ROBUST ADAPTIVE EVENT DETECTION IN NON-INTRUSIVE LOAD MONITORING FOR ENERGY AWARE SMART FACILITIES

Yuanwei Jin and Eniye Tebekaemi

Engineering and Aviation Sciences
University of Maryland Eastern Shore
Princess Anne, MD 21853
{yjin, etebekaemi}@umes.edu

Mario Berges and Lucio Soibelman[*]

Civil and Environmental Engineering
Carnegie Mellon University
Pittsburgh, PA 15213
{marioberges, lucio}@cmu.edu

ABSTRACT

This paper presents a robust adaptive goodness-of-fit (GOF) χ^2 test event detector for non-intrusive load monitoring applications. We derive a closed form for the decision threshold and a guideline for choosing the size of the detection data window. Using real-world power data collected in residential buildings, the proposed GOF detector shows superior performance compared to the conventional generalized likelihood ratio detector.

***Index Terms*—** Event Detection, Goodness-of-Fit

1. INTRODUCTION

Nonintrusive load monitoring (NILM) is an emerging technology that can disaggregate individual electrical loads due to various household appliances in individual buildings from measurements made at a centralized location, such as the electric utility service entry. Knowledge of electricity consumption and time of use in individual buildings is vital to consumers and utility companies. For utility service providers, this information provides the basis for billing and payments, while for consumers, the utilities information helps monitor and reduce energy consumption in buildings. Furthermore, the electrical load usage due to appliances can be related to the aggregated behavior of individual human beings, which typically exhibits a periodicity in time on a number of scales (daily, weekly, etc) that reflects the rhythms of the underlying human activity. Hence, NILM is an ideal platform for extracting useful information about electricity usage, daily human activity, and thus in turn enables potential changes of consumer behavior.

A typical electricity meter reports aggregated usage information, which is not very useful in determining the best method for energy conservation. On the other hand, although a metering system that explicitly meters individual circuits and appliances can provide much more detailed information regarding load energy consumption, the required additional hardware and sensors would be prohibitively expensive. In contrast to other systems, NILM reduces sensor costs by using relatively few sensors. From measurements of the voltage and aggregate current at the utility service point, NILM disaggregates and reports the operation of individual electrical loads such as lights and motors. The current and voltage data is often corrupted by a number of bursty periods of unusual behavior due to degrading or abnormal load conditions for appliances on a household electricity network. Hence, applying signal processing and data mining techniques in NILM applications for finding and extracting these anomalous events becomes difficult because of these elements, [1, 2].

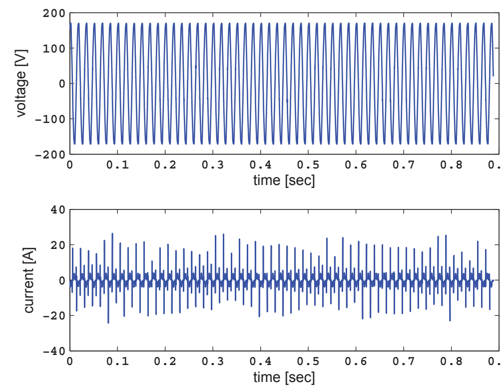

Fig. 1. Measured voltage and current signals over time.

In this paper, we develop a framework for robust adaptive appliance event detection. We consider the problem of event detection based upon near real-time power data stream that characterizes appliance activities in a household electricity grid. The goal of event detection is to raise an alarm after the onset of an event (i.e., on or off status of an appliance or appliance state-transitions), which would enable identification of the time-instant when the On or Off occurs. A simple way of addressing the event detection problem is to look for changes in the data stream and equate “change” with “onset of event”. Changes in the data stream can be detected by comparing the distribution of the most recent observations (the current set) with the distributions of previous observations (the reference set). This is often called “change detection.” Various methods for detecting changes in data streams have been proposed (see e.g., [3]). In the NILM application, conventional event detector often time requires periodic training to adjust the detection threshold due to the dynamics of electrical loads in order to achieve a high detection probability and a low false alarm rate. This condition imposes severe limits on the achievable accu-

*This joint work is supported in part by the National Science Foundation GOALI/CPS program under award no. CNS-093-868. Y. Jin was also supported in part by a 2010 Air Force Summer Faculty Fellowship award.

racy of the event detector, thus reducing the practical usage of a NILM system. This paper develops a robust adaptive detection scheme to detect appliance events with limited training while achieving high detection accuracy.

2. LOAD POWER CALCULATION

In an AC electric power system, harmonics have always been presented. Harmonics refer to spectral components present in a voltage or current waveform, whose frequencies are integer multiples of fundamental frequency (i.e., 60 Hz in the US) of the voltage form. Harmonic currents are created by nonlinear loads, such as variable speed drives (VSDs), electronic ballasts for fluorescent lighting, switching power supplies, rectifiers. In general, we may define the instantaneous voltage and current as follows:

$$v(t) = V_0 + \sum_{k=1}^{\infty} V_k \cos(k\omega t + \phi_{v_k}) \quad (1)$$

$$i(t) = I_0 + \sum_{k=1}^{\infty} I_k \cos(k\omega t + \phi_{i_k}) \quad (2)$$

where V_0, I_0 is the average value. V_k, I_k is the amplitude of the k-th harmonic of voltage and current, respectively. By integrating the instantaneous power $p(t) = v(t)i(t)$ over a window period T, where T is typically one or more periods of the fundamental frequency of the voltage waveform, we obtain the average power

$$P_{ave}(t) = \frac{1}{T}\int_{t-T}^{t} p(\tau)d\tau \quad (3)$$

$$= V_0 I_0 + \frac{1}{2}\sum_{k=1}^{\infty} V_k I_k \cos(\phi_{v_k} - \phi_{i_k}) \quad (4)$$

We note that (4) is related to many other power calculations reported in the literature. For example, the IEEE working group on power systems with nonsinusoidal waveforms and unbalanced loads, [4], recommends the following fundamental active power, $P_1(t)$, by separating the fundamental component V_1 and I_1 from the total voltage and current,

$$P_1(t) = \frac{1}{2} V_1 I_1 \cos(\phi_{v_1} - \phi_{i_1}) \quad (5)$$

Fig. 1 depicts a snapshot of experimentally measured voltage and current signals caused by activities of appliances. The plots clearly show that the voltage signal has a dominant first-harmonic component, V_1 and zero DC value, V_0. Hence we can approximate the voltage signal as

$$v(t) \approx \frac{1}{\sqrt{2}} V_1 \cos(\omega t + \phi_{v_1}) \quad (6)$$

Under this approximation, we can show that $P_{ave}(t) \approx P_1(t)$, i.e., the average power calculation given in (4) is consistent with the commonly used *active power* in the power industry. Furthermore, we should note that in a practical NILM system, $P_{ave}(t)$ or $P_1(t)$ varies over time t because of the continuous changes in the appliance operation status.

3. EVENT DETECTION

We consider the problem of event detection based upon the continuous power data stream collected from whole-house power meters. The data stream is recorded and divided into blocks (or windows) of n-samples. The goal of event detection is 1) to raise a flag after the onset of an event (i.e., on or off status of an appliance or other appliance state-transitions), and 2) to identify the time-instant where the change of appliance status occurs. The proposed event detection consists of two steps: 1) detect an event within each data block (or window) and 2) locate the time-instant of change for the event. Next, we discuss the data signal model for detection.

3.1. Signal Data Model

From the calculated power data given in (4), we model the discrete average power signal calculated as

$$x_i = e_i + w_i, i = 1, 2, \cdots, n \quad (7)$$

where w_i is the disturbance in power measurement and is assumed to be distributed as a white Gaussian process. The symbol e_i is considered as an indication that the data sample belongs to an appliance transition power signature. n is size of the observation window of the sampled power data. We should note that at a high sampling rate, the event signature signal e_i consists of the static state, transient state, and oftentimes a bursty period due to nonlinearities of the appliance load condition. Thus, the overall power signal behavior x_i is complicated and must be processed using statistical means. If there is no event, only white noise exists, which can be represented as $x_i = w_i$, $i = 1, \cdots, n$. The event detector operates on two sliding data windows defined as follows:

Pre-event window. The pre-event window is used as the reference for upcoming events. The pre-event window is defined as

$$W_{i,k} = \{x_n | i \leq n \leq k\} \quad (8)$$

Detection window. This is the window preceding the pre-event window. This is the working window of the GOF test. It is in this window we intend to detect the occurrence of an event, i.e., on or off of an appliance. The detection window is defined as

$$W_{l,m} = \{y_n | l \leq n \leq m\} \quad (9)$$

where $l = i + n$, and $m = k + n$. Next, we develop a framework for event detection based on the χ^2 test for goodness-of-fit. The parameters of detectors of the detection window are compared to the pre-event window, and a decision is reached based on the outcome of a event detector.

3.2. χ^2 Test of Goodness-of-Fit (GOF)

The goodness-of-fit test seeks to determine whether a set of data could reasonably have originated from some given probability distribution. Assume that we have n independent and identically distributed (iid) random samples $x_i, i = 1, 2, \cdots, n$, drawn from a distribution $G(x)$, which is *a priori* unknown. We have a supposed distribution function $F(x)$. The problem can be formulated as the binary hypothesis testing problem

$$\begin{array}{ll} \mathbb{H}_1: & G(x) \neq F(x) \\ \mathbb{H}_0: & G(x) = F(x) \end{array} \quad (10)$$

GOF tests will allow deciding between the two hypotheses in (10). In event detection, we will explain the GOF problem differently: There exist two sets of iid samples. The reference set (i.e., the pre-event window data set) consists of n samples x_i, $i = 1, \cdots, n$, with the distribution $G(x)$. The test set (i.e., the detection window data set) consists of n samples y_i, $i = 1, \cdots, n$ with distribution $F(y)$. Both $G(x)$ and $F(y)$ are unknown. The goal of GoF tests is to decide between the two hypotheses of (10). If the null hypothesis $\mathbb{H}_0$ is rejected, we claim that an appliance event occurs.

Among various goodness-of-fit tests, the χ^2 test has been widely used in statistics literature. In a standard application of the χ^2 test, the test procedure requires a random sample of size K from a population whose probability distribution is unknown. These K observations are arranged in a frequency histogram, having n bins or class intervals. Let $p_i(i = 1, \cdots, n)$ denote the probability of an observation falling into the ith bin, and y_i be the observed frequency in the ith bin, the test statistic is

$$X^2 = \sum_{i=1}^{n} \frac{(y_i - Kp_i)^2}{Kp_i}. \tag{11}$$

where the quantity Kp_i is the expected frequency in the i-th bin. If the observed frequency satisfies the supposed distribution, these observed values follow a multinomial distribution with p_i being probabilities, [5]. In power signal event detection, we consider the observation in the detection window $y_i \sim F(y)$ be the observed frequency in the i-th time-instant within the detection window [6]. Because the quantity Kp_i is unknown, it will be estimated from the data samples $x_i \sim G(x)$ in the pre-event window. We can show that the maximum likelihood estimate of Kp_i based on the pre-event window data $x_i, i = 1, \cdots, n$, is given by,

$$\widehat{Kp_i} = x_i. \tag{12}$$

Inserting (12) into (11), we obtain the χ^2 test for goodness-of-fit

$$\ell_{\text{GOF}} = \sum_{i=1}^{n} \frac{(y_i - x_i)^2}{x_i} \tag{13}$$

We would reject the $\mathbb{H}_0$ hypothesis that the distribution of the population is the hypothesized distribution if the calculated value of the test statistic [5]

$$\ell_{\text{GOF}} > \chi^2_{\alpha, n-1} \tag{14}$$

with $100(1-\alpha)\%$ confidence interval and $n-1$ degrees of freedom. We should note that $\chi^2_{\alpha,n-1}$ is the decision threshold that depends on the window size n and the detection confidence level α. This observation raises a question: how to determine the proper window size n?

Note that in the data model (7), the disturbance w_i implies the absence of event and is assumed to be a Gaussian process with mean μ_w and variance σ_w^2. These two quantities are unknown and need to be estimated from (pre-event) training data of n samples. If we use the sample mean $\bar{w} = \frac{1}{n}\sum_{i=1}^{n} w_i$ to estimate μ_w, we can be $100(1-\alpha)\%$ confident that the error $|\bar{w} - \mu_w|$ will not exceed a specified amount E when the minimum sample size is, [7]

$$n_0 = \left(\frac{z_{\alpha/2}\sigma_w}{E}\right)^2 \tag{15}$$

where $z_{\alpha/2}$ is the upper $100\alpha/2$ percentage point of the standard normal distribution. The quantity E can be chosen by the users. For example, if we decide to discard appliance events that are less than 30 Watts, we could set $E = 30$. Furthermore, the maximum window size, n_1, of the detection window should be limited by the maximum length of the state-transient transient of appliance signatures. Thus, we obtain

$$n_0 < n < n_1 \tag{16}$$

The importance of Eqns. (14) or (16) in GOF event detection is that they provide a guideline for choosing the window size, and then the decision threshold, based on no-event training data. Once the window size is chosen, repeated training or a data-dependent threshold becomes unnecessary. This is a significant advantage compared with the conventional generalized likelihood ratio test which will be discussed below.

3.3. Conventional Generalized Likelihood Ratio Test

The generalized log-likelihood ratio test (GLR) detector for k-th data block is defined as

$$\ell^k_{\text{GLR}} = \ln\frac{f(\mathbf{x}^k; \hat{\mu}_1, \hat{\sigma}_1^2)}{f(\mathbf{x}^k; \hat{\mu}_0, \hat{\sigma}_0^2)} = \sum_{j=1}^{n} \ln\frac{f(x_j^k; \hat{\mu}_1, \hat{\sigma}_1^2)}{f(x_j^k; \hat{\mu}_0, \hat{\sigma}_0^2)} \tag{17}$$

where the kth data vector $\mathbf{x}^k = [x_1^k, \cdots, x_n^k]^T$ are modeled as the random realizations of a Gaussian normal process. $\hat{\mu}_i, \hat{\sigma}_i^2$ are the maximum likelihood estimates of the mean and variance under hypothesis $\mathbb{H}_i$, respectively. Notice that the normal probability density function is

$$f(x_j^k; \hat{\mu}_i, \hat{\sigma}_i^2) = \frac{1}{\hat{\sigma}_i\sqrt{2\pi}} e^{-(x_j^k - \hat{\mu}_i)^2/2\hat{\sigma}_i^2} \tag{18}$$

Hence by straightforward derivation, we obtain

$$\ell^k_{\text{GLR}} = \sum_{j=1}^{n} \left(\frac{(x_j^k - \hat{\mu}_0)^2}{\hat{\sigma}_0^2} - \frac{(x_j^k - \hat{\mu}_1)^2}{\hat{\sigma}_1^2} \right) \tag{19}$$

where the maximum likelihood estimates are obtained from the pre-event window and the detection window as follows:

$$\hat{\mu}_1 = \frac{1}{n}\sum_{i=1}^{n} y_i \qquad \hat{\sigma}_1^2 = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{\mu}_1)^2 \tag{20}$$

$$\hat{\mu}_0 = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \hat{\sigma}_1^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \hat{\mu}_0)^2 \tag{21}$$

Using the GLR, we would reject the null hypothesis $\mathbb{H}_0$ if $\ell^k_{\text{GLR}} > \eta^k_{\text{GLR}}$, where η^k_{GLR} is the decision threshold for k-th data block. Due to the dynamic behavior of the appliance load conditions, we emphasize that this decision threshold shall be obtained based on periodic training in order to achieve a high correct detection rate and a low false alarm rate.

Table 1. Labeled Appliance Events

ID	Appliance Event Description
1	Microwave went to On
2	Microwave went to Off
3	Oven went to On
4	Oven went to Off
5	Lamp #1 in living room went to On
6	Lamp #1 in living room went to Off
7	Lamp #2 in living room went to On
8	Lamp #2 in living room went to Off
9	Lamp #3 in living room went to On
10	Lamp #3 in living room went to Off

Table 2. Detection Results

	GOF	GLR	GLR*
P_{D}	99.9%	72.7%	99.1%
P_{FA}	0.9%	0.58%	1.3%

4. EXPERIMENTAL RESULTS

A series of experimental week-long data-sets were collected in 3 different residential units. Through an initial training period, we obtained an initial set of state-transition transients for the appliances of interest in these homes. A set of plug-level power meters and other sensors were used to collect ground-truth data about the operation of the appliances during the period of study. From the measured voltage and current signals, we calculated the average power based on (4). We then conducted experiments on these power data using the developed event detector for performance evaluation.

Fig. 2 depicts a snapshot of power data with labeled appliance events over about a 250-second period. The label IDs for the appliance events (ID $1 \sim 10$) are summarized in Table 1. The detection results by the GOF and GLR detectors are marked with different symbols (□ for GOF and $\times$ for GLR, respectively). The results show that the GOF correctly detects all the events ID $1 \sim 10$, while the GLR fails to detect events $3, 7, 8, 9$, and 10. Note that event ID 11 is part of the state-transition transient signal for the oven (ID 4) but is detected as an event. Here, we consider this outcome as a false alarm since we focus on event detection. However, this error can be corrected in the subsequent classification stage using the turn-on state-transition signature waveform of the oven. Next, we calculate the correct detection rate P_{D}, defined as the ratio of the number of correctly detected events over the total number of events (ground truth), and the false alarm rate P_{FA}, defined as the ratio of the number of falsely detected events over the total number of events. The average detection rate and false alarm rate over 20 consecutive sets of collected power data from the 3 residential units are summarized in Table 2 for the three detectors: GOF (with a fixed threshold calculated by (14) and (16)), GLR (with a fixed threshold), and GLR* (with periodic training for adjusting threshold). The GOF shows the best performance in detection rate.

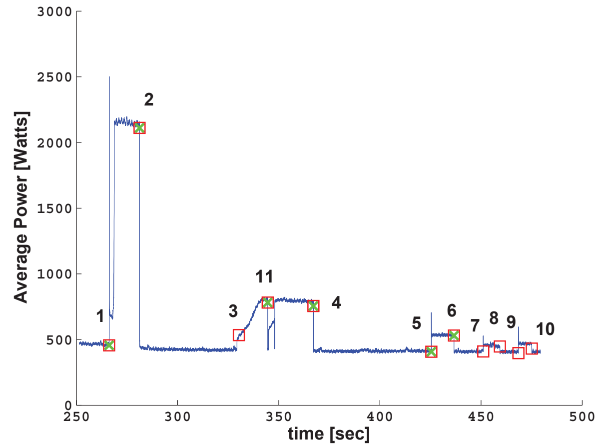

Fig. 2. Labeled and detected events by GOF (denoted by red □) and GLR (denoted by green $\times$).

5. CONCLUSION

In this paper we develop a robust adaptive goodness-of-fit χ^2 test for appliance event detection. The preliminary results based on the power data collected from three residential buildings show that the GOF test requires limited training and achieves a superior performance than the conventional generalized likelihood ratio detector.

6. REFERENCES

[1] C. Laughman, K. Lee, R. Cox, S. Shaw, S. Leeb, L. Norford, and P. Armstrong, "Power signature analysis," *IEEE Power and Energy Mag.*, pp. 56–63, March/April 2003.

[2] S. Shaw, S. Leeb, L. Norford, and P. Armstrong, "Non-intrusive load monitoring and diagnostic in power systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 7, pp. 1445–1454, July 2008.

[3] N. Vaswani, "The modified cusum algorithm for slow and drastic change detection in general hmms with unknown change parameters," in *ICASSP'05*, vol. 4. IEEE, 2005, pp. 701–704.

[4] "Practical definitions for powers in systems with nonsinusoidal waveforms and unbalanced loads: A decision (ieee working group on nonsinusoidal situations)," *IEEE Transactions on Power Delivery*, vol. 11, no. 1, pp. 79–87, January 1996.

[5] W. G. Cochran, "The χ^2 test of goodness of fit," *Annals of Math. Stat.*, vol. 23, pp. 315–415, 1952.

[6] H. Chernoff and E. L. Lehmann, "The use of maximum likelihood estimates in χ^2 tests for goodness of fit," *Annals of Math. Stat.*, vol. 25, no. 3, pp. 579–586, 1954.

[7] R. V. Hogg and A. Craig, *Introduction to Mathematical Statistics*. Upper Saddle River, NJ: Prentice Hall, 1994.