Load Disaggregation Using Harmonic Analysis and Regularized Optimization

Jerry T. Chiang* Tianzhu Zhang* Binbin Chen* and Yih-Chun Hu†
* Advanced Digital Sciences Center, Singapore
E-mail: {jerry.chiang, tz.zhang, binbin.chen}@adsc.com.sg
† University of Illinois at Urbana–Champaign, Urbana, IL, U.S.A.
E-mail: yihchun@illinois.edu

***Abstract*—In this paper, we present a load disaggregation technique that uses regularized optimization together with harmonic frequency signatures of appliances. The benefits of our technique are twofold: 1) The regularized optimization is faster than integer programming; and 2) The harmonic frequency signatures allow us to disaggregate the loads using as few as 10 cycles (equaling as little as 200 milliseconds) of samples, instead of having to wait for state changes from appliances or weekly usage pattern to emerge. We test our proposed technique in proof-of-concept experiments and show that our technique returns accurate disaggregation results.**

I. Introduction

With the rising cost and the pollution associated with fossil fuel, the safety concern of nuclear energy, and the inefficiency or inaccessibility of renewable energy sources, many governmental and private entities are focusing on reducing energy consumption in order to avert large-scale energy crisis. The United States Annual Energy Report estimates that the residential and commercial sectors (as opposed to the industrial sector) in the United States together account for over 70% of the total electricity consumption [1]. The Electric Power Research Institute further estimates that, if given the activity data of each household appliance, the residential sector can reduce its electrical consumption by 12% [2].

While it may be infeasible to install an activity monitor in each and every one of today's appliances, a smart-meter can reliably measure the total (aggregated) consumption of a household. Prior work has proposed to disaggregate the consumption measurements to obtain activity data of individual appliances [3], [4]: *Load disaggregation* is the task of measuring the aggregate power usage of multiple electric loads (e.g. all appliances in a household) over time, and determining the usage activity of the individual loads.

Hart describes several possible methods to derive the *signature* of each electric load, and subsequently use the load signatures to identify load activities [3]. In particular, Hart noted that the amount of real and reactive power consumed by each appliance in its steady state may differ greatly and can be used as its signature.

This work is supported by the research grant for the Human Sixth Sense Programme at the Advanced Digital Sciences Center from Singapore's Agency for Science, Technology and Research (A*STAR).

For example, in their steady states, an incandescent light bulb and an ice maker both consume around 250 watts of real power; the prior being a resistive load consumes nearly 0 reactive power while the latter consumes almost 200 volt-amperes of reactive power. When two electric loads are active, their power consumptions sum up; i.e. if both the light bulb and the ice maker are on, they would together consume a total of 500 watts of real power, and 200 volt-amperes of reactive power.

By searching all possible subsets of appliances, one can then determine the subset of appliances that are most likely active. However, the search space grows exponentially with the number of appliances, and an exhaustive search over all subsets is infeasible when many appliances are present. Prior studies suggest monitoring the power usage over time, and when there is a change in consumption, since the change is likely incurred by the state change of few appliances, the search space can be reduced significantly [3], [4]. While prior studies demonstrate that identification based on consumption change is computationally feasible and accurate, it comes at the expense of having to wait for a state change. If an appliance is never turned off (e.g. a security camera), we cannot identify it from the total electricity usage.

Hart notes that an electric load that is not purely resistive is likely to produce an assortment of harmonic currents; thus each appliance may have a unique *harmonic frequency signature*. In this paper, we explore using the harmonic frequency signatures of loads for disaggregation without having to wait for any consumption changes. The central idea of our protocol is to use regularized optimization to determine the *appliance state vector*, a vector whose k^{th} entry is a binary number indicating the activity state ("on" or "off") of the k^{th} appliance.

The rest of this paper is organized as follows: Section II presents our system model and our proposed disaggregation technique. We evaluate the proposed technique in Section III. We then provide a collection of related prior studies and conclude in Section IV and Section V, respectively.

II. Proposed Disaggregation Technique

A. System Assumptions

We assume there is no phase-shift between any pair of frequency-domain current measurements. Since a phase-shift in frequency-domain corresponds to a time-shift in the time-domain, this assumption is equivalent to assuming that any pair of time-domain current measurements are aligned in time, which is a standard assumption in prior disaggregation studies.

We also assume that the voltage is periodic with a period equaling the inverse of the *operating frequency*. If the voltage is aperiodic, then the electric current consumption most likely would also be aperiodic, and the harmonic current signature cannot reflect the steady state behavior of an appliance.

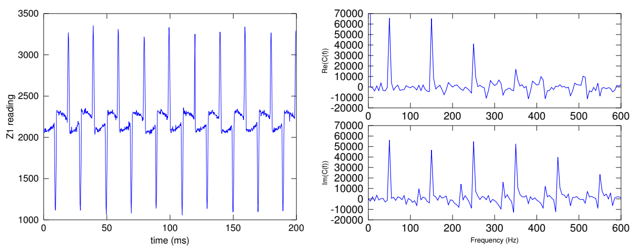
B. Definition and Problem Statement

In this section we define the terminologies and symbols used in this paper. Let N be the set of appliances, and $|N| = K$. Let $c_k(t)$ be the time-domain electric current measurement of the k^{th} appliance, and let $c_{\mathbb{K}}(t)$ denote the aggregated time-domain electric current measurement when only the set of appliances in $\mathbb{K} \subseteq 2^N$ are on. Without any time-offsets, Kirchoff's current law states that: $c_{\mathbb{K}}(t) = \sum_{k \in \mathbb{K}} c_k(t)$.

Let $C_k(f)$ be the fast Fourier transform (FFT) of $c_k(t)$: $C_k(f) = \mathfrak{F}\{c_k(t)\}$. Since the FFT is linear,

$$C_{\mathbb{K}}(f) = \sum_{k \in \mathbb{K}} C_k(f).$$

The *harmonic signature* of an appliance $H_k(f)$ is then a function of the frequency-domain representation of the current measurement: $H_k(f) = g(C_k(f))$. In our paper, we let $g(\cdot)$ be a sampling function that keeps F samples (corresponding to frequencies $f_1, \cdots, f_F$), and then takes the real part and concatenated with the imaginary part. The output of g is a $2F \times 1$ real vector that takes the form: $H_k(f) = [\mathfrak{Re}\,(C_k(f_1)| \cdots |C_k(f_F))\,|\mathfrak{Im}\,(C_k(f_1)| \cdots |C_k(f_F))]^T$, where $|$ represents concatenation.



(a) Time-domain measurement (b) Frequency-representation of the measurement

Fig. 1. Time-domain and frequency-domain representations of a load's current consumption

For example, if the measured time-domain current consumption of an appliance ($c(t)$) is as shown in Fig. 1(a), it's corresponding frequency-domain representation ($C(f)$) is shown in Fig. 1(b). Let $F = 3$, and $f_1 = 50$ Hz, $f_2 = 150$ Hz, and $f_3 = 450$ Hz, then the signature $H(f)$ equals to: $[65747, 65283, 326, 56140, 46682, 39856]^T$.

Since $g(\cdot)$, like the FFT, is a linear function,

$$\begin{aligned} H_{\mathbb{K}}(f) = g(C_{\mathbb{K}}(f)) &= g\left(\sum_{k \in \mathbb{K}} C_k(f)\right) \\ &= \sum_{k \in \mathbb{K}} g(C_k(f)) = \sum_{k \in \mathbb{K}} H_k(f). \end{aligned}$$

Let $\mathrm{I}_{k \in \mathbb{K}}$ be an indicator function that equals to 1 when the k^{th} load is on, and 0 when it is off. We define a $K \times 1$ appliance state vector S by equating it's k^{th} element to $\mathrm{I}_{k \in \mathbb{K}}$. We can then rewrite

$$H_{\mathbb{K}}(f) = [H_1(f)|H_2(f)| \cdots |H_K(f)]\, S.$$

The goal of disaggregation is then to correctly identify S.

C. Proposed Disaggregation Technique

It is standard to assume the subset of appliances most likely to be on is the subset of appliances that minimizes the *difference between the measured aggregate electric current consumption and the sum of the current signatures of the subset of appliances*. For a collection of on-off appliances, the disaggregation task is thus inherently a binary integer programming problem [8]:

$$\begin{gathered} \text{Find } \mathbb{K}^*, \\ \text{which minimizes } E = \left\| c_{\mathbb{K}}(t) - \sum_{k \in \mathbb{K}^*} c_k(t) \right\|. \end{gathered}$$

However, we would like to avoid integer programming since its complexity is generally exponential to the number of appliances aggregated.

We can approximate the solution using linear programming; however, there is generally not a unique solution, and the returned result may contain a large number of elements being neither 0 nor 1. To counter this drawback, we borrow the concept from sparse optimization and adds a *regularizer* to our objective function of the linear program that penalizes every non-zero element of the result. Depending on the weight of the penalty, the regularized optimization balances between the sparsity of the result (i.e. the number of appliances on) and the error. We then formulate our optimization as:

$$\begin{aligned} & \min \ \lambda||\widehat{S}||_1 + ||E||_1 \\ & \text{subject to } H_A(f) = [H_1(f)|H_2(f)| \cdots |H_K(f)]\widehat{S} + E \\ & \qquad \mathbf{0} \le \widehat{S} \le \mathbf{1}, \end{aligned}$$

where $\widehat{S}$ is our estimate of S, E is the error in estimation, $H_A(f)$ is the harmonic signature of the aggregated measurement (i.e. a corrupted version of $H_{\mathbb{K}}$), and λ is the weight of the L_1 regularizer.

Since the L_1-norm regularizer penalizes every non-zero element of $\widehat{S}$, the regularizer penalizes complicated solutions that uses multiple bases to fit a solution for which one basis would suffice. The regularizer thus tends to drive the solution ($\widehat{S}$) to the boundary of the solution space. Since in our formulation, an appliance is either on or off, we restrict the boundary space to $\{0, 1\}^K$, and the vertices form the space of desired solutions. We solve the above optimization problem using existing solver in the community[1] [5].

[1] http://web.eecs.umich.edu/~honglak/softwares/nips06-sparsecoding.htm

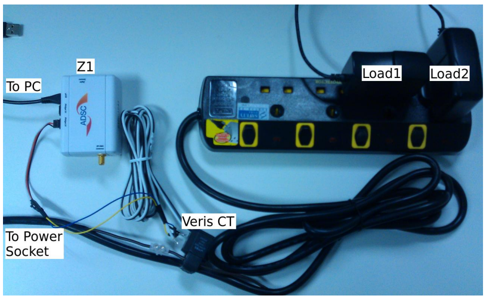


Fig. 2. Experimental setup including Veris current sensor, Z1 mote, and how we connect the loads

III. Evaluation

A. Methodology

As a proof-of-concept, we show how our protocol can be used to disaggregate appliances. We use a Veris split-core current sensor to transform the measured electric current into a voltage potential. We then use a Zolertia Z1 mote to digitize the voltage reading and send the reading to the computer. The Z1 mote can only measure positive voltage.

We perform two independent experiments. In the first experiment, we use three appliances that have similar power consumptions: a lamp, a fan, and a computer monitor. Similar power consumptions yield signatures that are not likely to shadow each other, and we expect our protocol to perform well. We also offset the voltage output of the Veris sensor so the Z1 mote is able to digitize the entire current waveform. In the second experiment, we consider a more adverse environment where one of the three appliances consumes significantly more power than others, and thus the signature of the powerful appliance is likely to shadow the signatures of other appliances. We also perform this experiment using only the half-wave rectified readings. Fig. 2 shows our experimental setup.

For our experiment, we first obtain the harmonic signature of each appliance by sampling its electric current 750 times at around 3400 hertz (corresponding to around 220 milliseconds of data). Fig. 3 and Fig. 4 show the electric current consumptions of each appliance in time-domain in the first and second experiments, respectively.

The time-domain representation is then transformed using FFT into the frequency-domain representation. We then sample the current's frequency representation at multiples of the electricity operating frequency (in Singapore, the operating frequency is 50 hertz; in the United States, 60 hertz). We perform our experiments in Singapore, sampling up to 10 harmonics in the frequency-domain, corresponding to using the harmonic currents at $f_1 = 50$ Hz, $f_2 = 100$ Hz, $\cdots, f_{10} = 500$ Hz. Since our sampling rate is at 3400 hertz, we can easily obtain the tenth harmonic without any aliasing. We let $\lambda = 10^5$, and perform 20 and 5 rounds of experiments for the first and second experiments, respectively.

B. Evaluation Results

For the r^{th} round of the i^{th} experiment, we obtain an $\widehat{S_{i,r}}$, each entry of which is between 0 and 1 with 0 indicating the "off" state and 1 indicating the "on" state. An entry that is close to 1 can be construed as *likely to be on*, and vice versa. We average the $\widehat{S_{i,r}}$ over the rounds for each $\mathbb{K}$ where $|\mathbb{K}| > 1$ (i.e., more than one appliance is on), and present the results in Table I and Table II. We also impose the decision rule that an appliance is "on" if its entry in the state vector is greater than 0.25.

TABLE I

S_1, AVERAGE OF $\widehat{S_1}$ OVER 20 ROUNDS, AND SUM OF DECISIONS OVER 20 ROUNDS. THE L COLUMN DENOTES WHETHER THE LAMP IS ON; THE F COLUMN, FAN; AND THE M COLUMN, COMPUTER MONITOR.

L	F	M	$\mathrm{E}\left[\widehat{\mathrm{L}}\right]$	$\mathrm{E}\left[\widehat{\mathrm{F}}\right]$	$\mathrm{E}\left[\widehat{\mathrm{M}}\right]$	$\sum \widehat{I_{\mathrm{L}}}$	$\sum \widehat{I_{\mathrm{F}}}$	$\sum \widehat{I_{\mathrm{M}}}$
0	1	1	0.019	0.799	0.968	0	20	20
1	0	1	0.891	0.030	0.974	20	0	20
1	1	0	0.945	0.749	0.072	20	20	0
1	1	1	0.917	0.831	0.952	20	20	20

TABLE II

S_2, AVERAGE OF $\widehat{S_2}$ OVER 5 ROUNDS, AND SUM OF DECISIONS OVER 5 ROUNDS. THE H COLUMN DENOTES WHETHER THE HAIR DRYER IS ON; THE F COLUMN, FAN; AND THE W COLUMN, WATER HEATER.

H	F	W	$\mathrm{E}\left[\widehat{\mathrm{H}}\right]$	$\mathrm{E}\left[\widehat{\mathrm{F}}\right]$	$\mathrm{E}\left[\widehat{\mathrm{W}}\right]$	$\sum \widehat{I_{\mathrm{H}}}$	$\sum \widehat{I_{\mathrm{F}}}$	$\sum \widehat{I_{\mathrm{W}}}$
0	1	1	0.008	0.974	0.94	0	5	5
1	0	1	1	0	0.958	5	0	5
1	1	0	0.862	1	0	5	5	0
1	1	1	0.98	0.2	0.848	5	1	5

The disaggregation result of the first experiment is very accurate. An "off" appliance results in less than 0.08 in the average disaggregation estimate, and is never determined to be on. An "on" appliance results in higher than 0.74 in the average estimate, and is always determined to be on.

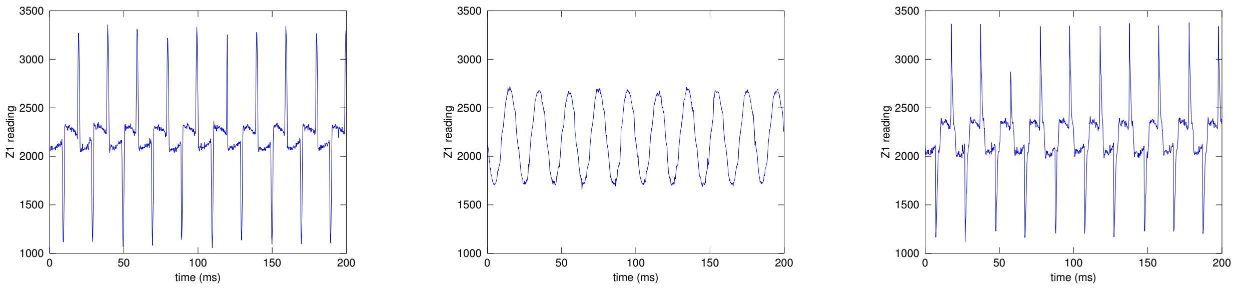
In the second experiment, we see that our protocol has its limitations: Since the water heater consumes a lot more power than the fan (see Fig. 4), its presence has a significant impact on the disaggregation performance. In particular, when all three appliances are on, the average estimated appliance state of the fan is only 0.2 (state correctly identified only one-out-of-five times). Besides this setting, all other cells of the table show that our technique returns accurate disaggregation results: an "off" appliance results in less than 0.1 in the average disaggregation estimate, and is never determined to be on; and an "on" appliance results in 0.8 or higher in the average estimate and is never determined to be off.

IV. Related Work

The research community has put in significant effort over the last two decades on load disaggregation.

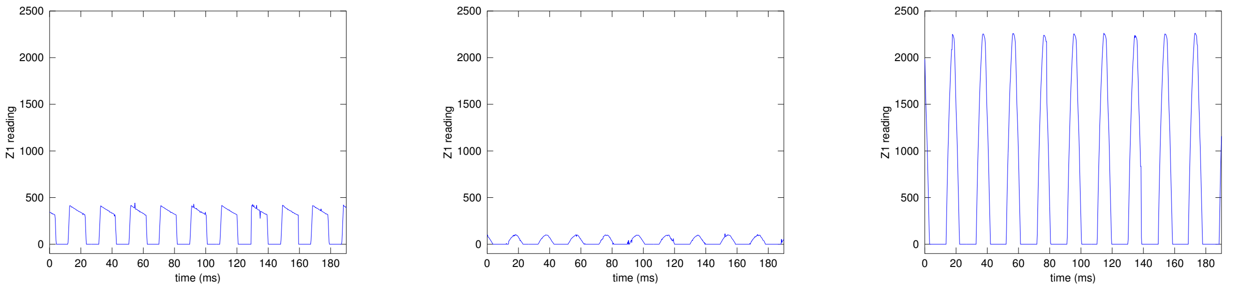
Hart's study presents numerous techniques in disaggregation [3]. Sultanem independently propose similar concepts [4]. Gupta et al. use a Universal Software Radio Platform to show that an appliance can generate much high-frequency electromagnetic interference (between 36 kilohertz to 500 kilohertz) that can also be used as part of the appliance's harmonic frequency signature [6].

Researchers have also sought to disaggregate loads using transient information. Leeb et al. proposed that when an



(a) Lamp (b) Fan (c) Computer Monitor

Fig. 3. Time-domain representation of each appliance's electric current consumption in our first experiment



(a) Hair dryer (b) Fan (c) Water Heater

Fig. 4. Time-domain representation of each appliance's electric current consumption in our second experiment

appliance switches on, its transient behavior is much different from its steady state behavior and presents another avenue of obtaining appliance signatures [7].

High resolution analysis using high resolution aggregate usage and not relying on appliance state changes promises to return disaggregation results faster than waiting for load profile changes. Suzuki et al. uses time-domain per-cycle data as the basis of formulating an integer programming optimization, the solution to which is the appliance state vector [8]. Our proposed technique is different in that our harmonic frequency signature is significantly smaller in size than the time-domain data, and our formulation of regularized optimization is able to further reduce the computation complexity compared to integer programming.

Some researchers have also adopted novel signal processing techniques for load disaggregation purpose. In particular, Kolter et al. propose using discriminative sparse coding to disaggregate loads using week-long time-domain load usage patterns [9]. Our proposed protocol differs from the work by Kolter et al. in that we do not try to identify any usage patterns, and thus we can reach the disaggregation result without having to collect weeks of aggregation data.

V. Conclusion

In this paper, we present a load disaggregation technique that uses regularized optimization together with harmonic frequency signatures of appliances. The benefits of our technique are twofold: 1) The regularized optimization allows us to perform the optimization faster; and 2) The harmonic frequency signatures allow us to disaggregate the loads using as few as 10 cycles (equaling to less than a quarter of a second) of samples, instead of having to wait for state changes from appliances or weekly usage pattern to emerge. We test our proposed technique in a proof-of-concept experiment and show that our technique returns accurate disaggregation results.

VI. Acknowledgment

We would like to thank Mr. Hoang Hai Nguyen and Ms. Sreejaya Viswanathan for collecting the experiment data.

References

[1] U.S. Energy Information Administration, "Annual energy review," Oct. 2011, page 266.

[2] Electric Power Research Institute, "Residential electricity use feedback: A research synthesis and economic framework," Feb. 2009, page 3–4.

[3] G. Hart, "Nonintrusive appliance load monitoring," *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870 –1891, Dec. 1992.

[4] F. Sultanem, "Using appliance signatures for monitoring residential loads at meter panel level," *IEEE Transactions on Power Delivery*, vol. 6, no. 4, pp. 1380 –1385, Oct. 1991.

[5] H. Lee, A. Battle, R. Raina, and A. Y. Ng, "Efficient sparse coding algorithms," in *Proceedings of the Advances in Neural Information Processing Systems (NIPS) 19 Conference*, 2007, pp. 801–808.

[6] S. Gupta, M. S. Reynolds, and S. N. Patel, "Electrisense: single-point sensing using EMI for electrical event detection and classification in the home," in *Proceedings of the 12th ACM international conference on Ubiquitous computing*, ser. Ubicomp '10, 2010, pp. 139–148.

[7] S. Leeb, S. Shaw, and J. Kirtley, Jr., "Transient event detection in spectral envelope estimates for nonintrusive load monitoring," *IEEE Transactions on Power Delivery*, vol. 10, no. 3, pp. 1200 –1210, Jul. 1995.

[8] K. Suzuki, S. Inagaki, T. Suzuki, H. Nakamura, and K. Ito, "Nonintrusive appliance load monitoring based on integer programming," in *SICE Annual Conference, 2008*, aug. 2008, pp. 2742–2747.

[9] J. Z. Kolter, S. Batra, and A. Y. Ng, "Energy disaggregation via discriminative sparse coding," in *Proceedings of the Advances in Neural Information Processing Systems (NIPS) 22 Conference*, 2010.