise signatures. Thus, these particular lights did not have sufficient differences nor were they far apart enough along the power line to differentiate between those lights.

Higher frequency resolution hardware may partially alleviate this problem, but with added costs. Also, for some applications, logically combining or clustering the lighting that are spatially co-located might be acceptable. Using this latter approach, Table 3 shows the performance of our classification scheme after clustering the same model of lights or devices that are 1-2 feet apart from each other as a

single event. This approach yields an increase in classification accuracy for H5 (92.4%) and H6 (91.8%) for an average accuracy of 93.82%.

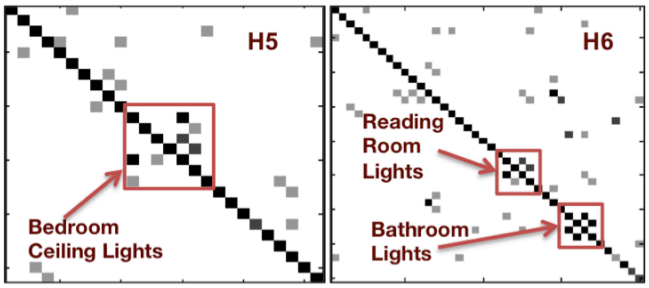

Figure 7: Visual confusion matrix highlighting misclassification due to physical proximity of similar fixtures in H5 and H6.

Home	10 Fold Cross Validation (%)		Minimal Training Set
	As is	Combined	
H1	96.29	96.29	93.46
H2	89.65	89.65	83.73
H3	96.47	96.47	93.52
H4	97.23	97.23	93.81
H5	85.35	92.42*	90.53
H6	84.41	91.83*	84.61
H7	92.85	92.85	85.13
Total	**91.75**	**93.82**	**89.25**

Table 3: The performance using 10-fold cross validation and minimal training for classification for each home. The results of combining the same devices that are spatially near each other as a single event are also shown (*).

The lower accuracy of H2 was due to classifier confusion between two devices of the same brand that were on the opposite phase of the installation point of the PLI. Part of the second floor exhibited very weak coupling between the electrical phases, which required us to plug in the PLI on that phase and thus causing some of the signatures to look very similar. We later discuss ways to address this problem, such as using two different PLIs, one on each of the two phases or installing a single PLI at a split-phase 240 V dryer outlet.

Since N-fold cross validation is generally optimistic and is not a true measure of the expected classification performance for a real world system, we perform a follow-up analysis using a minimal training set, i.e. a single event signature for each device to model and then apply it to a test set. For example, a homeowner would likely be only willing to provide a few training events for each appliance. Table 3 summarizes the classification accuracies when using a minimal training data set. We found that this approach has an accuracy of 89.25%

Stability of Signatures Across Homes

Since no two homes have the exact same electrical infrastructure and can have drastically different baseline noise present on the power line, we must examine the portability of noise signature across homes, which is a prerequisite for allowing signatures from one home to be applied to a similar device in another home.

We performed two experiments that together suggest the viability of applying a learned signal for a device in one home to a similar device in another. In the first, we show that the EMI signal of a device is independent of the home in which it is used, thus proving that the signal is intrinsic to the device's functioning. Second, the signal is consistent to within the variance limits for multiple, but similar devices, that is, the signals from the same brand and model of devices are similar.

Our first experiment comprised of collecting data for four of our own preselected devices in each of the seven homes and showing the signal portability through classification results across homes. The average accuracy of the ten trials for a 10-fold cross validation test was 96.87%. For 3 of the 4 devices, the classification accuracy was 100%, strongly suggesting that these devices generated similar signals. Only one device performed poorly. Table 4 summarizes these results.

Device	10 Fold Cross Validation (%)
Camera Charger	100
Laptop	87.5
23W CFL Lamp	100
12W CFL Lamp	100
Aggregate	**96.87**

Table 4: The performance of four of our own devices across different homes using a 10-fold cross validation classification.

We found that the laptop's power adaptor was harder to identify because the extracted feature vector for this device in H6 looked slightly different than the other homes. Upon closer inspection, the noise generated by the laptop's adaptor had a harmonic peak that was very close in amplitude to the center frequency. This marginal difference caused the peak detection algorithm to assign the first harmonic as the center frequency in H6, thus extracting a different feature vector. A simple approach to alleviate this problem is by employing an algorithm that classifies an event only if the distance between the event feature vector and the nearest neighbor in KNN is less than a certain threshold. If not so, the algorithm builds a new feature vector from the next strongest peak. Using this new method, classification accuracy for the laptop was 100%.

In our second experiment, we collected data for eight 20" Dell™ LCD monitors that were of the same model (from our research lab). This LCD model was also coincidently found in H5 and H7. We swapped out the EMI signature for one of the other 9 signatures obtained from a different home or building. For example, we trained the classifier using the signature from H7 and tested it in H5. This test ensured that if the signatures for any of the LCDs were different it would be misclassified as another device. All of our tests yielded

100% accuracy suggesting that similar make and model devices produce similar signature irrespective of the home or building they are in, implying the feasibility of crowd sourcing to obtain device signatures.

Temporal Stability of Noise Signatures

For any signature or fingerprint-based classification system, temporal stability is important. The classifier must perform well for months (or ideally, years) without requiring frequent re-training, which necessitates that the underlying features to remain stable over time.

To show the stability of our noise signatures over time, we chose devices from our long-term deployment dataset, in particular those that met two criteria: (1) devices were fixed in their location during the duration of the deployment and (2) they were not altered in any way, for example light bulbs were not replaced. For these devices, we extracted randomly selected EMI signature vectors spread over the period of the 6-month evaluation.

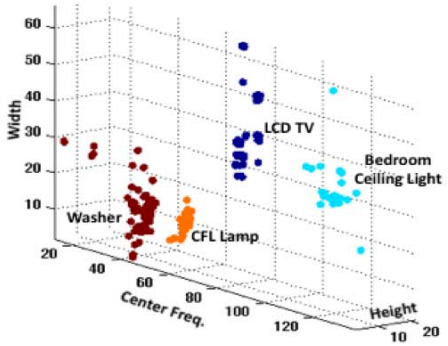

Figure 8: Variation of features over 6 months for four devices shown in the feature space. Note that no cluster overlaps.

Figure 8 shows the temporal stability or variation of the signatures over time for four randomly chosen devices by visualizing the feature vectors in the feature space. We observed that the long-term temporal variation was similar to what we observe in the short-term temporal variation in these devices and that none of the clusters overlapped.

To better understand how temporal variation effects the classification accuracy over time, we generated test sets for each device consisting of all events that happened more than one week prior to the last day of the deployment and a training set consisting of events from *all* devices that happened in the last week. This setup ensured that, if the EMI signatures in the test set deviated more than the distance between the device clusters that the classifier had computed, we would see misclassifications. We observed 100% accuracy with KNN classifier on our results, which indicate that the devices are largely stable over a long period of time. It is important to note that this long-term experiment was straddling the summer and winter seasons.

DISCUSSION AND NEW INSIGHTS

Using EMI for electrical event detection is a promising approach. In this section, we provide additional detail and insights that can shed some light on improving our overall approach as well the limitations and challenges we uncovered. This paper is primarily focused on event detection, but the EMI signal also provides rich information about the state of particular devices (i.e., the setpoint of a dimmer switch, the mode of a washing machine, the changing of TV channels, etc).

Multiple Similar Devices

Having a number of similar devices is a common occurrence in a home, such as having multiple TVs or, more commonly, lights that all use the same brand CFL bulbs. This can cause problems, especially if similar devices cannot be grouped into a single group. For example, grouping two ceiling lights in a bedroom may be acceptable, but grouping lights that are in different rooms or floors may not be. There are two potential solutions to this.

First, the tolerances in components that make up the switching circuitry of a device can introduce enough variability in switching frequency such that the mean of the Gaussian fits observed on the power line are also shifted. Figure 9(a) shows a subset of the spectrum observed by our system showing the spectra of the noise generated by four CFL lamps of the same model that were purchased as a pack of four, thus ensuring that they came from the same manufacturing batch. Note that the spectra do not overlap even among the same batch of CFLs.

Our current hardware is able to discern these subtle features only when observed in isolation, i.e., a line isolation transformer was used to create a noise free power line for this particular experiment. With higher ADC resolution and a larger FFT, this shortcoming may be overcome. Thus, increased resolution may give us better differentiability.

Second, as the conducted EMI travels through the power line, it is affected in several ways, but most prominently the signal is attenuated as a function of the line inductance between the source of noise and the point of sensing. Thus, two identical devices generating identical EMI may look different at the sensing source depending on where the devices are attached along the power line, which we observed in our in-home experiments.

To confirm this, we plugged a device in two different locations in a home and logged the raw spectrum data as sensed by our system. Figure 9(b) shows a spectrum of a small section of this data. The difference in amplitude can be used to differentiate between similar devices located in various parts of the home. This suggests that we might have found a way to determine the number of fixed devices present in the home (i.e., the number of CFL lights in the house or the number of a particular type of TV).

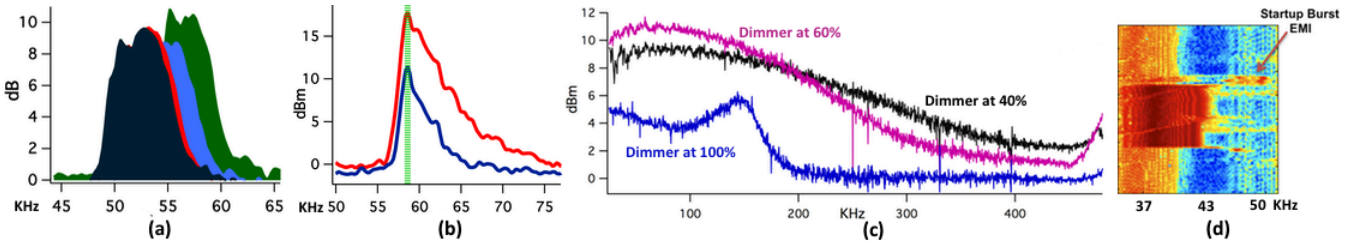

Figure 9: (a) Small, but discernable variation in the mean of the EMI peaks for four same model and brand CFLs. (b) Same CFL lamp plugged into different regions of the home producing EMI amplitude variations. (c) Band limited EMI generated by a dimmer shown at various dim levels. (d) Startup burst of EMI signal generated by CFL lamps on ignition.

It should be observed that only the amplitude varies prominently as a function of the position of the device in the home. Other features such as the mean and variance of the Gaussian remain relatively stable. It is these stable subset features that allow us to correctly identify a mobile appliance, such as a laptop adapter despite changes in its position in a home. It also allows reuse of signature from one home to another without re-training for similar devices. Thus, the stable subset features can be used to identify which *device* (CFL, TV etc.) and the amplitude can be used to resolve the *instance* (TV1, TV2 etc.) in a particular home.

The probability of two different types of devices having the same Gaussian fit mean and variance is small. For independent EMI signals, it is 1/(FFT Size) if we only consider the probability of two devices having the same mean frequency. In practice, since we use not only the location, but also parameters of a fitted Gaussian, the true probability is likely much lower.

Dimmers

Unlike SMPS based devices, dimmers generate a wide band signal from their internal triac switches. Since our system is designed around a Gaussian fit in the feature extraction phase, we model the broadband noise emitted by a dimmer with a Gaussian distribution, even though a band-limited uniform distribution would be more appropriate. Thus, we found that for dimmer controlled devices we observed Gaussian fits with very large variances. Figure 9(c) shows the EMI signal generated by a dimmer at various dim levels.

The difference in the signature generated at various dim levels opens up the possibility for our approach to identify as well as infer the dim of such devices device. Since it is not possible to train at different levels, we would need to build a model for the noise characteristic and how it is affected by the dim level or conduction angle of the triac.

Additional Features

The performance and robustness of a classification algorithm is only as good as the features being used, so it is desirable to explore and extract more advanced features from the underlying signals. During our experimentation we observed multiple potentially useful features that could be used in future classification work. Several devices produce a characteristic EMI pattern that lasts for a short time when they are switched on, i.e. they produce narrow band transient EMI while the SMPS is starting up. These transients are lower frequency than those described by Patel et al. [15]. Figure 9(d) shows a short burst of EMI that most CFL lamps produce when first powered up. The short burst comes from the ignition circuitry in a CFL, which is required to warm the lamp up for operation [10]. Similarly, other devices such as modern TVs and DVD players that have multiple power supplies and supporting electronics also produce start up noise signatures that may provide additional information for determining the category of the devices.

In our current implementation, the algorithm design assumes that the mean or location of EMI peaks do not change when the device is operational. This may not be true for certain devices. We have come across a particular brand of a LCD TVs, where the switching frequency of its power supply is a function of the screen brightness, thus causing the mean of EMI peaks to shift as the content on the screen changes. Our current algorithm will report a new event every time such a change happens, which could overwhelm the system. A new approach may be to lock onto the time varying noise peak and extract temporal features or templates.

Certain motor based devices such as a washing machine or a dishwasher generate low frequency periodic noise patterns in their motor controllers, which could be used to identify the device and its state. For example, in H7, we observed that the clothes washer in its wash cycle produced intermittent noise at roughly 0.1 Hz in contrast to constant noise while in its spin cycle. Such features can be used to build a finite state machine or a statistical model for more detailed classification.

Simultaneous Events

Our current implementation can detect near simultaneous events as close as 102 milliseconds (ms), that is, two events that occur more than 102 ms apart can successfully be detected as separate events by our current implementation. This is based on our current sampling frequency and averaging window size. If simultaneous events happen in shorter than 102 ms duration, they are detected as a single event, and the features extracted are a collection of features from multiple devices. Thus, the underlying features remain intact, but are reported as one event. A new classification approach may be able to separate out these compounded features to identify individual devices.

End User Calibration

Since ElectriSense employs a fingerprinting based approach for device identification, it requires a training process to learn the parameters of various devices in a home. Like any such approach, this is generally accomplished by having the user go through actuating each device at least once.

Fortunately, our approach benefits from the portability of signature across homes, which allows end users to share their device signatures, potentially through crowd sourcing. In other words, an end user may be able to access an online database of popular consumer electronic device signatures submitted by other users who have gone through a calibration process. Additionally, since the EMI that we sense is an engineered signal, it may be possible to generate such signature databases without physically actuating a device or having physical access to it. This can be realized by mining information from a device's FCC compliance report, which lists the frequency and raw magnitude in dBuV for various noise peaks it emits. An example compliance report for a PC power supply is shown here [12]. Mining the datasheets of the internal integrated circuits and oscillators found in consumer electronic devices are other sources for this information.

Phase Coupling

We observed strong signal coupling or crosstalk across phases in most homes, including larger homes like H5 and H6. However, for H2 we found that for some parts of the home, which were on the opposite phase from the sensor it was very difficult to detect events. Either installing two PLIs (one on each phase) or installing the PLI to an available 240 V outlet where both phases are present (typically for a dryer) can address this problem. Installing two PLIs would allow the system to capture events from both phases but also increases the chances of similar looking signatures for two similar devices. This can trivially be addressed by knowing which of the two PLIs detected the event.

CONCLUSION

We have presented a significant new advancement in single-point sensing solution for whole-home electrical event detection. This work leverages the trend towards more energy efficient and low-cost switch mode power supplies for consumer electronic devices. The EMI generated by these power supplies allow us to isolate and classify the occurrence of unique electrical events. Our initial results both validate the effectiveness of our approach and provide a basis for future analyses and improvements. Our new strategy shows significant promise as a practical, low-cost solution for providing disaggregating electrical information for energy monitoring and ubiquitous computing applications.

REFERENCES

1. Abbott, R.E. and S.C. Hadden, Product Specification for a Nonintrusive Appliance Load Monitoring System. EPRI Report #NI-101, 1990.
2. Chen, J., Kam, A.H., Zhang, J., Liu, N. and Shue, L.: Bathroom Activity Monitoring Based on Sound. In: Pervasive Computing 2005, pp. 47--61 (2005).
3. Cohen, G., Gupta, S., Froehlich, J., Larson, E., and Patel, S. GasSense: Appliance-Level, Single-Point Sensing of Gas Activity in the Home. In Pervasive 2010.
4. Energy Star. SMPS and linear power supplies compliance. http://www.efficientpowersupplies.org/pages/NRDC_power_supply_report.pdf. Last accessed March 10, 2010.
5. Fogarty, J., Au, C. and Hudson, S.E.: Sensing from the Basement: A Feasibility Study of Unobtrusive and Low-Cost Home Activity Recognition. In: UIST '06.
6. Froehlich, J., Larson, E., Campbell, T., Haggerty, C., Fogarty, J., Patel, S.N.: HydroSense: infrastructure-mediated single-point sensing of whole-home water activity. In: UbiComp 2009, pp. 235--244 (2009).
7. Functioning of a switched mode Power supply. http://www.smps.us/power-supply.html.
8. Hart, G. Advances in Nonintrusive Appliance Load Monitoring. In Proceedings of EPRI Information and Automation Conference, 1991.
9. Hart, G., Nonintrusive appliance load monitoring. In Proceedings of the IEEE, 1992. 80(12): p. 1870-1891.
10. High voltage driver IC for Compact Fluoroscent Lamps. http://www.nxp.com/documents/data_sheet/UBA2021.pdf.
11. In-line current monitoring device. http://www.p3international.com/products/special/P4400/P4400-CE.html. Last accessed March 10, 2010.
12. Jetway Computers, compliance report. Page 13, http://www.jetwaycomputer.com/download/resources/J7F4%20FCC%20report.pdf. Last accessed March 5, 2010.
13. Kim, Y., Schmid, T., Charbiwala, Z., Srivastava, M.B.: ViridiScope: design and implementation of a fine grained power monitoring system for homes. In: UbiComp 2009.
14. Lifton, J., Feldmeier, M., Ono, Y., Lewis, C., and Paradiso, J. A. A platform for ubiquitous sensor deployment in occupational and domestic environments. In Proceeding of Information Processing in Sensor Networks, 2007.
15. Patel, S.N., Robertson, T., Kientz, J.A., Reynolds, M.S. and Abowd, G.D.: At the Flick of a Switch: Detecting and Classifying Unique Electrical Events on the Residential Power Line. In: UbiComp 2007, pp. 271-288
16. Patel, S.N., Reynolds, M.S., Abowd, G.D. Detecting Human Movement by Differential Air Pressure Sensing in HVAC System Ductwork: An Exploration in Infrastructure Mediated Sensing. In Pervasive 2008. pp 1-18. 2008.
17. Patel, S.N., Truong, K.N., Abowd, G.D. PowerLine Positioning: A Practical Sub-Room-Level Indoor Location System for Domestic Use. In Ubicomp 2006.
18. Tapia, E.M., Intille, S.S., Lopez, L., and Larson, K. The design of a portable kit of wireless sensors for naturalistic data collection. In Proc. of Pervasive 2006. pp. 117-134. 2006.