Leveraging data from environmental sensors to enhance electrical load disaggregation algorithms

Mario Bergés, Lucio Soibelman, H. Scott Matthews
Carnegie Mellon University. 5000 Forbes Ave., Pittsburgh, PA 15213 USA

Appeared in: *Proceedings of the 13th International Conference on Computing in Civil and Building Engineering, Nottingham, UK, 2010.*

Abstract

The idea of sustainable or green buildings generally stops after the design and construction phases. Little effort is made to continuously monitor and control the energy profile throughout the life-cycle of these facilities. To effectively identify opportunities for consumption reduction, measurement and feedback of current energy use is necessary. Monthly utility bills are inadequate for planning conservation programs, or even for assessing their effectiveness once implemented. Extensive hardware sub-metering, although very expensive, is sometimes used to obtain more granular feedback. Non-Intrusive Load Monitoring (NILM), another method that has been studied for the past two decades, follows an inexpensive approach for obtaining appliance-specific consumption information. The idea behind this technique is that operation of individual appliances generates a distinct signature in the power distribution system of the building, which can be detected by carefully analyzing the overall voltage and current of the building. However, two of the main challenges keeping the technology from reaching wide adoption are: (a) finding simple ways to train the algorithms; and (b) obtaining robust appliance signatures that form spread-out clusters in the feature space, especially for small loads.

In this paper we explore the feasibility of utilizing data from separate environmental sensors (e.g., light intensity, sound level, etc.) present in the building, for improving the training process by enhancing the appliance signatures and providing an independent and trusted source of information about the operation of appliances. We exploit the fact that the operation of appliances will likely be reflected in both the power and environmental data streams. We present initial results from a case study where a prototype NILM system was deployed in an occupied apartment building, along with a number of environmental sensors. We also suggest two approaches for leveraging the environmental data and provide descriptions for possible future research in the area.

Keywords: non-intrusive load monitoring, energy efficiency, building automation.

1 Introduction

The idea of sustainable or green buildings generally stops after the design and construction phases. Little effort is made to continuously monitor and control the energy profile throughout the life-cycle of these facilities. Building owners and managers, when making energy-saving decisions, typically rely on a simple monthly bill from their utility summarizing their consumption. In the case of electricity, because of the variety of electrically-powered loads present in a typical building, these values are not informative about the effects of individual appliances or equipment on the total consumption. Thus, objective decisions about how to manage these loads to optimize the consumption are hard to make with aggregate monthly data. Given that residential and commercial buildings account for as much as 73% of the total electricity use in the U.S. (Energy Information Administration 2008), a solution to this problem is worth pursuing.

Real-time feedback has been shown to reduce consumption by 5-15% (Darby 2006), especially appliance-level feedback (Fischer 2008). Obtaining such measurements can become expensive with currently available commercial products due, in part, to having to separately measure each load in the building (Kim, Schmid, Srivastava et al. 2009). Non-Intrusive Load Monitoring (NILM) is a technique for obtaining appliance-level information from measurements taken at a single location: the main electrical feed of the building. This approach, which relies on signal processing and machine learning techniques, can be used as a lower-cost solution to obtaining granular feedback.

NILM traces its origins back to the late 1980s when it was first investigated (Hart 1989; Hart 1992). The early work focused on using changes in the total real and reactive power of residential buildings as signatures for individual appliance state-transitions. In that way, purely resistive loads such as heating coils or incandescent light bulbs could be distinguished from loads with capacitive or inductive components. Follow-up work extended the approach by extracting useful information from appliance start-up transients (Leslie K. Norford & Steven B. Leeb 1996; Steven Leeb et al. 1993; Khan et al. 1997) and from the harmonic content of the signals (Laughman et al. 2003). Many other improvements have been made in recent years, in particular towards applying the technique to commercial buildings, shipboard systems (Jones 2008; Ramsey et al. 2004), and automobiles (S. R. Shaw et al. 1999); as well as extending it for fault detection and diagnosis (Cox IV 2006). However, despite more than two decades of research in the area and while some researchers claim the technological problems for residential settings have been solved (Leslie K. Norford & Steven B. Leeb 1996), the technique has not resulted in wide commercialization and remains mostly a research topic.

To understand the reasons for this we have implemented prototype systems (Berges et al. 2009) based on the techniques described in the literature and have found that, in agreement with findings from other researchers (EPRI 1997; Kim, Schmid, Srivastava et al. 2009), two of the main challenges are: (a) finding simple ways to train the algorithms; and (b) obtaining robust appliance signatures that form spread-out clusters in the feature space, especially for small loads. In this paper we explore the effectiveness of leveraging data obtained from environmental sensors installed inside different spaces of the building as a solution for these problems.

1.1 Paper Organization

In section 2 we review the techniques we have employed to automatically disaggregate the total power of a building into the consumption of individual appliances. Particularly, we expand on the challenges mentioned above by providing concrete examples, and describe in more detail the concept of appliance signatures. Section 3 describes the systems used for collecting environmental data as well as introducing the experiments that will be performed. Then, in section 0 we explore different techniques for incorporating environmental sensor data into both the training and the classification processes. Early results from the prototype system are presented as a proof of concept. The last section is devoted to conclusions and a discussion of possible future work.

2 Appliance-level information from aggregate power measurements

Figure 1 shows the total real power of a residential apartment unit for a period of approximately 2.5 hours. During this time, the refrigerator and the central air unit fan (air blower) were switching on and off, as indicated in the figure. It is easy to see how the different appliances have different effects on this signal (e.g., the refrigerator has a higher turn-on spike than that of the air blower). Extending this concept to other power-related signals (e.g., reactive power), should also be intuitive. The main idea behind NILM is to automate the process of identifying these appliance state-transitions by making use of signal processing and machine learning techniques.

In our implementation of a NILM system, we used a four-step approach consisting of: (a) data collection; (b) pre-processing; (c) event detection and (d) classification. During the first two, overall voltage and current are sampled at close to 10kHz to compute power metrics. An event detector similar to the one described in (Luo et al. 2002) was used to detect abrupt changes of more than 30 watts in the mean of the real-power measurements, which we assume would indicate an appliance state-transition. Once an event (e.g., the on/off changes as indicated by the circles in Figure 1) was detected, a fixed-size window of real and reactive power samples around it were used to obtain the signature of the appliance state-transition. Coefficients of a kernel regression on these windowed signals were used as signatures, and a nearest neighbour approach in Euclidean space was used to perform the classification. More details of the system can be found in (Berges et al. 2009).

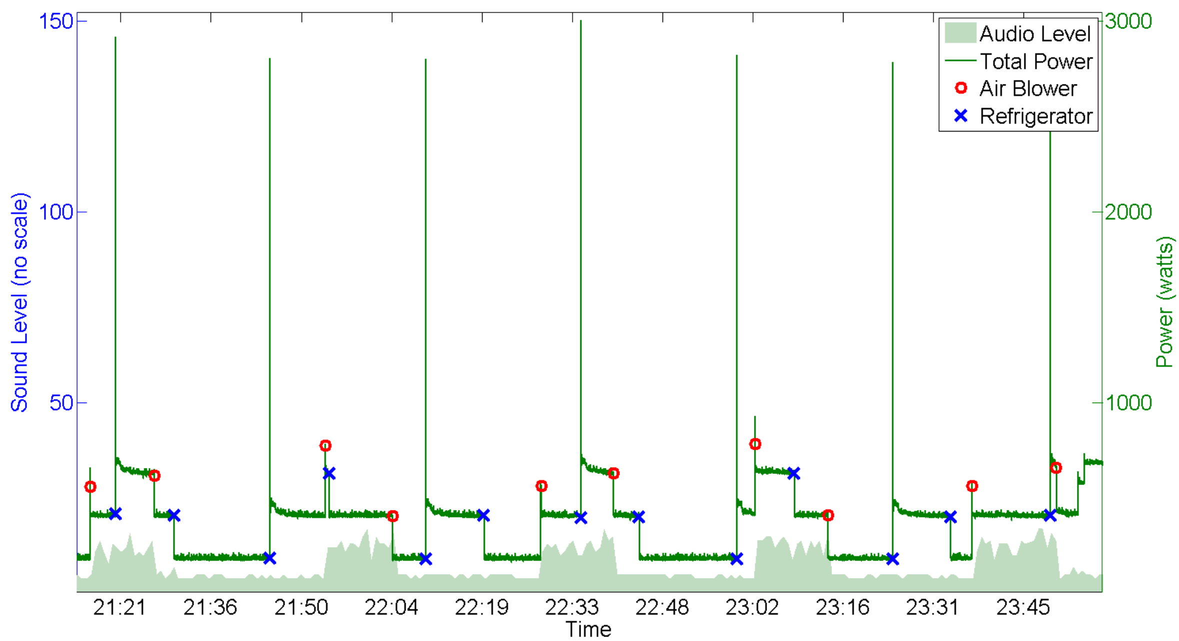

Figure 1, Measurements of the overall power consumption of an apartment unit with labelled events, overlaid on top of audio level measurements taken from a sensor in the living room.

When first deployed, the system needed to be trained on all the appliances of interest present in the building. We used a manual training process through which one person would operate each appliance (i.e., change its state), trigger a detection by the event detector, and then use a graphical user interface to provide an appropriate label to it. For the 800 sq. feet unit where we deployed the prototype system, which had 38 appliances of interest, this process took approximately 1.5 hours. Automating this process would greatly increase the opportunities of the technology to move out of research settings, into real homes. Towards this end, the authors developed a wireless clip-on meter (Berges et al. 2009). In the following sections we will describe another possible approach to achieve this.

Another challenge of NILM is that it is necessary to ensure that signatures from the same appliance state-transition can be separated from others in feature space. This condition is intrinsically linked to the choice of features used to compose the signatures, as well as to the similarity metric utilized to compare them. The first generation of NILM systems utilized changes in real reactive power (two features) as signatures and Euclidean distances to compare them (Hart 1992). This gave good results for large loads in a residential building, but other features such as the shape of the start-up transient and power-harmonics were needed to extend the approach to commercial buildings and other appliances (Laughman et al. 2003). Adding these new features increased the computational requirements of the metering system and with it the cost. We believe that data from environmental sensors present in the building, which are indirectly measuring the operation of appliances, could be leveraged to add new features to the signatures, at a low additional cost.

3 Leveraging environmental sensor-data

Research on leveraging heterogeneous sensor data for understanding the operational schedule of individual appliances in a facility, as applied to electricity metering, is relatively new. Recent technological advances in sensor networks and computing have reduced the costs of performing pervasive instrumentation of the physical world. This in turn, has opened the possibilities of utilizing sensor data about the environment as additional observations for solving a variety of problems. Some researchers have used the additional information to help disambiguate the context of electrical events (Lifton et al. 2007), for example. Others have proposed using such measurements as "indirect sensing" of an appliance's power consumption (Kim, Schmid, Charbiwala et al. 2009).

We outfitted the apartment unit where the NILM prototype system is installed with a wireless sensor network developed at Carnegie Mellon University (Rowe et al. 2006). The environmental sensor nodes deployed were capable of measuring light intensity, temperature, acceleration, and sound level, all represented as uncalibrated 8-bit values; and were polled for these values every 20 seconds. A number of plug-through power metering sensor nodes were also deployed as part of the network to obtain ground truth data and validate the predictions of the NILM system. More details on the sensor network can be found in (Rowe et al. 2009; Buevich et al. 2009). Figure 1 shows a chart for audio level, extracted from the web interface of the sensor network. Labels were superimposed to explain the source of the sound. As shown in Figure 1, the audio levels of a specific node have a good correlation with the operation of the air blower.

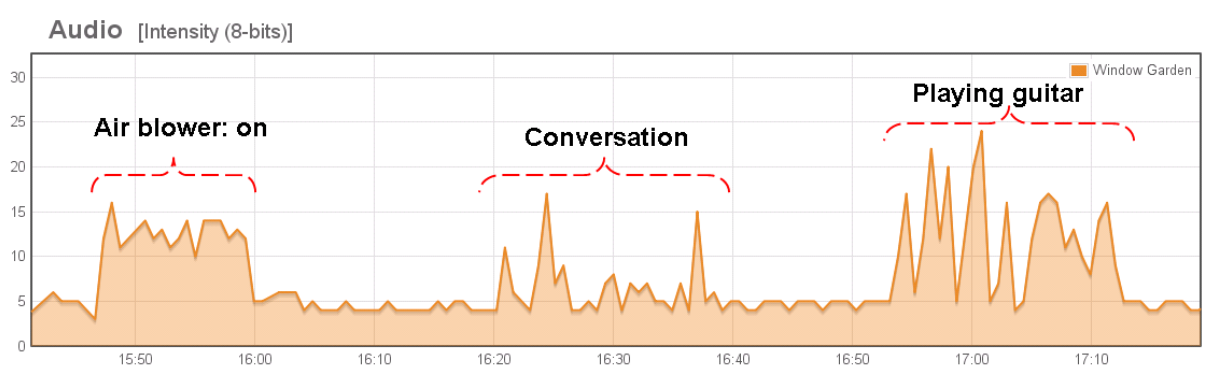

Figure 2, Audio intensity values from a sensor node placed near a window of the apartment. Labels are provided as explanations for the different audio patterns.

Our intention was to use the additional information to enhance the appliance signatures that the NILM utilizes, by treating these new streams of data as additional metrics besides the real and reactive power of the home. In this way, features can be extracted from all signals, once an event has been detected, and used to compose the signatures for an appliance state-transition. This process is better explained in the following section.

4 Merging different data-sources

In the approach presented in (Kim, Schmid, Charbiwala et al. 2009), a "fusion center" receives the signals from all the sensors and solves a numerical optimization problem which tries to find sensor calibration parameters that result in an overall power consumption that matches the one measured by the utility. For this to work, the appliance being *indirectly sensed* by each sensor must be made explicit (e.g., the user needs to specify which lamp a light intensity sensor is monitoring), given that there is no provision in the framework for automatically classifying appliances. We propose, instead, to treat the environmental sensor streams indifferently, and to not associate each one with a specific appliance. We exploit this new, more agnostic approach in two ways, explained below.

4.1 Multi-modal signatures

Incorporating information from the environmental streams into the appliance signatures can expand the feature space and help the system to better disambiguate. We would like for these new signatures to reflect the changes that the operation of appliances have on both the electric power and the other environmental phenomena. One way of achieving this is to apply the same process of obtaining kernel regression coefficients to these new signals. In essence, this would amount to finding the weights (ω) that solve Eq. 1, where Y is a matrix of size L x M containing sections of M signals, of L samples each; $\varphi(x)$ is the basis function and maps x into a higher dimensional space; and x is a vector of length L whose entries are the integers from 1 to L. The signals contained in Y can be power metrics and environmental streams, although they all need to be L samples long which can be achieved by re-sampling when needed.

$$Y = \varphi(x) \cdot \omega \tag{1}$$

One way to solve this equation is by applying a pseudo-inverse on $\varphi(x)$ and multiplying both sides by it. The resulting signature (ω) would summarize the effect of the appliance state-transition on all signals (environmental and power) without us having to explicitly model the relationship between each sensor and the appliance. For example, by including data from a light intensity sensor placed inside the bathroom, the overhead light can be more easily distinguished from the kitchen light even though their turn-off power trace is very similar as can be seen in Figure 3. In the same way, changes in audio levels from a sensor placed in the living room would help distinguishing the signature of the air blower unit, as shown in Figure 1.

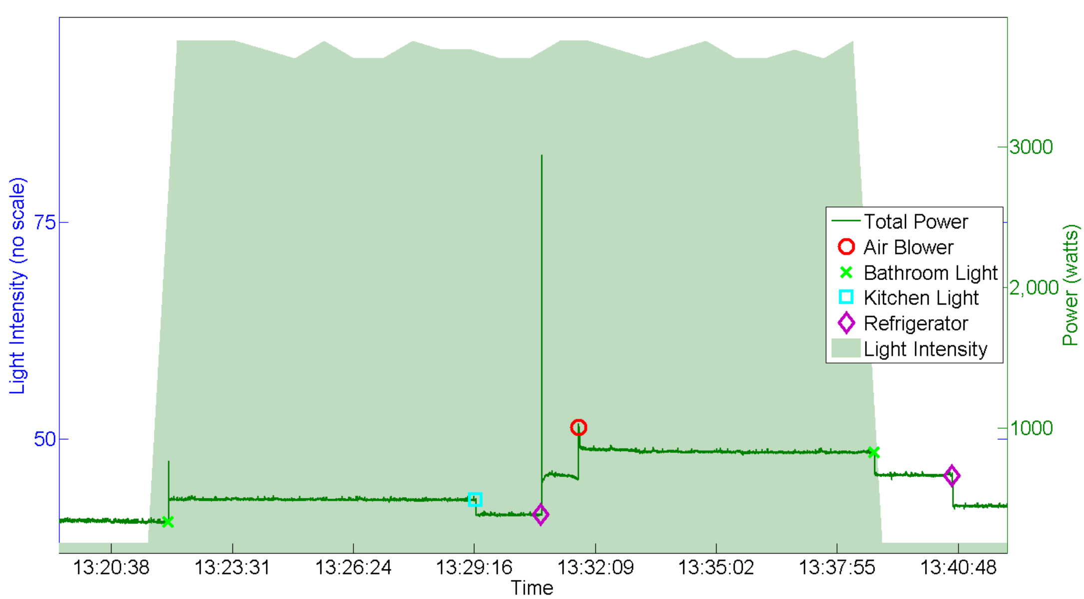

Figure 3, Light intensity measurements from a sensor in the bathroom, and power trace of the whole house overlaid on top.

4.2 Continuous automated training

If there are environmental signals that are strongly correlated with the operation of one individual appliance, then the system could exploit this feature and use it to automatically re-train itself when needed. For example, as shown in Figure 3, the light intensity sensor placed in the bathroom (shaded area) is highly correlated with the operation of the overhead light in the same part of the building (x marks), but not with the other appliances showed in the graph. If at some point in time the light bulbs are replaced with different types, the power signature of the light will change, but the relationship between the light intensity measurements and the operation of the light would remain. Thus, the system could potentially recognize this situation and adjust itself accordingly.

By running an event detection algorithm on all environmental signals the NILM system could try to find events that occur simultaneously or very close to each other, in order to determine which signals are correlated. It could then establish a relationship between an environmental stream and an appliance and store this information for future use.

In contrast with the multi-modal signatures, this approach for leveraging environmental sensor data tries to keep the information contained in the new streams separate from the power signatures. In this way, the new information can be used as a trusted source after determining the relationship that exists between appliances and the environmental phenomena. Once this trusted relationship has been established, the events detected on either power or environmental data can be used to confirm one another. One possible benefit of this is that the system can continuously assess its performance in an automated fashion, and could even use this as additional training data. Beyond these issues, environmental sensors are becoming more pervasive in homes from new devices (e.g., sound sensors in home security systems) and the potential to leverage these already deployed sensors is compelling.

5 Conclusions and Future Work

We have presented new ideas for making use of environmental sensor data to improve the performance of non-intrusive load monitoring algorithms. A prototype system to evaluate their feasibility was developed and installed in an occupied apartment unit. The prototype consisted of both a NILM system and a wireless sensor network capable of monitoring four environmental parameters in different parts of the dwelling.

The data obtained from this early prototype revealed that there is a good potential for extracting relevant appliance-related information from the environmental sensors. In contrast with other approaches, we propose to treat the new sensor streams as uncalibrated and unlabeled signals, and simply extract the relationships that exists between these new streams and the operation of appliances by analyzing the raw data. Two possible applications of this approach were discussed, but experiments need to be performed in order to validate their effectiveness.

A first logical step following the work presented here is to implement the multi-modal signature extraction algorithm described in section 4 and evaluate the performance of the system by comparing it with the power-only approach. Similarly, the ideas for using the environmental measurements as trusted sources after establishing a relationship between power events and environmental ones, should also be tested.

A more careful examination into the time-synchronization problem that arises when merging data from different sources should also be performed. Furthermore, even though we did not discuss about the interoperability issues related to acquiring the data from different sensing platforms, it is easy to see how this could prove to be a considerable obstacle when trying to deploy these systems in other real-world scenarios. Finally, a careful analysis on how to simplify the initial deployment of systems like these, from a human-computer interaction standpoint, should be made.

Acknowledgements

The authors would like to gratefully acknowledge the support from the Robert Bosch LLC Research and Technology Center North America and the National Science Foundation (NSF) grant #09-30868. The opinions expressed herein are those of the authors and not of the NSF.

References

BERGES, M. et al., 2009. Learning Systems for Electric Consumption of Buildings. In *2009 ASCE International Workshop on Computing in Civil Engineering*. Austin, Texas, USA.

BUEVICH, M. et al., 2009. Sensor Andrew Gateway Agent (SAGA). *SAGA*. Available at: http://www.andrew.cmu.edu/user/agr/projects/saga/ [Accessed December 19, 2009].

COX IV, R.W., 2006. *Minimally intrusive strategies for fault detection and energy monitoring*. Massachusetts Institute of Technology.

DARBY, S., 2006. *The effectiveness of feedback on energy consumption*, Oxford, UK: Environmental Change Institute, University of Oxford. Available at: http://www.eci.ox.ac.uk/research/energy/electric-metering.php [Accessed August 17, 2009].

ENERGY INFORMATION ADMINISTRATION, 2008. *Annual Energy Review*, Energy Information Administration. Available at: http://www.eia.doe.gov/emeu/aer/pdf/aer.pdf [Accessed February 18, 2009].

EPRI, 1997. *Low-Cost NIALMS Technology: Market Issues & Product Assessment*, Palo Alto, California: Electric Power Research Institute. Available at: http://mydocs.epri.com/docs/public/TR-108918-V1.pdf [Accessed April 30, 2009].

FISCHER, C., 2008. Feedback on household electricity consumption: a tool for saving energy? *Energy Efficiency*, 1(1), 79-104.

HART, G., 1992. Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, 80(12), 1870-1891.

HART, G., 1989. Residential energy monitoring and computerized surveillance via utility power flows. *Technology and Society Magazine, IEEE*, 8(2), 12-16.

JONES, R.A., 2008. *Improving shipboard applications of non-intrusive load monitoring*. Thesis. Available at: http://dspace.mit.edu/handle/1721.1/44843 [Accessed December 16, 2009].

KHAN, U., LEEB, S. & LEE, M., 1997. A multiprocessor for transient event detection. *Power Delivery, IEEE Transactions on*, 12(1), 51-60.

KIM, Y., SCHMID, T., CHARBIWALA, Z.M. et al., 2009. ViridiScope: design and implementation of a fine grained power monitoring system for homes. In *Proceedings of the 11th international conference on Ubiquitous computing*. pp. 245–254.

KIM, Y., SCHMID, T., SRIVASTAVA, M.B. et al., 2009. Challenges in Resource Monitoring for Residential Spaces. In *First ACM Workshop On Embedded Sensing Systems For Energy-Efficiency In Buildings*. Berkeley, CA.

LAUGHMAN, C. et al., 2003. Power signature analysis. *Power and Energy Magazine, IEEE*, 1(2), 56-63.

LEEB, S. et al., 1993. Development and Validation of a Transient Event Detector. *AMP Journal of Technology*, 3(November), 69-74.

LIFTON, J. et al., 2007. A platform for ubiquitous sensor deployment in occupational and domestic environments. In *Proceedings of the 6th international conference on Information processing in sensor networks*. Cambridge, Massachusetts, USA: ACM, pp. 119-127. Available at: http://portal.acm.org/citation.cfm?id=1236377 [Accessed December 18, 2009].

LUO, D. et al., 2002. Monitoring HVAC Equipment Electrical Loads from a Centralized Location - Methods and Field Test Results. *ASHRAE Transactions*, 108(1), 841-857.

NORFORD, L.K. & LEEB, S.B., 1996. Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms. *Energy and Buildings*, 24(1), 51-64.

RAMSEY, J.S. et al., 2004. *Shipboard applications of non-intrusive load monitoring*, FLORIDA STATE UNIV TALLAHASSEE DEPT OFPHYSICS.

ROWE, A., MANGHARAM, R. & RAJKUMAR, R., 2006. *FireFly: A Time Synchronized Real-Time Sensor Networking Platform, Wireless Ad Hoc Networking: Personal-Area, Local-Area, and the Sensory-Area Networks*, CRC Press Book Chapter.

ROWE, A. et al., 2009. Demo abstract: The Sensor Andrew infrastructure for large-scale campus-wide sensing and actuation. In *Proceedings of the 2009 International Conference on Information Processing in Sensor Networks - Volume 00*. IEEE Computer Society, pp. 415-416.

SHAW, S.R. et al., 1999. Non-intrusive load monitoring and identification in an automotive environment. In Lisboa, Portugal: IMACSFRANCE.