Semi-Intrusive Load Monitoring for Large-Scale Appliances

Guoming Tang*, Kui Wu*, and Jingsheng Lei†
*Dept. of Computer Science, University of Victoria, B.C., Canada
†School of Computer and Information Engineering, Shanghai University of Electric Power, Shanghai, China

***Abstract*—Non-intrusive appliance load monitoring (NIALM) is to identify major energy guzzlers in a house or building without introducing extra metering cost. To develop an easy-to-use and scalable solution to energy disaggregation for contemporary large-scale appliance groups, we propose a Semi-Intrusive Appliance Load Monitoring (SIALM) approach in this paper. Based on a simple power model, a Sparse Switching Event Recovering (SSER) model is established to recover appliance states from their aggregated load data, and the necessary conditions for unambiguous state recovery of multiple appliances are provided. Under the constraints of necessary conditions, the minimum number of required smart meters is pursued via a greedy clique-covering algorithm. We evaluate the performance of SIALM with Monte Carlo simulation. The results show that our method achieves high accuracy not only in appliances' state recovery but also in energy disaggregation.**

I. Introduction

Non-intrusive appliance load monitoring (NIALM) is to find the energy consumption of individual appliances from only a single measure of household electricity consumption. Without introducing extra metering cost, accurate NIALM helps identify major energy guzzlers in the house. It motivates users to take proper actions for energy saving and greatly facilitates demand response (DR) programs.

Because of its significance, tremendous research efforts have been devoted to NIALM and a broad spectrum of approaches have been proposed since 1980s [6], [16]. Nevertheless, NIALM still remains an open challenge for *large-scale* appliances, since the number of possible appliances' states grows exponentially.

Existing solutions may suffer from the scalability problem and may pose non-realistic demand to end users. First, large-scale, diverse appliance groups consisting of tens or hundreds of appliances are common in modern houses and buildings. Many NIALM approaches were developed and validated upon small-scale appliance groups [1], [3], [4], [7], [8], [10], [11], [12], [13], and their accuracy with large-scale appliance groups may be unclear. Second, some NIALM methods require expert knowledge or use auxiliary measurement devices to extract appliance signatures, *i.e.*, special features that can be used to identify an appliance. This requirement poses a hurdle to end users who normally may not have enough knowledge to carry out such tasks.

To pursue an easy-to-use and scalable solution to energy disaggregation for contemporary large-scale appliance groups, we propose (1) relying only on readily available information of appliances and (2) using multiple low-cost meters that optimally form a monitoring network. On the former, we observe that the *rated power*[1] of an appliance is a readily available feature in practice. Such information could be easily found in the appliance's manual or some public websites [2]. Regarding the latter, low-cost energy meters can be plug-and-play and have become popular on current market. With such meters, we can easily monitor the power consumption of a single appliance or the aggregated power consumption of a small group of appliances.

Based on the above observations, we make the following contributions in the paper:

- For a large-scale appliance group, instead of using only one meter, we deploy multiple meters, each measuring a sub-group of appliances. We call this approach *semi-intrusive*.
- To infer the states of different appliances, instead of relying on sophisticated appliance signatures from massive load data, we make use of the appliances' rated power.
- For a group of appliances, we give the necessary conditions for unambiguous state monitoring. The conditions provide theoretical evidence for accurate energy disaggregation results.
- Under the constraints of necessary conditions, we develop an algorithm to achieve the minimum number of meters for unambiguous appliance state monitoring.

II. Problems Definition

A. Appliance Power Model

Most household appliances have one or multiple operating modes when they are turned on [13]. Furthermore, for a certain operating mode of an appliance, the value of its *rated power* can be easily found in the appliance's manual or some open database [2], and the *power deviation*[2] of certain operating mode can be easily derived from the power reading samples, without applying sophisticated machine learning approaches.

Without loss of generality, we consider a group of N appliances. For the n-th appliance with m operating modes, we construct a *power vector* to denote corresponding rated powers:

$$p^n := [p_1^n, p_2^n, \cdots, p_m^n]^T \,, \tag{1}$$

[1]The rated power here refers to the mean value of real power consumption of an appliance under a certain operating mode, with unit of *Watt*.

[2]The power deviation here refers to the maximum difference between the real power and rated power, with unit of *Watt*. Thus, the real power consumption of a running appliance with rated power p and power deviation θ is bounded by $[p - \theta, p + \theta]$.

TABLE I
TABLE OF NOTATIONS

Symbol	Explanation
p	rated power vector of all appliances
p^n	rated power vector of the n-th appliance
p^n_m	rated power of the n-th appliance at the m-th operating mode
θ	power deviation vector of all appliances
θ^n	power deviation vector of the n-th appliance
θ^n_m	power deviation vector of the n-th appliance at the m-th operating mode
λ^n	power model of the n-th appliance
Λ	power model of a group of appliances
Λ^n	power model of a sub-group of appliances
$s(t)$	state vector of all appliances at time t
s^n	state vector of the n-th appliance along the timeline
$s^n(t)$	state vector of the n-th appliance at time t
$s^n_m(t)$	on/off state of the m-th operating mode of the n-th appliance at time t
S	state matrix of all appliances along the timeline
S^n	recovered state matrix of the n-th appliance along the timeline
x	aggregated power vector of all appliances
$x(t)$	aggregated power reading of all appliances at time t
$x^n(t)$	power reading of the n-th appliance at time t

a *power deviation vector* to represent the power deviations at corresponding operating modes:

$$\theta^n := [\theta^n_1, \theta^n_2, \cdots, \theta^n_m]^T, \tag{2}$$

To denote the rated power of all the N appliances with multiple operating modes, we construct the following power vector:

$$p := \left[\left(p^1\right)^T \quad \left(p^2\right)^T \quad \cdots \quad \left(p^N\right)^T \right]^T. \tag{3}$$

and power deviation vector:

$$\theta := \left[\left(\theta^1\right)^T \quad \left(\theta^2\right)^T \quad \cdots \quad \left(\theta^N\right)^T \right]^T. \tag{4}$$

Definition 1. *The* ***power model*** *of the n-th appliance is defined as:*

$$\lambda^n := \{p^n, \theta^n\}, \tag{5}$$

and that of a group of N appliances is defined as:

$$\Lambda := \{p, \theta\}. \tag{6}$$

B. Semi-Intrusive Appliance Load Monitoring (SIALM)

Definition 2. ***Unambiguous Appliance State Monitoring:*** *Given the aggregated power readings of multiple appliances in a time interval, identify the exact state of each appliance at each time instant.*

With the above definition, the Semi-Intrusive Appliance Load Monitoring (SIALM) problem is to find the minimum number of meters to achieve unambiguous appliance state monitoring. Note that this problem is different from the real-time appliance state monitoring problem raised before in [5], [14], [15].

III. SPARSE SWITCHING EVENT RECOVERING (SSER)

Considering a group of N appliances, we use a *state vector* to denote states of the n-th appliance's corresponding operating modes at an arbitrary time instant, t, as:

$$s^n(t) := [s^n_1(t), s^n_2(t), \cdots, s^n_m(t)]^T, \tag{7}$$

where $s^n_m(t)$ represents the on/off state of the m-th operating mode of the n-th appliance at time instant t. Thus, $s^n_m(t) \in \{0,1\}$ and $\|s^n(t)\|_1 \leq 1$, with $s^n_m(t) = 1$ indicates the n-th appliance is on the m-th operating mode at time t, and 0 otherwise.

Then, with (7), we have the state vector of all the N appliances at time t, as:

$$s(t) := \left[\left(s^1(t)\right)^T \quad \left(s^2(t)\right)^T \quad \cdots \quad \left(s^N(t)\right)^T \right]^T, \tag{8}$$

which is an M-dimension vector, where $M = \sum_{n=1}^N m^n$ and m^n denotes the number of operating modes of the n-th appliance.

Lemma 1. *Given the power model of N appliances in (6), the aggregated power reading of the N appliances at an arbitrary time instant t, denoted as $x(t)$, should be bounded by:*

$$s^T(t)\left(p-\theta\right) \leq x(t) \leq s^T(t)\left(p+\theta\right). \tag{9}$$

With (8), we can construct a *state matrix* to represent states of all appliances from time $t=1$ to $t=K$:

$$S := [s(1) \quad s(2) \quad \cdots \quad s(K)]. \tag{10}$$

Since $s(t)$ is an M-dimension column vector, S is an M-by-K matrix, where $M = \sum_{n=1}^N m^n$ with m^n denoting the number of operating modes of the n-th appliance.

Lemma 2. *Given the power model of N appliances in (6), based on* ***Lemma 1****, their aggregated power readings from time $t=1$ to $t=K$, denoted as:*

$$x := [x(1), x(2), \cdots, x(K)]^T, \tag{11}$$

should be bounded by:

$$S^T\left(p-\theta\right) \leq x \leq S^T\left(p+\theta\right). \tag{12}$$

For the appliances in a house or building, their state switch events along the timeline are sparse, compared with the total load samples [14], [15]. Thus, given a group of N appliances with power model Λ, and aggregated power readings from time $t=1$ to $t=K$, x, we can establish an optimization model of sparse switching event recovering (SSER):

$$\begin{aligned} \min_{S} \quad & \mathbf{TV}(SD) \\ s.t. \quad & x - S^T(p+\theta) \leq \mathbf{0}, \\ & S^T(p-\theta) - x \leq \mathbf{0}, \\ & HS \leq \mathbf{1}, \end{aligned} \tag{13}$$

where $\mathbf{TV}(A)$ denotes the *total variation* of matrix A, calculated by

$$\mathbf{TV}(A) := \sum_i \sum_j |a_{i,j}|, \tag{14}$$

D is a K-by-$(K-1)$ *difference matrix* defined by:

$$D := \underbrace{\begin{bmatrix} -1 & & & & \\ 1 & -1 & & & \\ & 1 & \ddots & & \\ & & \ddots & -1 & \\ & & & 1 & -1 \\ & & & & 1 \end{bmatrix}}_{K-1} \tag{15}$$

and H is an N-by-M *permutation matrix* defined by:

$$H := \underbrace{\begin{bmatrix} \underbrace{1 \ \cdots \ 1}_{m^1} & & & \\ & \underbrace{1 \ \cdots \ 1}_{m^2} & & \\ & & \ddots & \\ & & & \underbrace{1 \ \cdots \ 1}_{m^N} \end{bmatrix}}_{M} \tag{16}$$

where $M = \sum_{n=1}^{N} m^n$ with m^n denoting the number of operating modes of the n-th appliance. In (13), **0** is a K dimensional all 0 vector, and **1** is an N-by-K all 1 matrix.

IV. Necessary Conditions of Unambiguous State Recovery

With the SSER optimization model, we provide the necessary conditions to achieve unambiguous appliance state monitoring. We omit the proof to save space.

Theorem 1. *Given a group of N appliances with power model Λ, and aggregated power readings in a time interval, the* ***necessary conditions*** *to achieve unambiguous appliance state recovery under SSER model are:*

C-1: $\forall i, j \in \{1, 2, \cdots, N\}, i \neq j$

$$\left[p^i - \theta^i, p^i + \theta^i\right] \nsubseteq \left[p^j - \theta^j, p^j + \theta^j\right], \tag{17}$$

C-2: $2\left\|\theta\right\|_1 < min\{p^1 - \theta^1, p^2 - \theta^2, \cdots, p^N - \theta^N\}$. (18)

V. Optimal Smart Meter Deployment

We partition all appliances into exclusive sub-groups such that within each sub-group, the power model of appliances fits for the necessary conditions and can achieve unambiguous state monitoring. In order to lower the cost as much as possible, it is desirable to minimize the total number of sub-groups, *i.e.*, to minimize the number of smart meters. This optimization problem can be formulated as follow:

$$\begin{aligned} \min_{\{\Lambda^1, \Lambda^2, \cdots, \Lambda^n\}} \quad & n \\ \textit{s.t.} \quad & \bigcup \Lambda^i = \Lambda, \\ & \Lambda^i \bigcap \Lambda^j = \varnothing, i \neq j, \\ & \Lambda^i = \{\forall \{p, \theta\} \text{ fits } \textbf{\textit{C-1}} \& \textbf{\textit{C-2}}\}. \end{aligned} \tag{19}$$

In order to solve the above problem, a graph $G = (V, E)$ is constructed, where each vertex $v \in V$ represents the power model of an appliance and an edge is built between two vertices if the power models fit constraints ***C-1*** and ***C-2***. It is easy to see that the problem is equivalent to the *clique-covering* problem, which is NP-complete. Hence, a *greedy clique-covering algorithm* is adopted to obtain an approximate solution. The basic idea of the algorithm is to find cliques that cover more vertices that have not been clustered so far and run the process iteratively until all vertices are clustered. It is easy to prove that in *the worst case* the running complexity of this algorithm is $O(N^3)$ where N is the total number of appliances.

VI. Numerical Evaluations

A. Data Generation via Monte Carlo Simulation

TABLE II
Parameter Setting in Monte Carlo Simulation

Parameter	*Setting*
Number of Appliances (N)	$50, 100, 200$
Number of Operating Modes (M)	3
Total Samples/Time Interval (K)	8640
Lowest Appliance Power (p_{min})	$20w$
Highest Appliance Power (p_{max})	$2000w$
Power Range Ratio (r)	$[0.01, 0.1]$
Poisson Parameter (τ)	180

We apply the *Monte Carlo* simulation to generate the load data for large-scale appliances. The parameters used to generate the appliance power model and synthetic data are listed in Table II.

- Given the number of appliances, N, the operating mode number of each appliance is uniformly assigned between 1 and M.
- Given the lowest power (p_{min}) and the highest power (p_{max}), the lower bound of one operating mode of an appliance (p_l) is a random variable uniformly distributed between p_{min} and p_{max}. The upper bound of the operating mode (p_u) is determined by a parameter called *power range ratio* (r) and is calculated by $p_u = min\{p_l + random([0, r \cdot p_l]), p_{max}\}$, where $random([0, r \cdot p_l])$ returns a random number uniformly distributed in the range $[0, r \cdot p_l]$.
- Validate the appliance's power model with unambiguous state monitoring necessary constraints ***C-1*** and ***C-2***, and referring to the necessary conditions, partition the N appliances into multiple sub-groups using the greedy clique-covering algorithm.
- To generate a sample, each appliance reports its operating mode, which is a random number uniformly selected in its mode range, and its power reading, which is a random number uniformly distributed between the appliance's power bounds. It reports 0 if its state is *off*. At the end, the aggregated power reading of appliances (*i.e.*, the sum of appliances' power readings in the sub-group) is recorded.
- The occurrence of state switching event of an appliance follows a Poisson distribution with parameter τ.

TABLE III
PERFORMANCE RESULTS: NON-INTRUSIVE APPLIANCE LOAD MONITORING (NIALM) VS. SEMI-INTRUSIVE APPLIANCE LOAD MONITORING (SIALM)
(*In EDA and SRA, ∗ is labeled for the optimal value from the global optimization, and others are approximate ones from sequential local optimizations.*)

# of Appliances	**Non-Intrusive Appliance Load Monitoring (NIALM)**			**Semi-Intrusive Appliance Load Monitoring (SIALM)**		
	# of Smart Meters	*EDA*	*SRA*	*# of Smart Meters*	*EDA*	*SRA*
30	1	54.07%	51.40%	6	∗98.60%	∗100.00%
50	1	46.07%	41.03%	8	88.71%	89.63%
80	1	44.05%	42.81%	16	89.33%	90.65%
100	1	42.09%	42.71%	18	87.48%	85.45%
200	1	39.54%	37.92%	35	85.66%	84.30%

B. Performance Metrics and Evaluations

- *Energy Disaggregation Accuracy (EDA):* which indicates the accuracy of assigning correct power values to corresponding appliances [9].

$$EDA := 1 - \frac{\sum_{n=1}^{N} \|p_r^n - S^n p^n\|_1}{2\|x\|_1}, \qquad (20)$$

  in which N is the number of appliances, p_r^n, S^n and p_n represent the real power consumption vector, the recovered state matrix, and the rated power vector of the n-th appliance, respectively, and x is the aggregated power vector.
- *State Recovery Accuracy (SRA):* which indicates the accuracy of recovering the states of appliances.

$$SRA := 1 - \frac{\sum_{n=1}^{N} \|S_r^n - S^n\|_1}{N \cdot K}, \qquad (21)$$

  in which N is the number of appliances, S_r^n, S^n represent the real state matrix and the recovered state matrix of the n-th appliance, respectively, and N, K represent the number of appliances and the number of samples, respectively.

The energy disaggregation accuracy, state recovery accuracy, and the minimum number of required meters are summarized in Table III. We have proved that solving the SSER model is NP-hard. So except in the first case where the number of appliances is small, we used a sequential local optimization algorithm to find approximation solutions in other cases. This is the reason why the values of *SRA* except the first one did not reach 100%.

The results show that with the help of a few more smart meters, the accuracy of energy disaggregation for large-scale appliances can be significantly improved. Furthermore, from the first case where the number of appliances is 50, we have shown that with unambiguous necessary conditions and the SSER model, the accuracy of recovered appliance states can reach as high as 100%.

VII. CONCLUSIONS

We proposed a Semi-Intrusive Appliance Load Monitoring (SIALM) approach to energy disaggregation for large-scale appliances. Instead of using only one meter, multiple meters were deployed to collect the load data from sub-groups of appliances. Based on a simple power model, we established a Sparse Switching Event Recovering (SSER) model to recover appliance states from the aggregated load data, and provided the necessary conditions for unambiguous state recovery of multiple appliances. Furthermore, we found the minimum number of meters via a greedy clique-covering algorithm. Our simulation results disclosed that for large-scale appliances, SIALM performs much better than NIALM.

REFERENCES

[1] M. Dong, P. C. Meira, W. Xu, and W. Freitas. An event window based load monitoring technique for smart meters. *Smart Grid, IEEE Transactions on*, 3(2):787–796, 2012.

[2] EPA. A tool of home product finder by energy star. http://www.energystar.gov/productfinder/, 2013. [Online; accessed in 18-December-2013].

[3] M. Figueiredo, B. Ribeiro, and A. M. de Almeida. On the regularization parameter selection for sparse code learning in electrical source separation. In *Adaptive and Natural Computing Algorithms*, pages 277–286. Springer, 2013.

[4] 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*, pages 139–148. ACM, 2010.

[5] X. Hao, Y. Wang, C. Wu, A. Y. Wang, L. Song, C. Hu, and L. Yu. Smart meter deployment optimization for efficient electrical appliance state monitoring. In *Smart Grid Communications (SmartGridComm), 2012 IEEE Third International Conference on*, pages 25–30. IEEE, 2012.

[6] G. W. Hart. Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, 80(12):1870–1891, 1992.

[7] H. Kim, M. Marwah, M. F. Arlitt, G. Lyon, and J. Han. Unsupervised disaggregation of low frequency power measurements. In *SDM*, pages 747–758, 2011.

[8] J. Z. Kolter, S. Batra, and A. Ng. Energy disaggregation via discriminative sparse coding. In *Advances in Neural Information Processing Systems*, pages 1153–1161, 2010.

[9] J. Z. Kolter and M. J. Johnson. Redd: A public data set for energy disaggregation research. In *proceedings of the SustKDD workshop on Data Mining Applications in Sustainability*, pages 1–6, 2011.

[10] H. Lam, G. Fung, and W. Lee. A novel method to construct taxonomy electrical appliances based on load signatures. *Consumer Electronics, IEEE Transactions on*, 53(2):653–660, 2007.

[11] O. Parson, S. Ghosh, M. Weal, and A. Rogers. Non-intrusive load monitoring using prior models of general appliance types. In *AAAI*, 2012.

[12] S. R. Shaw, S. B. Leeb, L. K. Norford, and R. W. Cox. Nonintrusive load monitoring and diagnostics in power systems. *Instrumentation and Measurement, IEEE Transactions on*, 57(7):1445–1454, 2008.

[13] K. Suzuki, S. Inagaki, T. Suzuki, H. Nakamura, and K. Ito. Nonintrusive appliance load monitoring based on integer programming. In *SICE Annual Conference, 2008*, pages 2742–2747. IEEE, 2008.

[14] Y. Wang, X. Hao, L. Song, C. Wu, Y. Wang, C. Hu, and L. Yu. Tracking states of massive electrical appliances by lightweight metering and sequence decoding. In *Proceedings of the Sixth International Workshop on Knowledge Discovery from Sensor Data*, pages 34–42. ACM, 2012.

[15] Y. Wang, X. Hao, L. Song, C. Wu, Y. Wang, C. Hu, and L. Yu. Monitoring massive appliances by a minimal number of smart meters. *ACM Transactions on Embedded Computing Systems (TECS)*, 13(2s):56, 2014.

[16] M. Zeifman and K. Roth. Nonintrusive appliance load monitoring: Review and outlook. *Consumer Electronics, IEEE Transactions on*, 57(1):76–84, 2011.