At the Flick of a Switch: Detecting and Classifying Unique Electrical Events on the Residential Power Line
(Nominated for the Best Paper Award)

Shwetak N. Patel, Thomas Robertson, Julie A. Kientz,
Matthew S. Reynolds, and Gregory D. Abowd

College of Computing, School of Interactive Computing, & GVU Center
Georgia Institute of Technology
85 5th Street NW, Atlanta GA 30332-0280 USA
{shwetak,troomb1,julie,mattr,abowd}@cc.gatech.edu

Abstract. Activity sensing in the home has a variety of important applications, including healthcare, entertainment, home automation, energy monitoring and post-occupancy research studies. Many existing systems for detecting occupant activity require large numbers of sensors, invasive vision systems, or extensive installation procedures. We present an approach that uses a single plug-in sensor to detect a variety of electrical events throughout the home. This sensor detects the electrical noise on residential power lines created by the abrupt switching of electrical devices and the noise created by certain devices while in operation. We use machine learning techniques to recognize electrically noisy events such as turning on or off a particular light switch, a television set, or an electric stove. We tested our system in one home for several weeks and in five homes for one week each to evaluate the system performance over time and in different types of houses. Results indicate that we can learn and classify various electrical events with accuracies ranging from 85-90%.

1 Introduction and Motivation

A common research interest in ubiquitous computing has been the development of inexpensive and easy-to-deploy sensing systems to support activity detection and context-aware applications in the home. For example, several researchers have explored using arrays of low-cost sensors, such as motion detectors or simple contact switches [15, 16, 18]. Although these solutions are cost-effective on an individual sensor basis, they are not without some drawbacks. For example, having to install and maintain many sensors may be a time-consuming task, and the appearance of many sensors may detract from the aesthetics of the home [3, 7]. In addition, the large number of sensors required for coverage of an entire home may increase the number of potential failure points. To address these concerns, recent work has focused on sensing through existing infrastructure in a home. For example, researchers have looked at monitoring the plumbing infrastructure in the home to infer basic activities [6] or using the residential power line to provide indoor localization [13]. Inspired by the theme of leveraging existing infrastructure to support activity detection, we present an approach that uses the home's electrical system as an information source to

J. Krumm et al. (Eds.): UbiComp 2007, LNCS 4717, pp. 271–288, 2007.

observe various electrical events. The detection and classification of these events can be used later for a variety of applications, such as healthcare, entertainment, home automation, energy monitoring, and post-occupancy research studies.

A principal advantage of the approach presented in this paper is that it requires only the installation a single, plug-in module that connects to an embedded or personal computer. The computer records and analyzes electrical noise on the power line caused by the switching of significant electrical loads. Machine learning techniques applied to these patterns identify when unique events occur. Examples include human-initiated events, such as turning on or off a specific light switch or plugging in a CD player, as well as automatic events, such as a compressor or fan of an HVAC system turning on or off under the control of a thermostat.

By observing actuation of certain electrical devices, the location and activity of people in the space can be inferred and used for applications that rely on this contextual information. For example, detecting that a light switch was turned on can be an indication that someone entered a room, and thus an application could adjust the thermostat to make that room more comfortable. We can also detect other human-initiated kitchen events, such as a light turning on inside a refrigerator or microwave when its door is opened. The combination of these events may indicate meal preparation. Our approach also has implications for providing a low-cost solution for monitoring energy usage. An application could log when particular electrical loads are active, revealing how and when electrical energy is consumed in the household, leading to suggestions on how to maintain a more energy-efficient household. In addition, because our approach is capable of differentiating between the on and off events of a particular device in real time, those events can be “linked” to other actuators for a variety of home automation scenarios. One can imagine a home automation system that maps the actuation of a stereo system to an existing light switch without having to install additional wiring.

In this paper, we first present a review of related work in event detection for indoor settings, identifying the inspiration for our work and how it complements and extends past results. We then describe the underlying theory and initial implementation details of our approach to powerline event detection. We report the results of a series of tests to determine the stability of our approach over time and its capability of sensing electrical events in different homes. These tests consisted of installing our device in a single location of a house and collecting data on a variety of electrical events within that house. Results show our support vector machine system can learn and later classify various unique electrical events with accuracies ranging from 85-90%. Finally, we discuss the results, current limitations and potential improvements for this powerline event detection approach.

2 Related Work

We can classify research in activity and behavior recognition in a home setting by examining the origin of the proposed sensing infrastructure. The first area of classification includes approaches that introduce new, independent sensors into the home that directly sense various activities of its residents. This classification includes infrastructures where a new set of sensors and an associated sensor network (wired or wireless) is deployed. A second area encompasses those approaches that take advantage of existing home infrastructure, such as the plumbing or electrical busses in a home, to sense various activities of residents. The goal of the second approach is to lower the adoption barrier by reducing the cost and/or complexity of deploying or maintaining the sensing infrastructure.

Some research approaches use high-fidelity sensing to determine activity, such as vision or audio systems that capture movements of people in spaces [2, 10]. Chen *et al.* installed microphones in a bathroom to sense activities such as showering, toileting, and hand washing [5]. While these approaches may provide rich details about a wide variety of activities, they are often very arduous to install and maintain across an entire household. Furthermore, use of these high fidelity sensors in certain spaces raise concerns about the balance between value-added services and acceptable surveillance, particularly in home settings [3, 7, 9].

Another class of approaches explores the use of a large collection of simple, low-cost sensors, such as motion detectors, pressure mats, break beam sensors, and contact switches, to determine activity and movement [15, 16, 18]. As an example of this low-cost approach, Tapia *et al.* discussed home activity recognition using many state change sensors, which were primarily contact switches [15, 16]. These sensors were affixed to surfaces in the home and logged specific events for some period of time. The advantage of this approach is being able to sense physical activities in a large number of places without the privacy concerns often raised for high-fidelity sensing (*e.g,.* bathroom activity). There are also some disadvantages to this add-on sensor approach, which include the requirements of powering the sensors, providing local storage of logged events on the sensor itself, or a wireless communication backbone for real-time applications. These requirements all complicate the design and maintenance of the sensors, and the effort to install many sensors and the potential impact on aesthetics in the living space may also negatively impact mass adoption of this solution. As an example of the often difficult balance of the value of in home sensing and the complexity of the sensing infrastructure, the Digital Family Portrait is a peace of mind application for communicating well-being information from an elderly person's home to a remote caregiver [14]. In their deployment study, movement data was gathered from a collection of strain sensors attached to the underside of the first floor of an elder's home. The installation of these sensors was difficult, time-consuming, and required direct access to the underside of the floor. Though the value of the application was proven, the complexity of the sensing limited the number of homes in which the system could be easily deployed.

Other approaches, which are similar to ours, are those that use existing home infrastructure to detect events. Fogarty *et al.* explored attaching simple microphones to a home's plumbing system, thereby leveraging an available home infrastructure [6]. The appeal of this solution is that it is low-cost, consists of only a few sensors, and is sufficient for applications, such as the Digital Family Portrait, for which the monitoring of water usage is a good proxy for activity in the house. This approach requires relatively long timescales over which events must be detected, sometimes up to ten seconds. This longer time increases the likelihood of overlapping events, which are harder to distinguish. In contrast, powerline event detection operates over timescales of approximately half a second and thus overlapping is less likely. Some water heaters constantly pump hot water through the house, complicating the detection of some on-demand activities. Detecting noise on water pipes introduced by other household infrastructure requires careful placement of the microphone sensors. Some homes may not have plumbing infrastructure that is easily accessible, particularly those with a finished basement or no basement at all. Despite these limitations, this solution is very complementary to our approach, as some events revealed by water usage, such as turning on a faucet in a sink or flushing a toilet, do not have direct electrical events that could serve as predictive antecedents. The converse also holds, as a light being turned on often does not correlate with any water-based activity. Another "piggybacking" approach is to reuse sensing infrastructure or devices in the home that may be present for other purposes. For example, ADT Security System's QuietCare [1] offers a peace of mind service that gathers activity data from the security system's motion detectors.

There are several other techniques that employ electrical power use to sense activity. For example, some researchers have monitored electrical current flow to infer the appliances or electrical equipment being used in the house as a proxy for detecting activity [12, 16]. The MITes platform supports the monitoring of current consumption of various appliances of interest. Changes in current flow indicate some change in state for the instrumented appliance, such as a change from on to off. This solution requires a current sensor to be installed inline with each appliance or around its power cord and thus only works well if it is sufficient to study the usage of a small subset of appliances and those appliances' power feeds are easy accessible. An extension to the MITes work would be to install current sensors on major branch circuits of the power lines, but this may require professional installation to provide an acceptable level of safety. Our solution can detect a larger number of appliances with less instrumentation and with a much easier deployment phase. Our approach is influenced by our previous work in PowerLine Positioning system [13], which uses existing powerline infrastructure to do practical localization within a home. The main difference between that work and the present work is that we are passively sensing electrical events using simple events, whereas our previous work senses the location of actively tagged objects.

3 Our Approach and System Details

Our prototype system consists of a single module (see Figure 1) that is plugged into any electrical outlet in the home. Although not necessarily required, we installed it in a convenient, central location in the home while experimenting with the setup. The other end of the plug-in unit is connected via USB to a computer that collects and performs the analysis on the incoming electrical noise. The system learns certain characteristics from electrical noise produced by switching an electrical device on or off and later predicts when those devices are actuated based on the learned phenomena. Note that we present an approach for countries that use 60 Hz electrical systems, but our approach can easily be extended to different frequencies used in other countries (*i.e.,* those that use 50 Hz).

3.1 Theory of Operation

Our approach relies on the fact that abruptly switched (mechanical or solid-state) electrical loads produce broadband electrical noise either in the form of a transient or

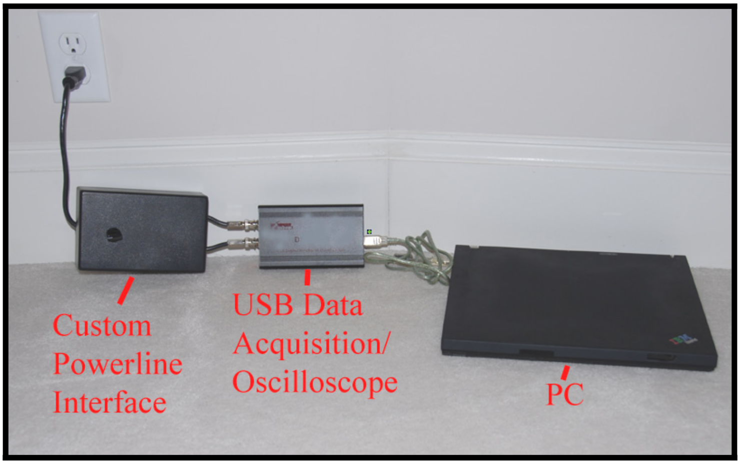

Fig. 1. Our prototype system consists of a powerline noise analyzer plugged in to an ordinary wall outlet and connected to a PC

continuous noise. This electrical noise is generated either between hot and neutral (known as normal mode noise) or between neutral and ground (known as common mode noise). Transient and continuous noise on the residential power line is typically high in energy and may often be observed with a nearby AM radio. The types of electrical noise in which we are interested are produced within the home and are created by the fast switching of relatively high currents. For example, a motor-type load, such as a fan, will create a transient noise pulse when it is first turned on and will then produce a continuous noise signal until it is turned off. In addition, the mechanical switching characteristics of a light switch itself can generate transient electrical noise [8]. Other examples of noisy events include using a garage door opener, plugging in a power adaptor for an electric device, or turning on a television. Marubayashi provides a more complete description of this electrical noise phenomenon [11].

In the case of transient noise, the impulses typically last only a few microseconds and consist of a rich spectrum of frequency components, which can range from 10 Hz to 100 kHz. Thus, it is interesting to consider both the temporal nature (duration) of the transient noise and its frequency components. Depending on the switching mechanism, the load characteristics, and length of transmission line, these impulses can be very different. For example, Figure 2a shows a sample frequency domain graph of a light switch being toggled in a house (light on followed by light off). Note the rich number of high amplitude frequency components for each pulse and their relative strengths. Also, notice that the signature of a device being turned on is different from the same device being turned off. Figure 2b shows the same switch being actuated in the same order, but taken 2 hours later, and Figure 2c shows it taken 1 week later. The amplitudes of individual frequency components and the duration of the impulse produced by each switch are similar between the three graphs, although there are a few high frequency regions that are different across the samples. Even similar light switches produce different signatures, which is likely due to the mechanical construction of each switch and the influence of the power line length connected to each switch. For example, we observed that three-way wall switches connected to the same light each produced discernable signatures. The main difference was in the relative amplitudes of the frequencies being observed. For devices that produce continuous noise, they are bounded by some transient phenomena, but also exhibit electrical noise during their powered operation. For this class of noises, it is possible to not only identify it based on its transient response but also its continuous noise signature.

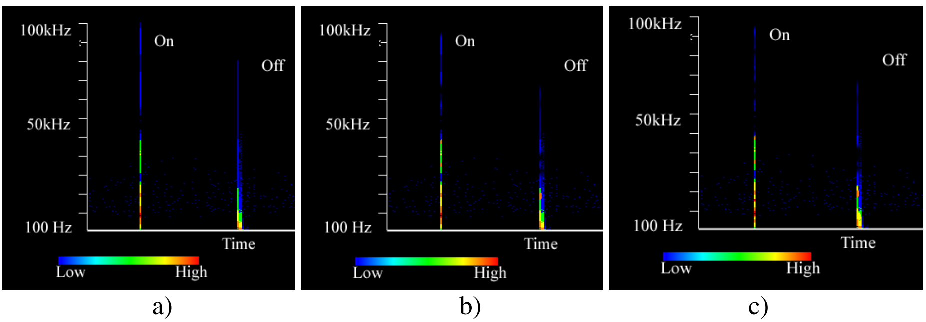

Fig. 2. Frequency spectrum of a particular light switch being toggled (on and off events). The graphs indicate amplitudes at each frequency level. Events in (b) were captured two days after (a), and events in (c) were captured one week after (a). Each sample is rich in a broad range of frequencies. On and off events are each different enough to be distinguished. In addition, the individual on and off events are similar enough over time to be recognized later.

Because we assume the noise signature of a particular device depends both on the device and the transmission line behavior of the interconnecting power line, we have attempted to capture both contributions in a single model. Figure 3 depicts a high-level overview of our simplified model of a home's electrical infrastructure and where particular noise transfer functions occur, denoted as H(s). These transfer functions reflect our expectation that both the electrical transmission lines and the data collection apparatus connected to that line all contribute to some transformation of the noise from the source to the collection apparatus. The observed noise results from the imposition of all the transfer functions against the generated noise. The influence of the transmission line’s transfer function is an important contributor to the different electrical noise signatures we observed, which explains why similar device types (*e.g.,* light switches) can be distinguished and why the location of the data collection module in the house impacts the observed noise.

In our simplified model, three general classes of electrical noise sources may be found in a home (see Figure 3): resistive loads, inductive loads such as motors, and loads with solid state switching. Purely resistive loads, such as a lamp or an electric stove, do not create detectable amounts of electrical noise while in operation, although as a resistor, they can be expected to produce trace amounts of thermal noise (Johnson noise) at an undetectable level. In this particular case, only a transient

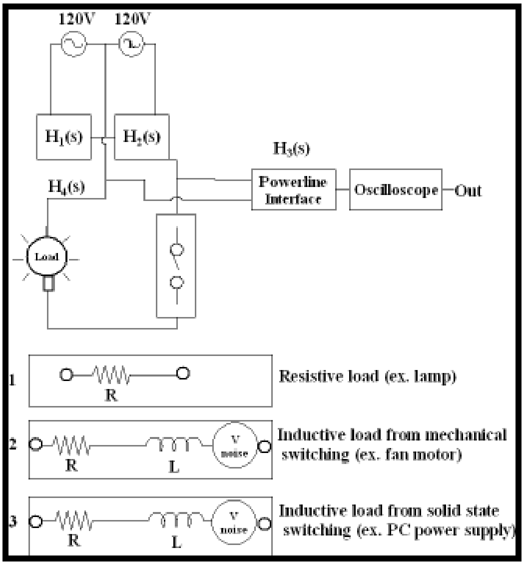

Fig. 3. Overview of the powerline infrastructure and location of particular signal/noise transfer functions, $H_n(s)$. The bottom of the figure shows three general types of loads found in a home, a purely resistive, an inductive where voltage noise is generated from a continuous mechanical switching (motors), and an inductive load where voltage noise is generated by an internal oscillator of a solid state switch.

noise is produced by minute arcing in the mechanical switch itself (wall switch) when the switch is turned on or off. A motor, such as in a fan or a blender, is modeled as both a resistive and inductive load. The continuous breaking and connecting by the motor brushes creates a voltage noise synchronous to the AC power of 60 Hz (and at 120 Hz). Solid state switching devices, such as MOSFETs found in computer power supplies or TRIACs in dimmer switches or microwave ovens, emit noise that is different between devices and is synchronous to an internal oscillator. Thus, the latter two classes contribute noise from both the external power switching mechanism (transient) and the noise generated by the internal switching mechanism (continuous).

In the United States, the Federal Communications Commission (FCC) sets guidelines on how much electrical noise AC-powered electronic devices can conduct back onto the power line (Part 15 section of the FCC regulations). Device-generated noise at frequencies between 150 kHz-30 MHz cannot exceed certain limits. Regulatory agencies in other countries set similar guidelines on electronic devices. Although this mainly applies to electronic devices, such as those that have solid state switching power supplies, this gives us some assurance about the type and amount of noise we might expect on the power line.

It is often extremely difficult to analytically predict the transient noise from the general description of a load and its switching mechanism because ordinary switches are usually not well characterized during their make-and-break times. However, it is possible to take a mapping approach by learning these observed signatures using supervised machine learning techniques. The challenge then becomes finding the important features of these transient pulses and determining how to detect the relevant ones of interest.

3.2 Hardware Details

To explore the idea of detecting and learning various electrical events in the home, we first built a custom data collector that consisted of a powerline interface with three outputs (see Figures 4 and 5). One output was the standard 60 Hz AC power signal, which we used during our initial testing and exploratory phase. The second output was an attenuated power line output that has been bandpass-filtered with a passband of 100 Hz to 100 kHz. The third output was similarly attenuated and was bandpass-filtered with a 50 kHz to 100 MHz passband. We chose these different filtered outputs to have the flexibility to experiment with different frequency ranges (see Figure 6). Both filtered outputs have a 60 Hz notch filter in front of their bandpass filters to remove the AC power frequency and enhance the dynamic range of the sampled data. We built our interface so that we could monitor the power line between hot and neutral, neutral and ground, or hot and ground. For the work reported here, we chose to observe the noise between hot and neutral (normal mode) because many loads that we would like to observe (such as table lamps and small appliances) do not have a ground connection.

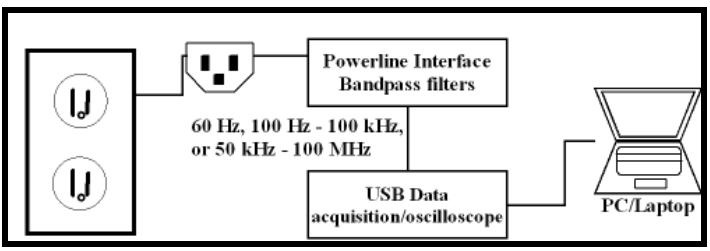

Fig. 4. Block diagram of our powerline interface system

We further chose to interface with only one 120V leg or branch of the electrical system. Most residential houses and apartments in North America and many parts of Asia have a single-phase or a split single-phase electrical system. This means there are two 120V electrical branches coming into the house to supply 240V appliances, but the two branches are still in phase. We found that the noises generated by devices of interest connected to the other electrical branch were already being coupled to the electrical branch we interfaced to, and so were detectable by our system. While this approach was practical and sufficient for our research prototype, we could also plug a coupler into a 240V outlet to ensure we have direct access to both electrical branches.

Finally, the outputs of the powerline interface are connected to a dual-input USB oscilloscope interface (EBest 2000) that has a built-in gain control. Each input has 10-bit resolution with a full scale voltage of 1V, so the least significant bit represents a

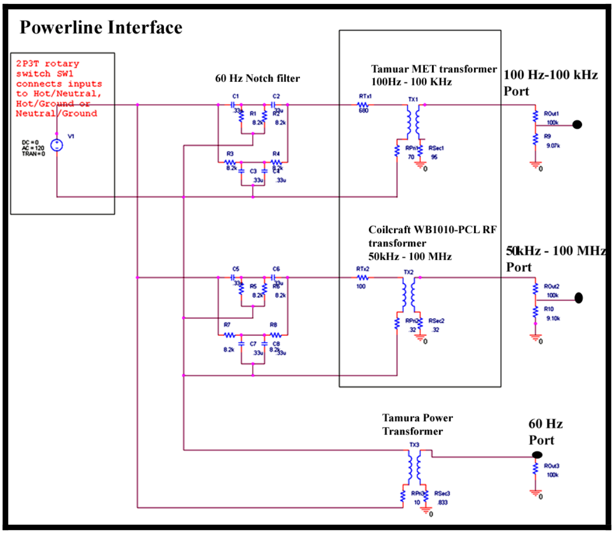

Fig. 5. The schematic of our powerline interface device

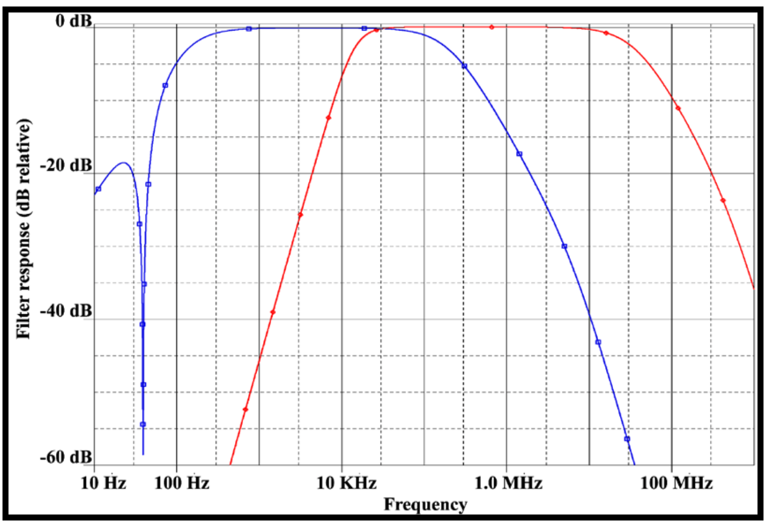

Fig. 6. A model of the frequency response curve of our powerline data collection apparatus at the 100 Hz – 100 kHz and the 50 kHz – 100 MHZ outputs. The 60 Hz dip is from the notch filter.

voltage of 4 mV. The oscilloscope interface has a real-time sampling rate of 100 million samples/sec. A C++ API is provided, resulting in a simple software interface to the sampled signal.

3.3 Software Details

For the software components of our prototype, we wrote a C++ application to sample the USB oscilloscope interface and perform a Fast Fourier Transform (FFT) on the incoming signal to separate the component frequencies for our analysis. The application also produces a waterfall plot, a commonly used frequency domain visualization in real-time used for visual inspection (such as in Figure 2). The application performs this analysis in nearly real-time, and it has the ability to record the data stream for post processing. A second application, written in Java, performs the machine learning and provides the user interface for the system. The Java application connects via a TCP connection to the FFT application and reads the data values. The Java application provides the user interface for surveying the home and remotely accessing the data from the powerline interface. We used the Weka [17] toolkit for our machine learning implementation.

3.4 Electrical Events That Can Be Recognized

Having built our data collection apparatus, we first wanted to identify the variety of electrical devices we could detect with our apparatus and see which electrical devices would produce recognizable signatures that can be used for our machine learning software. For this exploration, we installed our apparatus in a single fixed location throughout the data collection process. We collected data with both the low frequency (100 Hz – 100 KHz) and high frequency (50 kHz – 100 MHz) ports. We took care to ensure no major electrical devices were activated (such as the HVAC, fridge, water pumps, *etc.*) by turning them off for the duration of the testing so we knew which devices were causing which response. For each electrical device of interest, we visually observed and collected noise signatures for turning the device on, turning it off, and its stable on state. Table 1 shows the various devices we were able to detect and the events we were able to observe for each device (on, off, continuously on state). Although we could have observed many more devices, we only show a representative sample of commonly used devices.

After initial experimentation, we found that most loads drawing less than 0.25 amps were practically undetectable Loads above that amount produced very prominent electrical noise (transient and/or continuous). This is related to the dynamic range of our data collection device—a collection device with more than 10 bits of resolution would be able to detect lower current devices. The devices listed in Table 1 showed not only strong but also consistently reproducible signatures. However, we did observe a limitation in how quickly we could switch a given device (*i.e.,* the delay between toggles). Depending on the device, we observed that approximately 500 ms delay between subsequent toggles was required for our data collection apparatus to detect a noise impulse successfully. This is largely attributed to the sampling and processing latency from our device (*e.g.,* USB latency plus processing delays on the PC).

While most devices produced a transient pulse only a few microseconds in duration in their energized state, certain devices continuously produced electrical noise while they were powered, as expected. For example, lamp dimmers or wall-mounted dimmer switches produced noise that was very rich in harmonics while they were

Table 1. Electrical devices we tested and which events we were able detect. These devices also consistently produced detectable event signatures.

Device Class/Type	Devices Observed	On to Off Transition Noise?	Off to On Transition Noise?	Continuously On Noise?
Resistive	Incandescent lights via a wall switch	Y	Y	N
	Microwave door light	Y	Y	N
	Oven light/door	Y	Y	N
	Electric stove	Y	Y	N
	Refrigerator door	Y	Y	N
	Electric Oven	Y	Y	N
Inductive (Mechanically Switched)	Bathroom exhaust fan	Y	Y	N
	Ceiling fan	Y	Y	N
	Garage door opener	Y	Y	N
	Dryer	Y	Y	N
	Dishwasher	Y	Y	N
	Refrigerator compressor	Y	Y	N
	HVAC/Heat Pump	Y	Y	N
	Garbage disposal	Y	Y	N
Inductive (Solid State Switched)	Lights via a dimmer wall switch	Y	Y	Y
	Fluorescent lights via a wall switch	Y	Y	N
	Laptop power adapter	Y	N	N
	Microwave Oven	Y	Y	Y
	Television (CRT, plasma, or LCD)	Y	Y	N

activated. Similarly, microwave ovens also coupled broadband noise back on the power line during its use. These devices tended to produce strong continuous noise above 5 kHz and reaching up to 1 MHz. We also found that switching power supplies, such as from a laptop or PC, produced considerably higher noise in the 100 kHz – 1 MHz area than at the lower 100 Hz – 5 kHz range.

To understand devices that produced continuous noise, we tested various switching power supplies in isolation from other electrical line noise (see Figure 7). Using the higher 50 kHz – 100 MHz output on our data collection apparatus, we found that many of these devices produced more detectable continuous noise at the higher frequencies. At the lower 100 Hz – 5 kHz range, we saw fairly low amplitude, continuous noise, and a higher transient noise effect (from the flipping of the switch).

In the 100 Hz – 100 kHz range, motor-based devices, such as a ceiling or bathroom exhaust fan, exhibited slightly longer duration transient pluses when activated with a switch, but did not show continuous normal mode noise which would have been expected from the repeated electromechanical switching from the motor brushes. We attribute this difference to our 60 Hz notch filter, which blocked the 60 Hz power frequency. To confirm this hypothesis, we conducted another experiment in which we isolated various mechanically-switched devices (*e.g.,* fans) and looked at their noise output (see Figure 7). In the case of the fan, our data collection apparatus did indeed show the transient pulse, but not the continuous electrical noise.

From these observations, we are able to characterize the noise characteristics produced by different devices. We observed that transient noise produced from a single abrupt switching event (e.*g.,* a wall switch) tended to produce signals rich in high amplitude components in the lower frequency range (100 Hz – 5 KHz). Inductive loads featuring a solid state switching mechanism generally produced continuous noise in the 5 kHz – 1 MHz range. Inductive loads with mechanically switched voltages produce noise near 60 Hz, but our data collection apparatus filtered out much of that noise. We thus observed that the analysis of the frequency spectrum may be broken up into two parts. The lower frequency space (100 Hz – 5 kHz) is effective for analysis for transient noise events, such as those produced by wall switches. The higher frequency is better for continuous noise events, such as those produced by TRIACs and switching power supplies. We even observed that dim levels can also be gathered from the continuous noise frequency generated by the TRIACs. For this particular paper, we primarily focus on exploring transient noise events. Similar analysis and learning could be applied to continuous noise events.

Fig. 7. The setup we constructed for isolating and testing the noise response for various electrical devices on an individual basis

3.5 Detecting and Learning the Signals

Our detection approach requires detection of the transient pulse of electrical noise followed by extraction of relevant features for learning classification.

3.5.1 Detecting Transient Pulses

The filtering hardware in the powerline interface removes most of the high frequency noise. Some broadband noise is always present, but typically at low amplitudes. To detect the transient pulses, we employ a simple sliding window algorithm to look for drastic changes in the input line noise (both beginning and end). These drastic changes, lasting only a few microseconds, are labeled as candidate signals and processed further. The sliding window acquires a 1-microsecond sample, which is averaged from the data acquired after performing the FFT on data from the data acquisition hardware. Each sample consists of frequency components and its associated amplitude values in vector form. Each vector consists of amplitude values for frequency intervals ranging between 0 and 50 kHz. We then compute the

Euclidean distance between the previous vector and the current window's vector. When the distance first exceeds a predetermined threshold value, the start of the transient is marked. The window continues to slide until there is another drastic change in the Euclidean distance (the end of the transient). Although the threshold value was determined through experimentation, we can imagine learning and adapting the thresholds over time.

After having isolated the transient, we are left with N vectors of length L, where N is the pulse width in 1 microsecond increments and L is the number of frequency components (2048 in our case). A new vector of length $L + 1$ is then constructed by averaging the corresponding N values for each frequency components. The $(L + 1)$st value is simply N, the width of the transient. This value then serves as our feature vector for that particular transient.

3.5.2 Learning the Transients

For our learning algorithm, we employed a support vector machine (SVM) [4]. SVMs perform classification by constructing an N-dimensional hyperplane that optimally separates the data into multiple categories. The separation is chosen to have the largest distance from the hyperplane to the nearest positive and negative examples. Thus, the classification is appropriate for testing data that is near, but not identical, to the training data as is the case for the feature vectors for the transients. SVMs are appealing because our feature space is fairly large compared to our potential training set. Because SVMs employ overfitting protection, which does not necessarily depend on the number of features, they have the ability to better handle large feature spaces. The feature vectors are used as the support vectors in the SVM. We used the Weka Toolkit to construct an SVM, using labeled training data to later classify the query points.

4 Feasibility and Performance Evaluation

To evaluate the feasibility and performance of our approach, we tested it in six different homes of varying styles, age, sizes, and locations. We first tested our transient isolation scheme in a single home. Next, we conducted a feasibility study in that home for a six-week period to determine the classification accuracy of various electrical events over an extended period of time. Finally, for the five other homes, we conducted a one-week study to reproduce the results from the first home.

4.1 Transient Isolation Evaluation

To evaluate the feasibility of our automatic transient detection scheme, we collected data from one home for a four-hour period and had our software continuously isolate transient signals. During that period, we actuated various electrical components and made a note of their timestamps. A total of 100 distinct events were generated during this period. For each event, we then determined if a transient was isolated successfully at the noted times. Table 2 shows the results of five different four-hour sessions. We report the percentage of successfully identified transients out of the number of event triggers. We believe the reason for the missed events was because of our static threshold algorithm. An adaptive threshold approach would mitigate this problem.

Table 2. Percentage of successfully identified transient pulses using our transient isolation scheme. Each test lasted for a four-hour period with approximately 100 possible transient events in each period.

Test 1 (% found)	Test 2 (% found)	Test 3 (% found)	Test 4 (% found)	Test 5 (% found)
98	93	91	88	96

4.2 Classifying Transient Events in Various Homes

The aim of our extended 6-week evaluation was to determine the classification accuracy of various types of electrical devices and how often we had to retrain the system (signal stability over time). The other five deployments were used to show that we could detect events similar to those of the initial home and to show that the transient noise signatures were temporally stable in other homes as well. Despite the small number of homes, we tried to test a variety of homes and sizes, including older homes with and without recently updated electrical systems (see Table 3). We also included an apartment home in a six-story building, as we expected its electrical infrastructure to be somewhat different from that of a single family home. We were interested in testing the types of electrical devices listed in Table 1, so we ensured that the homes in which we deployed had most of these devices.

For the entire testing period, we installed our data collection apparatus in the same electrical outlet. For Home 1, we collected and labeled data at least three times per week during the 6-week period. The data collection process involved running our system and toggling various predetermined electrical devices (see Table 1 for examples). For each device toggled, we manually labeled each on-to-off and off-to-on event. In addition, we captured at least two instances of each event during each session. For Home 1, we selected 41 different devices for testing (82 distinct events) and collected approximately 500 instances during each week. Thus, approximately 3000 labeled samples were collected during the 6-week period.

We collected and labeled data in a similar manner for the shorter 1-week deployments. We collected training data at the beginning of the week and collected additional test data at the end of the week. At least 4 instances of each event were gathered for the training set. Because we had control over the events, the number of distinct events were fairly equally distributed among the data and not biased towards a single device or switch for all the 6 homes.

Tables 4 and 5 show classification accuracies for the different homes we tested. For Home 1, we show the classification accuracy of test data gathered at various times during the six weeks using the training set gathered during the first week. The average overall classification accuracy in Home 1 was approximately 85% (Table 4). We also show the accuracy of the classification for varying training set sizes. Because there can potentially be many events of interest in the home, making the training process an arduous task, we wanted to find the minimum number of samples that would provide reasonable performance. The results suggest that there is only a slight decrease in classification over the 6 week period. The results also suggest that a small number of

Table 3. Descriptions of the homes in which our system was deployed. Home 1 is where we conducted the long-term 6 week deployment.

Home	Year Built	Electrical Remodel Year	Floors/ Total Size (Sq Ft)/ (Sq M)	Style	Bedrooms/ Bathrooms/ Total Rms.	Deply. Time (weeks)
1	2003	2003	3/4000/371	1 Family House	4/4/13	6
2	2001	2001	3/5000/464	1 Family House	5/5/17	1
3	1999	1999	1/700/58	1 Bed Apartment	1/1/4	1
4	2002	2002	3/2600/241	1 Family House	3/3/12	1
5	1935	1991	1/1100/102	1 Family House	2/1/7	1
6	1967	1981	1/1500/140	1 Family House	2/1/7	1

Table 4. Performance results of Home 1. The accuracies are reported based on the percentage of correctly identified events. Training happened during Week 1, and we reported the accuracies of the classifier for test data from subsequent weeks using that initial training set from week 1. Overall classification accuracy of a simple majority classifier was 4%.

	SVM accuracies during specific weeks of testing					
Training Set Size/Instances per event	**Week 1 (%)**	**Week 2 (%)**	**Week 3 (%)**	**Week 4 (%)**	**Week 5 (%)**	**Week 6 (%)**
164/2	83	82	81	79	80	79
246/3	86	84	85	84	82	83
328/4	88	91	87	85	86	86
410/5	90	92	91	87	86	87

training instances result in lower classification accuracies. In addition, the majority classifier had accuracies of only about 4% on average, because of the equal distribution of the distinct events in the training and test data,

As reported, increasing the number of training instances did increase the classification accuracy. A small number of training samples makes it very important to have accurate training data. Mislabeling of a single training sample can have major impacts on the learned model. We even caught ourselves accidentally mislabeling a few events. For example, the on and off event labels we noted were sometimes flipped for a particular electrical device. Thus, this highlights the importance of designing a training or calibration scheme that mitigates human error during the training and labeling process.

The results from the one-week deployments in the five other homes are shown in Table 5, and the test data from the end of the week showed promising results. We did not see any significant differences in accuracy between old and new homes. The lower classification accuracy for Home 5 was the result of a low frequency noise that interfered with our transient events. Although we could not find the origin of that noise, we can imagine building a smarter system that learns these erroneous noise events to avoid incorrect classifications.

Table 5. Performance results of various homes. The accuracies are reported based on the percentage of correctly identified toggled light switches or other events in the test data set. The results of a majority classifier are also shown. For each home, the training of the data occurred at the beginning of the week and the test data set was gathered at the end of that week.

Home	Distinct events	Training set (events)	Test set (events)	Accuracy (%)	Majority classif. (%)
2	82	328	100	87	4
3	48	192	96	88	6
4	76	304	103	92	3
5	64	256	94	84	3
6	38	152	80	90	8

5 Discussion of Limitations and Potential Improvements

Although we found promising results with our system, it is not without some limitations and some future considerations. In the current implementation, we purposely analyzed the lower frequency spectrum where solid-state switching devices would produce the lowest interference from potential continuous noise. However, at the same time, this choice limits our feature space. Looking at a larger frequency spectrum could provide better classification for certain transient events. In addition, a fully functional system must be able to detect and to adapt to random noise events when looking for transient pulses. In the future, we plan to improve the feature extraction step. We focused on only the amplitudes of the component frequencies. Phase difference between component frequencies, however, should be considered as part of a feature extraction scheme. In addition, the exploration of other machine learning techniques and application of more domain knowledge of the transient signals may also prove valuable in building a better classifier.

Another consideration is the scaling of our approach. Although unlikely in domestic settings, compound events, such as two lights flipped simultaneously, can produce errors in classification because their combined transient noises produce different feature vectors. This type of event is more of a concern in an extremely large home with many residents or in an apartment building that does not have individually metered units. If users regularly flip light switches nearly simultaneously, this could be trained as a separate event from the individual switches.

We have been primarily focused on domestic environments, but this type of system can also be applied to commercial settings. However, compound events and electrical noise in these settings may become a more significant issue. Another issue is that the electrical lines may be so long that the noise does not reach the analyzer. Commercial buildings typically have multiple electrical legs, and to mitigate problems with compound events and line distance, we could install multiple line noise analyzers throughout an office building to isolate the analysis to certain sections of the building. Our approach will have some difficulty differentiating between individual events among a dense collection of proximal devices that have similar switching and load characteristics. For our approach to scale to these environments, the entire frequency band may needs to be considered. Another drawback of commercial buildings is that they tend to have more noisy components, such as large HVAC systems, connected to the power line that can produce many other transients and mask the pulses of interest..

Our system is more appropriate for detecting and learning fixed electrical devices than mobile devices or portable devices. Though we could support them, portable devices require training the system on any possible outlet that they may be plugged into. In addition, plugging the device into an extension cord or power strip might produce a different fingerprint than plugging it into an electrical outlet directly. With a well-defined set of events that should be detected, a suitable training plan can be devised, but it may become time-consuming as the set grows larger.

In some respects, this system represents a tradeoff between the two categories of systems we mentioned in Section 2. Unlike the first category of prior work, our system does not require the deployment of a large number of sensing units throughout the home. A single data collection module is certainly easier to physically deploy and maintain than a large array of distributed sensors, though one could argue that a single point of failure has been introduced (*e.g.,* what if someone accidentally unplugs the data collection module?). On the other hand, this simplicity of physical installation and maintenance has its cost in terms of training the machine learning algorithm to recognize a significant number of electrical loads. The appropriateness of this tradeoff is thus expected to be application dependent.

6 Conclusion

We presented an approach for a low-cost and easy-to-install powerline event detection system that is capable of identifying certain electrical events, such as switches that are toggled. This system has implications for applications seeking simple activity detection, home automation systems, and energy usage information. We showed how our system learns and classifies unique electrical events with high accuracy using standard machine learning techniques. Additionally, a deployment of our system in several homes showed long-term stability and the ability to detect events in a variety of different types of homes. We also discussed specific events our system can detect and which events may have problems when used for specific applications. Our system has the potential to be integrated easily into existing applications that aim to provide services based on detection of various levels of activity.

Acknowledgments

This work is sponsored in part by the National Science Foundation Graduate Fellowship and the Intel Research Council. The authors would also like to thank the members of the Ubicomp Research Group and Sooraj Bhat at Georgia Tech.

References

1. ADT QuietCare, (2007) http://www.adt.com/quietcare/
2. Bian, X., Abowd, G.D., Rehg, J.M.: Using Sound Source Localization in a Home Environment. In: Proc of the International Conference on Pervasive Computing (2005)
3. Beckmann, C., Consolvo, S., LaMarca, A.: Some Assembly Required: Supporting End-User Sensor Installation in Domestic Ubiquitous Computing Environments. In: Davies, N., Mynatt, E.D., Siio, I. (eds.) UbiComp 2004. LNCS, vol. 3205, pp. 107–124. Springer, Heidelberg (2004)

4. Burges, C.J.C.: A Tutorial on Support Vector Machines for Pattern Recognition. Journal of Data Mining and Knowledge Discovery 2(2) (June 1998)
5. Chen, J., Kam, A.H., Zhang, J., Liu, N., Shue, L.: Bathroom Activity Monitoring Based on Sound. In: Gellersen, H.-W., Want, R., Schmidt, A. (eds.) PERVASIVE 2005. LNCS, vol. 3468, pp. 47–61. Springer, Heidelberg (2005)
6. Fogarty, J., Au, C., Hudson, S.E.: Sensing from the Basement: A Feasibility Study of Unobtrusive and Low-Cost Home Activity Recognition. In: The Proc of ACM Symposium on User Interface Software and Technology (UIST 2006), ACM Press, New York (2006)
7. Hirsch, T., Forlizzi, J., Hyder, E., Goetz, J., Kurtz, C., Stroback, J.: The ELDer Project: Social, Emotional, and Environmental Factors in the Design of Eldercare Technologies. In: Proc of the ACM Conference on Universal Usability, pp. 72–79 (2000)
8. Howell, E.K.: How Switches Produce Electrical Noise. IEEE Transactions on Electromagnetic Compatibility 21(3), 162–170 (1979)
9. Iachello, G., Abowd, G.D.: Privacy and Proportionality: Adapting Legal Evaluation Techniques to Inform Design In Ubiquitous Computing. In: the Proc of CHI 2005, pp. 91–100 (2005)
10. Koile, K., Tollmar, K., Demirdjian, D., Howard, S., Trevor, D.: Activity Zones for Context-Aware Computing. In: Dey, A.K., Schmidt, A., McCarthy, J.F. (eds.) UbiComp 2003. LNCS, vol. 2864, Springer, Heidelberg (2003)
11. Marubayashi, G.: Noise Measurements of the Residential Power Line. In: The Proceedings of International Symposium on Power Line Communications and Its Applications, pp. 104–108 (1997)
12. Paradiso, J.A.: Some Novel Applications for Wireless Inertial Sensors. In: Proc of NSTI Nanotech 2006, Boston, MA, May 7-11, 2006, vol. 3, pp. 431–434 (2006)
13. Patel, S.N., Truong, K.N., Abowd, G.D.: PowerLine Positioning: A Practical Sub-Room-Level Indoor Location System for Domestic Use. In: The Proccedings of Ubicomp (2006)
14. Rowan, J., Mynatt, E.D.: Digital Family Portrait Field Trial: Support for Aging in Place. In: Proc of the ACM Conference on Human Factors in Computing Systems (CHI 2005), pp. 521–530. ACM Press, New York (2005)
15. Tapia, E.M., Intille, S.S., Larson, K.: Activity recognition in the home setting using simple and ubiquitous sensors. In: Ferscha, A., Mattern, F. (eds.) PERVASIVE 2004. LNCS, vol. 3001, pp. 158–175. Springer, Heidelberg (2004)
16. Tapia, E.M., Intille, S.S., Lopez, L., Larson, K.: The design of a portable kit of wireless sensors for naturalistic data collection. In: Fishkin, K.P., Schiele, B., Nixon, P., Quigley, A. (eds.) PERVASIVE 2006. LNCS, vol. 3968, pp. 117–134. Springer, Heidelberg (2006)
17. Weka. Weka 3: Data Mining Software in Java (2007), http://www.cs.waikato.ac.nz/ml/weka/
18. Wilson, D.H., Atkeson, C.G.: Simultaneous Tracking and Activity Recognition (STAR) Using Many Anonymous, Binary Sensors. In: Gellersen, H.-W., Want, R., Schmidt, A. (eds.) PERVASIVE 2005. LNCS, vol. 3468, pp. 62–79. Springer, Heidelberg (2005)