A Distributed and Scalable Approach to Semi-Intrusive Load Monitoring

Guoming Tang, *Student Member, IEEE*, Kui Wu, *Senior Member, IEEE*, and Jingsheng Lei

Abstract—Non-intrusive appliance load monitoring (NIALM) helps identify major energy guzzlers in a building without introducing extra metering cost. It motivates users to take proper actions for energy saving and greatly facilitates demand response (DR) programs. Nevertheless, NIALM of large-scale appliances is still an open challenge. To pursue a scalable solution to energy monitoring for contemporary large-scale appliance groups, we propose a distributed metering platform and use parallel optimization for semi-intrusive appliance load monitoring (SIALM). Based on a simple power model, a sparse switching event recovering (SSER) model is established to recover appliance states from their aggregated load data. Furthermore, the sufficient conditions for unambiguous state recovery of multiple appliances are presented. By considering these conditions as well as the electrical network topology constraint, a minimum number of meters are obtained to correctly recover the energy consumption of individual appliances. We evaluate the performance of both SIALM and NIALM with real-world trace data and synthetic data. The results demonstrate that with the help of a small number of meters, the SIALM approach significantly improves the accuracy of energy disaggregation for large-scale appliances.

Index Terms—Semi-intrusive appliance load monitoring, energy disaggregation, optimization, measurement

1 Introduction

According to a recent report, the residential and commercial buildings consumed about 74 percent of US electricity consumption in 2013 [1]. To save energy, it is critical to not only monitor the energy consumption of buildings, but also better understand where and how the energy is consumed. For this reason, energy monitoring for appliances in buildings has attracted more and more attention and cheap solutions for non-intrusive appliance load monitoring (NIALM) are highly demanded. NIALM does not use a dedicated power meter for each appliance. Instead, it finds the energy consumption of individual appliances from only a single point of measurement of a building's electricity consumption, and as such NIALM is also known as energy disaggregation. Without introducing extra metering cost, accurate NIALM helps identify major energy guzzlers in a building, as illustrated in Fig. 1. It motivates the facility management to take proper actions for energy saving and greatly facilitates demand response (DR) programs [2], [3].

With a single point of measurement, NIALM on one hand simplifies the task of load monitoring, but on the other hand its accuracy suffers as the number of appliances increases. While large-scale, diverse appliance groups consisting of hundreds or thousands of appliances are common in commercial buildings, many NIALM approaches were developed and validated upon small-scale appliance groups, and their accuracy with large-scale appliance groups may be unclear. Recently, low-cost energy meters can be plug-and-play and have become popular on market, e.g., [5]. With such low-cost meters, we can easily monitor the power consumption of a single appliance or the aggregated power consumption of a small group of appliances *without changing existing electric circuitry* in the building. As such, we see no convincing need to stick with a single point of measurement.

In this paper, we explore the benefit of introducing low-cost energy meters for monitoring large-scale appliance groups. This effort is not targeted at providing a complete, universal solution to NIALM; instead it is to demonstrate, in theory and in practice, that the accuracy of NIALM can be improved significantly with a small number of extra meters. While this is intuitively true, an in-depth analysis is needed to better understand the tradeoff between the benefit and the metering overhead. In addition, design problems such as where and how many meters should be installed must be addressed.

For such an investigation, we adopt a simple power model relying only on readily-available or easily-available information of appliances, such as the *rated power*[1] of an appliance, which is often found in the appliance's manual and some public websites [6], or can be easily measured via the plug-and-play power meter [5]. Compared to other complex NIALM methods that require expert knowledge or use auxiliary measurement devices to extract appliance signatures [7], this model promotes the simplicity of NIALM.

Targeting at developing a solution to energy disaggregation for large-scale appliance groups, we make the following contributions in the paper.

- *G. Tang and K. Wu are with the Department of Computer Science, University of Victoria, Victoria, B.C, Canada. E-mail: {guoming, wkui}@uvic.ca.*
- *J. Lei is with School of Computer and Information Engineering, Shanghai University of Electric Power, Shanghai, China. E-mail: jshlei@shiep.edu.cn.*

Manuscript received 25 Apr. 2015; revised 27 June 2015; accepted 9 Aug. 2015. Date of publication 18 Aug. 2015; date of current version 18 May 2016. Recommended for acceptance by J. Chen.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TPDS.2015.2470238

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*.

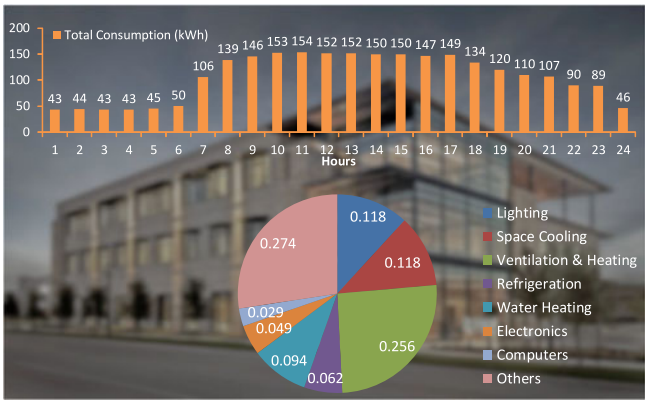

Fig. 1. An example of energy disaggregation for one-day energy consumption of a commercial building. The disaggregated results are referred to the (predicted) 2015 building energy end-use splits in US [4].

- For a large-scale appliance group, instead of using only one meter, we deploy multiple meters distributed over the power network, each measuring a sub-group of appliances. We call such an approach *semi-intrusive appliance load monitoring (SIALM)* in this paper.
- 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 sufficient conditions for unambiguous state recovery. The conditions provide theoretical evidence for accurate energy disaggregation results. Accordingly, a parallel optimization algorithm is developed for accurate appliance state recovery. In addition, we quantitatively investigate the impact on energy disaggregation results when the sufficient conditions are violated.
- We perform comprehensive performance test and robustness test with real-world data and synthetic data. The experimental results show that with the help of a small number of extra meters, our SIALM approach can significantly improve the accuracy of energy disaggregation for large-scale appliances. Such improvement is resilient to inaccurate power estimation and network topology changes.

The rest of this paper is organized as follows. In Section 2, we first review current approaches applied for non-intrusive load monitoring and real-time appliance state monitoring (RTASM), respectively. The problem of semi-intrusive appliance load monitoring is defined based on our appliance power model in Section 3. Then, according to the temporary sparsity of appliances' switching events, we establish a sparse switching event recovery (SSER) model in Section 4, and provide the sufficient conditions to achieve unambiguous state recovery in Section 5. Subsequently, the meter deployment optimization strategies are developed in Section 6 to partition a number of appliances into sub-groups and ensure the unambiguous sufficient conditions. Comprehensive experiments are performed using real-world data and synthetic data in Section 7 to evaluate both the accuracy and robustness of our SIALM approach. We further clarify the misconception regarding NIALM in Section 8 and conclude the paper in Section 9.

2 Related Work

2.1 Non-Intrusive Appliance Load Monitoring (NIALM)

Tremendous research efforts have been devoted to NIALM and a broad spectrum of approaches have been proposed since 1980s [8], [9], [10]. Recently, it has also drawn attention from both large companies and small start-ups, including for example Intel, Belkin, GetEmme, and Navetas. According to [9], however, there is no best solution working well for all types of appliances. So far, no single solution stands out to solve all the problems in NIALM.

Signature based approaches are most popular among those solving the energy disaggregation problem. Either steady or transient signal features of running appliances are extracted from the labeled training dataset and treated as signatures to identify the appliances from aggregated load data [11]. Then, in the testing dataset, some event detecting schemes are developed to estimate whether or not appliances are on/off, and the detected events are ascribed to the appliances' activities via classification strategies [12], [13], [14]. In addition to the time-domain signal features, spectral analysis has also been adopted to discover appliances' signatures in the frequency domain [7], [15], [16]. Nevertheless, the signatures used in those methods may be hard to obtain in practice without using sophisticated machine learning methods or auxiliary measuring devices [7]. Furthermore, the performance greatly depends on the uniqueness of appliances' signatures. So far, no complete set of robust, widely accepted appliance signatures has been identified [9].

Some NIALM methods make use of state transition in appliances' activities. Hidden Markov model (HMM) is applied to solve energy disaggregation, in which the state transition patterns of appliances are trained and the hidden states of each appliance are predicted by inference strategies (such as Viterbi algorithm) with the observed emission probabilities [17], [18]. Some variants are also developed based on HMM, such as Factorial HMM (FHMM) [19], Additive Factorial Approximate MAP (AFMAP) [20], and Conditional Factorial Hidden Semi-Markov Model (CFHSMM) [21]. Non-negative sparse coding is proposed to solve the energy disaggregation problem in [22] and further explored in [23], in which a training process is performed to obtain the basis vector related to the state transition patterns among appliances. Zoha et al. [10] give a comprehensive survey for the state transition based methods. Those approaches usually need a large number of trainings and are thus time consuming. In addition, since the solutions rely on the patterns of appliances' activities in certain training datasets, their performance may vary with different datasets, especially for large-scale appliances.

2.2 Real-Time Appliance State Monitoring (RTASM)

In addition to the approaches to NIALM, there are some related solutions for real-time appliance state monitoring/recovering from their aggregated power readings with multiple meters. In [24], the power distribution network is decomposed to multiple mono-meter trees, and the real-time states of appliances are predicted according to their aggregated power readings by a minimum number of meters without errors. In [25], a lightweight metering and

TABLE 1
Table of Notations

Symbol	*Explanation*
p	power vector of all appliances
p^n	power vector of the nth appliance
p_m^n	power of the nth appliance at the mth operating mode
θ	power deviation vector of all appliances
θ^n	power deviation vector of the nth appliance
θ_m^n	power deviation vector of the nth appliance at the mth operating mode
λ^n	power model of the nth appliance
Λ	power model of a group of appliances
Λ^n	power model of the nth sub-group of appliances
$s(t)$	state vector of all appliances at time t
s^n	state vector of the nth appliance along the timeline
$s^n(t)$	state vector of the nth appliance at time t
$s_m^n(t)$	on/off state of the mth operating mode of the nth appliance at time t
S	state matrix of all appliances along the timeline
S^n	recovered state matrix of the nth sub-group of appliances
x	aggregated power vector of all appliances
$x(t)$	aggregated power reading of all appliances at time t
x^n	aggregated power vector of appliances in the nth sub-group

sequence decoding approach is developed to recover the real-time states of massive appliances with multiple meters, in which an HMM-based state sequence decoding model is adopted to infer the hidden states of appliances. This method is further validated in [26] with extensive experiments on simulated data and real PowerNet data. Since RTASM focuses on states prediction in real time, it only makes use of current or part of the historical aggregated data. Consequently, with limited information, the required number of meters may be too large.

Aiming at different targets, the principles in RTASM are different from those in NIALM. Timeliness is the most important requirement for RTASM methods. Therefore, only current consumption data or a small part of the historical consumption data are used for the short-time prediction. Nevertheless, accuracy instead of timeliness is the key for NIALM methods. In order to improve the accuracy, NIALM methods usually require more detailed energy consumption models and unambiguous features of appliances, and take advantage of aggregated load data during the whole time interval.

3 Problem Representation and Definition

For ease of reference, we list the notations in Table 1.

3.1 Appliance Power Model

According to [8], [9], [13], most appliances have one or multiple operating modes. Furthermore, for a certain operating mode of an appliance, the value of its *rated power* can be easily accessed, by either referring to the user's manual of the appliance or searching the Internet. For example, Fig. 2 shows the rated power information of a microwave with multiple operating modes, which is specified in the user's


NN-CT579SBPQ
Slimline Combination Microwave

Electricity

Rated Voltage	230-240V
Power Frequency	50Hz
Power Consumption	
Microwave	1140W
Grill	1360W
Convection	1375W
max.	2385W

Fig. 2. Power consumption information of a microwave specified in the user's manual.

manual. Furthermore, the *power deviation*[2] of certain operating mode can be easily measured from the power readings without applying sophisticated machine learning approaches, e.g., using the cheap plug-and-play meters from [5].

Without loss of generality, we consider a group of N appliances. For the nth appliance with m operating modes, we denote the corresponding rated powers with a *power vector* as:

$$p^n := \left[p_1^n, p_2^n, \ldots, p_m^n\right]. \tag{1}$$

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

$$\theta^n := \left[\theta_1^n, \theta_2^n, \ldots, \theta_m^n\right], \tag{2}$$

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

$$p := \left[p^1, p^2, \ldots, p^N\right]^T. \tag{3}$$

and power deviation vector:

$$\theta := \left[\theta^1, \theta^2, \ldots \theta^N\right]^T. \tag{4}$$

Definition 1. *The* ***power model*** *of the nth appliance can be defined as:*

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

and a group of N appliances can be denoted as:

$$\Lambda := \{\lambda^1, \lambda^2, \ldots, \lambda^N\}. \tag{6}$$

Furthermore, for the nth appliance, we use a *state vector* to denote the states of its corresponding operating modes at an arbitrary time instant, t, as:

$$s^n(t) := \left[s_1^n(t), s_2^n(t), \ldots, s_m^n(t)\right], \tag{7}$$

where $s_m^n(t)$ represents the on/off state of the mth operating mode of the nth appliance at time instant t. Thus, $s_m^n(t) \in \{0, 1\}$ and $\|s^n(t)\|_1 \leq 1$, where $\|.\|_1$ means l_1 norm

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 $[p - \theta, p + \theta]$.

(i.e., $\sum_{j=1}^{m} |s_j^n(t)|$) and $s_m^n(t) = 1$ indicating that the nth appliance is on the mth operating mode at time t and 0 otherwise.

3.2 Semi-Intrusive Appliance Load Monitoring

Definition 2 (Unambiguous State Recovery). *Given the aggregated power readings of multiple appliances in a time interval, provide the exact state of each appliance at each time instant.*

Definition 3 (Predictable Energy Disaggregation). *Given the aggregated power readings of multiple appliances in a time interval, estimate the energy consumption of each appliance during the time interval with a bounded error.*

With the above definitions, the semi-intrusive appliance load monitoring problem can be defined as follow.

Definition 4 (Semi-Intrusive Appliance Load Monitoring (SIALM)). *Given a group of appliances and their power models, provide the minimum number of meters to ensure both unambiguous state recovery and predictable energy disaggregation.*

4 Parallel Sparse Switching Event Recovering

In this section, we first design a sparse switching event recovering (SSER) model based on total variation minimization. Then, a parallel local optimization algorithm is proposed to solve the (NP-hard) SSER problem.

4.1 SSER Model

Considering a group of N appliances with states in (7), at an arbitrary time instant, t, the *state vector* of all appliances can be denoted as:

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

which is an M-dimension column vector, where $M = \sum_{n=1}^{N} \|p^n\|_0$ and $\|p^n\|_0$ denoting the number of operating modes of the nth appliance since $\|.\|_0$ counts the nonzero elements in a vector and all values in the vector p^n are larger than 0.

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:*

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

With (8), we can construct a *state matrix* of all appliances in a time interval, without loss of generality, from time $t = 1$ to $t = K$, as:

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

Since $s(t)$ is an M-dimension column vector, S is an M-by-K matrix.

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:

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

should be bounded by:

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

It has been known that, for most appliances, a general feature of their state (operating mode) switching events is *sparsity* [8], [25], [26], i.e.,

- During a short time interval, the number of state switching events for all appliances is quite small.
- For the whole time interval, the total number of state switching events is significantly smaller than the number of samples.

For a further clarification of sparsity, refer to Section 8.

Based on the above sparsity feature of appliances' state switching, we establish an optimization model of sparse switching event recovering:

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

In the above model, i) $\mathbf{TV}(S)$ denotes the *total variation* of the state matrix S, calculated by

$$\mathbf{TV}(S) := \|SD\|_{\mathcal{F}} = \sum_i \sum_j \left|(SD)_{i,j}\right|, \tag{14}$$

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

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

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

$$H := \underbrace{\begin{bmatrix} \underbrace{1 \cdots 1}_{\|p^1\|_0} & & & \\ & \underbrace{1 \cdots 1}_{\|p^2\|_0} & & \\ & & \ddots & \\ & & & \underbrace{1 \cdots 1}_{\|p^N\|_0} \end{bmatrix}}_{M}. \tag{16}$$

and iii) $\mathbf{0}$ is a K dimensional all 0 vector, and $\mathbf{1}$ is an N-by-K all 1 matrix.

Remark 1. In the above SSER model, the objective function of total variation (TV) minimizing is applied as a classical approach to recovering a sparse matrix. It has been widely used in signal restoration, image denoising, and compressive sensing [27]. In our case, by using TV minimization, we actually minimize the total number of state changes which is regarded sparse along the timeline. For the constraints, the first two are the power reading constraints derived in equation (12); the third one represents the fact that an appliance can be only on one operating mode at any instant of time.

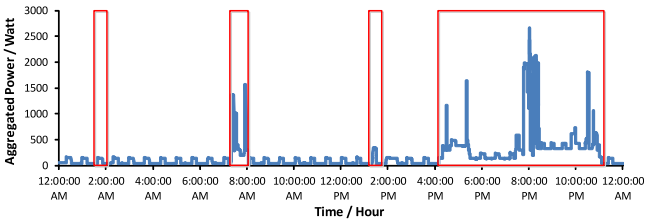

Fig. 3. A real example of one-day power consumption from a residential house and four types of split time windows.

4.2 Parallel Optimization for SSER Model

To solve the SSER model, however, is proved to be NP-hard (refer to the supplemental document, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TPDS.2015.2470238, of this paper). Thus, an approximate approach has to be proposed to solve the problem, especially when the dataset size is large.

We have observed that the aggregated power of a house increases and varies when the occupants are at home and use appliances, while it decreases to the baseline when the house is unoccupied or no appliances are used. With this observation, we can split the whole timeline into smaller time windows based on whether the house is occupied or the power consumption is close to the baseline.

As a real case shown in Fig. 3, the one-day aggregated power curve can be split into four kinds of smaller windows: i) periodical windows along the timeline for the activity of always-on appliances (e.g., the fridge in this example), ii) morning activities window, iii) noon/afternoon activities window, and iv) evening activities window. Note that the occupants are not always at home and even when they are at home, the appliance usage is always in burst. Thus, the length of split time windows can be quite short.

For a certain time window starting from $t = k$ and with length ℓ, we can solve a sub-problem of SSER model with the following form:

$$\begin{aligned} \min_{S_{k:k+\ell}} \quad & \mathbf{TV}(S_{k:k+\ell}) \\ s.t. \quad & x_{k:k+\ell} - S^T_{k:k+\ell}(p+\theta) \leq \mathbf{0}, \\ & S^T_{k:k+\ell}(p-\theta) - x_{k:k+\ell} \leq \mathbf{0}, \\ & H_{k:k+\ell} S_{k:k+\ell} \leq \mathbf{1}, \end{aligned} \tag{17}$$

in which the size of original matrices (i.e., S and H) and vector (i.e., x) appearing in (13) is cut down to ℓ. As ℓ is expected to be way shorter than the whole time length, the sub-problem can be solved efficiently using tools such as CVX with a Gurobi engine [28].

Assuming that we have split the whole time period into Q windows, for the qth ($1 \leq q \leq Q$) window with length w_q. Thus, for each split time window we can perform a local optimization by solving a sub-problem like (17). Algorithm 1 shows the pseudo code of the parallel optimization procedure, in which *parfor*[3] is a notation of parallel computing instead of *for*.

3. With support of *Parallel Computing Toolbox*, *parfor* loop can be applied in *Matlab* and bring in multiple workers to execute codes in parallel.

Algorithm 1. Parallel Optimization to Solve SSER Model

Input: Aggregated power vector $x_{1:K}$, appliances power vector $p_{1:M}$ and deviation vector $\theta_{1:M}$, split time windows $w_{1:Q}$.
Output: State matrix S.
1: **parfor** $q \leftarrow 1 : Q$ **do**
2: $t \leftarrow (w_q).start$
3: $\ell \leftarrow (w_q).end - (w_q).start$
4: Get $S_{t:t+\ell}$ by solving (17).
5: **end parfor**
6: **return** $S_{1:K}$.

By applying the parallel optimization for SSER model, we can recover the state matrix S, i.e., the on/off states of different operating modes of each appliance along the timeline. Then, the energy consumption of individual appliances can be estimated with the rated power information. Meanwhile, based on the power deviations of different operating modes, the upper and lower bounds of the energy consumption of individual appliances can be provided. Eventually, we achieve the initial objective of energy disaggregation.

5 Unambiguous State Recovery and Predictable Energy Disaggregation

With the SSER optimization model, we first provide the sufficient conditions to achieve unambiguous appliance state recovery. When the sufficient conditions are violated, we define the ambiguous degree to quantify the violations and investigate the impact on the accuracy of energy disaggregation.

5.1 Sufficient Conditions of Unambiguous State Recovery

Theorem 1. *Given a group of appliances and their power models, the sufficient conditions for unambiguous state recovery with the SSER model are:*

$$\begin{cases} \textbf{\textit{C}-0} : \|s(t+1) - s(t)\|_1 \leq 1, \forall t \\ \textbf{\textit{C}-1} : [(p-\theta)_i, (p+\theta)_i] \nsubseteq [(p-\theta)_j, (p+\theta)_j], \forall i \neq j \\ \textbf{\textit{C}-2} : 2 \cdot \|\theta\|_1 < (p-\theta)_i, \forall i \end{cases}$$

where $(p \pm \theta)_i$ denotes the ith element of vector $p \pm \theta$.

Proof. Assume that $\mathcal{F}$ is the set of feasible solutions satisfying the constraints in SSER model, the optimal solution from SSER model is $S_{opt} \in \mathcal{F}$, which makes SD sparsest, and the ground-truth state matrix is S_{true} and $S_{true} \in \mathcal{F}$. Then, $S_{opt} \neq S_{true}$, if and only if: (A) S_{opt} is not a unique solution of (13), or (B) S_{opt} is sparser than S_{true}, i.e.,

$$S_{opt} \neq S_{true} \Longleftrightarrow A \vee B. \tag{18}$$

With constraint of ***C*-0**, no more than one appliance could change its states during a sampling interval. Thus, conditions A and B can only be caused by two kinds of power range settings, respectively:

- A is caused by the ambiguity of power range settings illustrated in Fig. 4, i.e., the power range of one device is covered by that of another device.
- B is caused by an over-sparse optimal solution, i.e., the switch events of a device cannot be

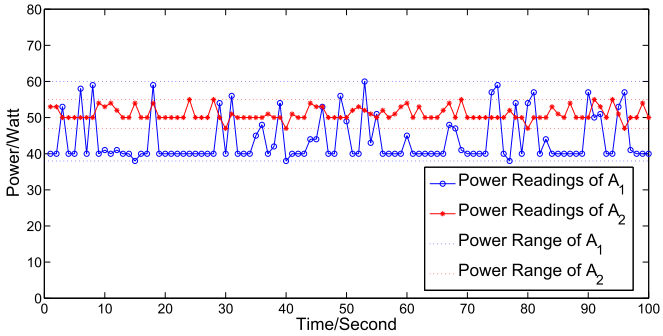

Fig. 4. The power range of A_2 is covered by that of A_1. Therefore, when A_2 is actually on, there will be two optimal solutions, as the switching events of A_1 and the switching events of A_2 are indistinguishable.

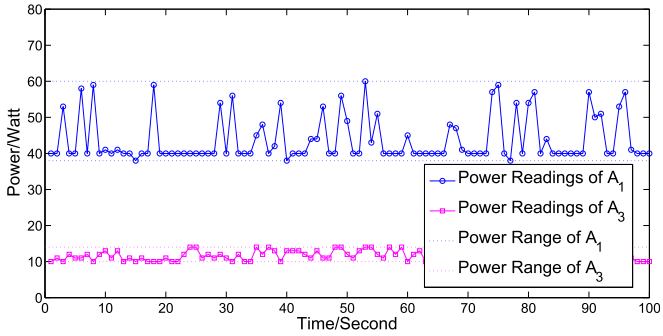

Fig. 5. The power range of A_3 is submerged by the power deviation of A_1. Therefore, when A_1 is on, the state switching events of A_3 only result in small power changes, and due to the objective function of (13), such small changes will be considered as part of power deviation of A_1. In other words, the switching events of A_3 will not be identified, resulting in an optimal solution sparser than the ground-truth.

identified since they are submerged as part of power deviation of another device, as illustrated in Fig. 5.

The constraint ***C-1*** is to eliminate condition A and the constraint ***C-2*** is to avoid condition B.

Therefore, we have

$$C\text{-}0 \wedge C\text{-}1 \wedge C\text{-}2 \Longrightarrow \neg A \wedge \neg B \Longleftrightarrow S_{opt} = S_{true}, \tag{19}$$

which means SSER achieves unambiguous state recovery. □

The physical meanings of the three conditions in Theorem 1 are as follow:

- ***C-0***: no more than one appliance could change its states during the same sampling interval, which is also known as one-at-a-time constraint and has been used before [20].
- ***C-1***: the power range of an appliance should not be completely covered by that of another (to eliminate the situation appearing in Fig. 4).
- ***C-2***: the power deviation of any appliance should be smaller than the power range of any other appliance (to eliminate the situation appearing in Fig. 5).

Remark 2. The constraints stated in the theorem are sufficient conditions for achieving unambiguous state recovery. The theorem does not necessarily mean in practice all the conditions will be met, *e.g.*, the estimated power deviation may not be accurate. The violation of these conditions is the main reason why we cannot achieve 100 percent accuracy in practice. Nevertheless, the theorem provides us with good heuristics to search for high-accuracy solutions.

After recovering the state of appliances at each time instant, we estimate the real power of each appliance at the time instant as the rated power at the recovered state. With unambiguous state recovery under the SSER model, we can easily provide the error bounds for energy disaggregation as follows.

Corollary 1. *Energy Disaggregation Error Bounds for Individual Appliance: with unambiguous appliance state recovery under the SSER model, for the nth appliance with power model $\lambda^n = \{p^n, \theta^n\}$, the energy disaggregation error ϵ satisfies 1) at each time instant: $\epsilon \leq \theta^n$, and 2) from time $t = 1$ to $t = K$: $\epsilon \leq K \cdot \theta^n$.*

5.2 Violation of the Sufficient Conditions

In the sufficient conditions, **C-0** can be enforced by making the sampling interval sufficiently small. Since current power meters [5] can sample data in seconds, this condition generally holds in practice. As such, we mainly focus on the violation of Conditions ***C-1*** and ***C-2*** in the following.

Conditions ***C-1*** and ***C-2*** may be violated when 1) the power range of an appliance is submerged by another's power range or 2) the power range of an appliance is submerged by the power deviation of another appliance. To quantify the degree of such violations, we introduce the concept of ambiguous degree.

Definition 5. *Given the power models of all appliances $\Lambda = \{\lambda^1, \lambda^2, \ldots, \lambda^N\}$, the **ambiguous degree** is defined as:*

$$d(\Lambda) := \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} I(\lambda^i, \lambda^j), \tag{20}$$

where $I(\lambda^i, \lambda^j)$ is an indicator function defined by:

$$I(\lambda^i, \lambda^j) = \begin{cases} 1, & if\ \lambda^i, \lambda^j\ violate\ C\text{-}1\ or\ C\text{-}2 \\ 0, & otherwise. \end{cases} \tag{21}$$

The ambiguous degree $d \in \left[0, \binom{N}{2}\right]$, with $d = 0$ indicating no violation of the sufficient conditions, and $d = \delta$ indicating δ pairs of appliances that violate the conditions. We will investigate the impact of ambiguous degree on the accuracy of state recovery and energy disaggregation in Section 7.4.3.

6 Meter Deployment Optimization

To achieve unambiguous state recovery and predictable energy disaggregation, we can partition the appliances into exclusive sub-groups and ensure that within each subgroup, the power model of appliances meets the sufficient conditions.

6.1 Meter Deployment Under Topology Constraint

We have noticed that it would be inconvenient or even impossible to arbitrarily partition appliances, because it is desirable not to change existing circuitry in a building, e.g., the dryer and the television may be located in different rooms. Therefore, appliances can be grouped together for measuring

purpose when they are either proximal to share the same socket, or connected to the same power line so that the low-cost meter [5] can be used easily. We call such a requirement as *network topology constraint*. When two appliances cannot be clustered in the same sub-group, we call them *incompatible* (otherwise, they are compatible). Correspondingly, network topology constraint can be formally defined as:

$$\textbf{\textit{C-3}}: \text{ If appliances } i \text{ and } j \text{ are incompatible and } \lambda^i \in \Lambda^k, \text{ then } \lambda^j \notin \Lambda^k, \forall k. \tag{22}$$

Remark 3. While the need to consider topology constraint is obvious, optimization with network topology constraint may cause two concerns: the availability of topology information and the overhead of re-installing meters when a new appliance is introduced or when an existing appliance moves. Regarding the former concern, the topology information should be known to facility management of commercial buildings; in the worst case, a coarse-grained topology information, such as the level of floors, is easily obtained. Regarding the latter concern, it is practically unnecessary to re-deploy meters if the changes are not significant. This is due to the fact that once deployed, meters can tolerate a certain level of appliance changes, without causing obvious performance degradation in energy disaggregation results, as shown later in Section 7.4.

In order to lower the cost, it is desirable to minimize the total number of sub-groups, or equivalently, the minimum number of meters. This optimization problem can be formulated as follow:

Input. A group of N appliances with power model Λ, and network topology constraints **C-3**.

Output. The minimum number of sub-groups, denoted by $\{\Lambda^1, \Lambda^2, \ldots, \Lambda^n\}$.

$$\begin{aligned} \min_{\{\Lambda^1,\Lambda^2,\ldots,\Lambda^n\}} \quad & n \\ s.t. \quad & \bigcup \Lambda^i = \Lambda, \\ & \Lambda^i \bigcap \Lambda^j = \varnothing, i \neq j, \\ & \Lambda^i = \{\text{appliances fit } \textbf{\textit{C}}\text{-1}, \textbf{\textit{C}}\text{-2\&}\textbf{\textit{C}}\text{-3}\}. \end{aligned} \tag{23}$$

Note that we exclude ***C-0*** from the constraints since it may depend on human behaviour and thus cannot be controlled. 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***, and the two appliances are compatible. It is easy to see that the problem is equivalent to the *clique-covering* problem, which has been proven to be NP-hard [29]. 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. Heuristically, the vertices with larger degrees may have a better chance of resulting in a smaller number of cliques. Thus, the search starts from the vertex with the largest degree, until all vertices are covered. Obviously, a resulted cluster is a clique in the graph. Since each vertex represents an appliance, a clique represents a sub-group, within which unambiguous state recovery is ensured (if ***C-0*** holds in the dataset). We omit the pseudocode of the greedy algorithm for briefty.

6.2 Complexity of Deployment Optimization

Lemma 2. *The computational complexity of meter deployment optimization is lower bounded by $O(N^2)$ and upper bounded by $O(N^3)$, where N is the number of appliances.*

Proof. For a sub-group of n appliances, $1 \leq n \leq N$, to check the sufficient conditions of unambiguous state recovery and network topology constraints, i.e., ***C-1***, ***C-2*** and ***C-3***, the number of comparisons is up to:

$$\underbrace{\binom{n}{2} \times 2}_{\text{for } \textbf{\textit{C}}\text{-1}} + \underbrace{n}_{\text{for } \textbf{\textit{C}}\text{-2}} + \underbrace{\binom{n}{2}}_{\text{for } \textbf{\textit{C}}\text{-3}} = 3n^2/2 + n/2. \tag{24}$$

In the greedy clique-covering algorithm, a clique represents a sub-group of appliances. As a clique gets smaller, the number of appliances in the sub-group gets smaller and the number of comparisons decreases. Hence, two extreme situations yield the minimum and maximum numbers of comparisons, respectively: 1) all the appliances can be cluster into one group, and 2) each appliance becomes a sub-group. The first situation results in N^2 comparisons, and the second $O(N^3)$ comparisons, because

$$\sum_{n=1}^{N} \left(3n^2/2 + n/2\right) = \frac{N(N+1)^2}{2} \sim O(N^3). \tag{25}$$

Therefore, the total number of comparisons in the meter deployment optimization is in $[O(N^2), O(N^3)]$. □

6.3 Parallel Energy Disaggregation

After optimizing deployment of meters by solving (23), we can obtain n sub-groups of appliances and then perform energy disaggregation for each of them. Here, to speed up energy disaggregation among the n sub-groups, a simple yet efficient strategy is performing parallel energy disaggregation. With notations and outcomes from (23), the pseudo code of the straightforward parallel process is shown in Algorithm 2, in which *parfor* is the same parallel *for* loop in Algorithm 1.

Algorithm 2. Parallel State Recovery among Sub-groups

Input: Power models of sub-groups $\{\Lambda^1, \Lambda^2, \ldots, \Lambda^n\}$, aggregated power vectors of sub-groups $\{x^1, x^2, \ldots, x^n\}$.
Output: Recovered state matrices of sub-groups $\{S^1, S^2, \ldots, S^n\}$.

```
parfor i ← 1 : n do
    x ← x^i
    Λ ← Λ^i
    Get S^i by applying Algorithm 1.
end parfor
return {S^1, S^2, ..., S^n}.
```

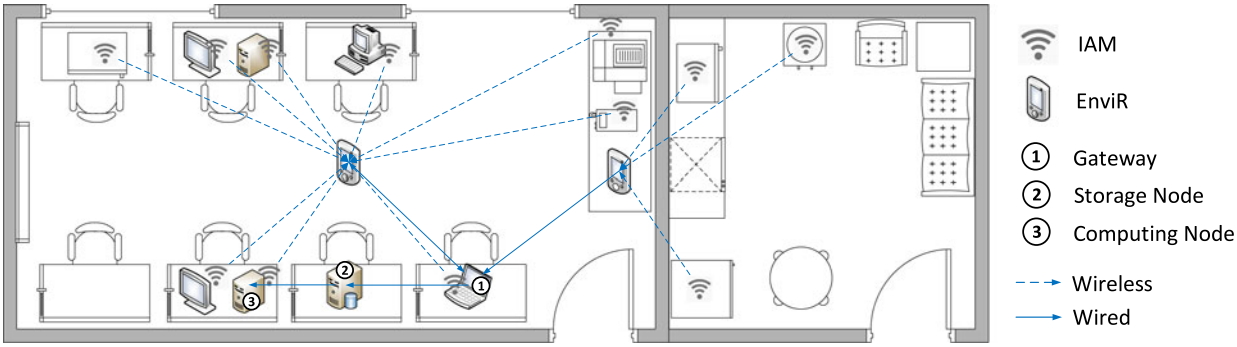

Fig. 6. Distributed appliances power monitoring platform at the fifth floor of ECS building, University of Victoria, Canada. In the platform, we use a dedicated server (computing node) to execute the energy disaggregation algorithm (parallel SSER). The power consumption data of each appliance is collected via an individual appliance monitor (IAM) and stored in a storage node via the gateway. Note that the energy data of individual appliance is used only as the ground truth to evaluate the energy disaggregation results. The computing node retrieves the aggregated energy data from the storage node when performing energy disaggregation. The detailed architecture and implementation of the power monitoring platform can be found in our demo paper [30].

After recovering the state matrix for each sub-group of appliances, the energy consumption of individual appliances can be estimated with the rated power information. Thus, we eventually achieve energy disaggregation for all the appliances. Furthermore, as we have mentioned, based on the power deviations of appliances, the upper and lower bounds of estimated energy consumption of individual appliances can also be provided.

7 Experimental Evaluation

In this section, benchmark NIALM approaches are selected and performance metrics are introduced. We first evaluate different approaches using real-world trace data from household appliances. Then synthetic load data from large-scale appliances are generated by Monte Carlo simulations and used to evaluate different approaches.

7.1 Benchmark Approaches and Performance Metrics

To compare the performance, we also implemented other NIALM approaches. Using the power information the same as in the SSER model, we implemented 1) a signature based approach, the Least Square Estimation (LSE) based integer programming method [13], and 2) a state transition based approach, the iterative Hidden Markov Model [18].

To evaluate the error of energy disaggregation, the *Disaggregation Error* is usually used [18], [19], [23]. Furthermore, to validate our state recovery strategies, we also evaluate the accuracy of recovered appliances' states via *Hamming Loss* [31]. Accordingly, we use $1 - Disaggregation\ Error$ and $1 - Hamming\ Loss$ to get the *accuracy* of energy disaggregation and state recovery, respectively. The performance metrics are defined as follows.

7.1.1 Energy Disaggregation Accuracy (EDA)

It indicates the accuracy of assigning correct power values to corresponding appliances,

$$EDA := 1 - \frac{\sum_{n=1}^{N} \left\| p_r^n - s^n p^n \right\|_1}{2\|x\|_1}, \tag{26}$$

in which N is the total number of appliances, p_r^n, s^n and p_n represent the real power consumption vector, the recovered state vector, and the rated power vector of the nth appliance, respectively, and x is the aggregated power vector.

7.1.2 State Recovery Accuracy (SRA)

It indicates the accuracy of recovering the states of appliances,

$$SRA := 1 - \frac{\sum_{n=1}^{N} \left\| s_r^n - s^n \right\|_1}{N \cdot T}, \tag{27}$$

in which s_r^n and s^n represent the real state vector and the recovered state vector of the nth appliance, respectively, and N, T represent the number of appliances and the number of samples, respectively.

7.2 Real-World Evaluations

As shown in Fig. 6, we construct a real-world distributed appliances power monitoring platform using the off-the-shelf solution provided by *Current Cost* (www.currentcost.com). Individual power consumption data (real power measurements) of 12 household appliances (including microwave, refrigerator, desktop, and so on) are collected using the individual appliance monitors (IAMs) and pour into our database via the *EnviR* display (a sink node) every 10 seconds. The IAMs, EnviR display, gateway (a laptop) as well as our client application (named *CC Data Collector*) are illustrated in Fig. 7.

One-week load data was collected and used for performance evaluation. The prior knowledge of appliances' rated power is learned from the users' manual or from the public website [6]. The power deviations are estimated from the collected power readings of each appliance. To address the concern on the accuracy of estimated power deviation, we will perform robustness test in Section 7.4.

The deployment and performance results with our SIALM approach are summarized in Table 2, and the performance comparison between SIALM and NIALM approaches is summarized in Table 3. In addition, as illustrated in Fig. 8, we can check the *overall energy disaggregation accuracy* of the three methods, which indicates the energy contribution of each appliance to the total energy

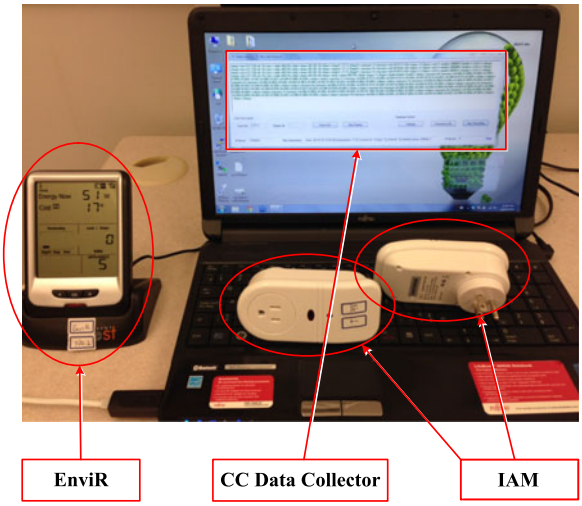

Fig. 7. Plug-and-play wireless power meters (IAM), sink node and display (EnviR), gateway (a laptop) and client application (CC Data Collector) developed for real-time power consumption data collection [30].

consumption in the whole time period. From the results, we can see that with only three extra meters, the performance of SIALM approach is much better than that of the two NIALM approaches. Specifically, the accuracy in both energy disaggregation and appliance state recovery with our SIALM approach is over 90 percent.

7.3 Scalability Test with Synthetic Data

Note that currently public datasets for energy disaggregation (such as [19]) are from small-scale appliances, and most of them are circuit-oriented and do not contain detailed appliances' information. Therefore, in this section, synthetic data of large-scale appliances are generated and numerical evaluations are performed to validate the efficiency of our SIALM approach.

There is no standard model to generate load data of appliances, since the energy consumption actually results from a complex process related to human activities. We thus apply the *Monte Carlo* simulation to generate the load data for large-scale appliances:

TABLE 2
Deployment and Performance Results from the SIALM Approach with Real Data

Sub-group	*Appliance*	*EDA*	*SRA*
1	Water Cooler Microwave-1 Refrigerator	94.66%	96.47%
2	Microwave-2 Monitor-1 Laptop Desktop-1	91.19%	90.01%
3	Monitor-2 Desktop-2 Printer	93.40%	95.17%
4	Desktop-3 Kettle	93.55%	96.79%

Note: Appliances in the first sub-group are located in the lounge room, while others are in a laboratory.

TABLE 3
Performance Results from Experiments with Real Data

Methods \ Metrics		*#Meters*	*EDA*	*SRA*
NIALM	*LSE*	1	33.40%	45.67%
	HMM	1	55.27%	67.47%
SIALM	*SSER*	4	**93.20%**	**94.61%**

Note: The performance of LSE and HMM with multiple meters is listed in Table 5.

- Given the number of appliances, N, the number of operating modes of each appliance is uniformly assigned between 1 and $\Delta(\Delta \geq 1)$, and $K(K < N/2)$ pairs of incompatible appliances are manually chosen.
- Given the lowest power (p_{min}) and the highest power (p_{max}) of all appliances, 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 unambiguous state recovery necessary constraints with the appliance's power model. Referring to the sufficient conditions, partition the N appliances into multiple sub-groups using the greedy algorithm introduced in Section 6.
- For each sub-group of appliances: every appliance reports its current operating mode (a random number uniformly distributed between 1 and the maximum mode number of the appliance) and power reading (a random number uniformly distributed between the appliance's power bounds). It also reports 0 if its state is *off*. Then the aggregated power reading of all 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[4] with parameter τ.

The values of parameters used to generate the synthetic data are listed in Table 6.

7.3.1 Effectiveness of SIALM Approach

The synthetic data is used to test the performance of NIALM and SIALM approaches. The accuracy of energy disaggregation and state recovery, and the number of meters are summarized in Table 4. We can see that, with the help of a few extra meters, the accuracy of energy disaggregation can be significantly improved. From the first case where the number of appliances is 30, we can see that when the sufficient conditions for unambiguous state recovery are hold, the accuracy of recovered appliance states with SSER model can reach as high as 100 percent.

4. Poisson distribution is a good model for situations where the total number of items is large and the probability that each individual item changes its state is small. It has been broadly adopted to simulate events related to human behavior, such as the number of telephone calls in a telephone system and the number of cars on high way.

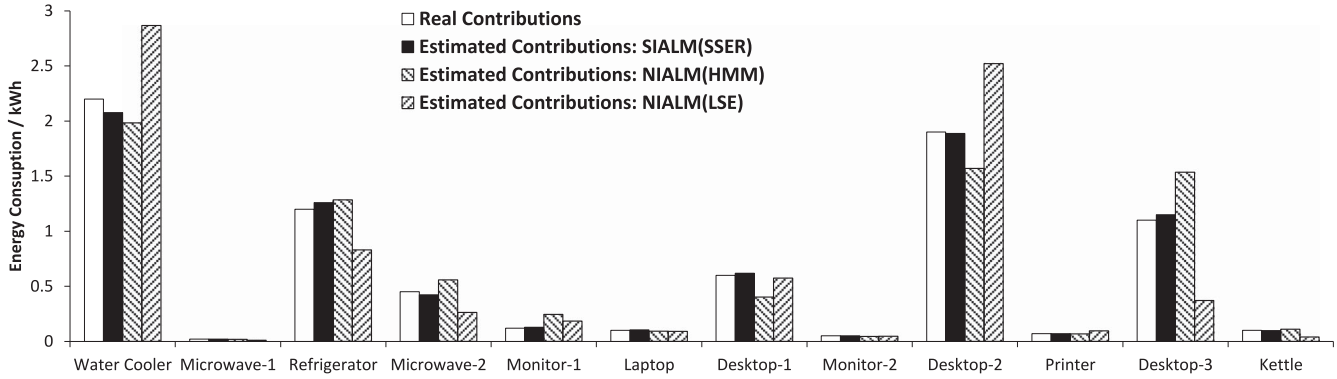

Fig. 8. Real and estimated energy contributions of each appliance to the total consumption for one week.

TABLE 4
Performance Results of Appliances Deployment Optimization and Energy Disaggregation: NIALM versus SIALM

method	metric	30 Appliances			50 Appliances			100 Appliances			200 Appliances		
		#meter	EDA	SRA	#meter	EDA	SRA	#meter	EDA	SRA	#meter	EDA	SRA
NIALM	*LSE*	1	52.85%	40.81%	1	47.61%	38.43%	1	43.31%	37.30%	1	40.35%	39.47%
	IHMM	1	62.02%	56.23%	1	59.77%	58.40%	1	59.01%	58.33%	1	52.95%	52.81%
SIALM	*Random*	7	*90.91%	*85.02%	10	77.08%	66.20%	18	75.97%	70.94%	35	75.46%	68.74%
	Optimal	7	***98.60%**	***100%**	10	**89.70%**	**89.94%**	18	**87.48%**	**85.45%**	35	**85.66%**	**84.30%**

Note: For EDA *&* SRA *resulted from SSER model, $*$ is labeled for the optimal value and others are approximate ones from the approximate parallel optimizations. It may not be fair to only compare the performance of SIALM with multiple meters with that of NIALM with one meter. Therefore, we also provide the performance of LSE and HMM with multiple meters and list the results in Table 5.*

TABLE 5
Performance Results of Energy Disaggregation for Sub-Groups of 30 Appliances Using Different Disaggregation Models

model	group	Subgroup-1		Subgroup-2		Subgroup-3		Subgroup-4		Subgroup-5	
		EDA	SRA	EDA	SRA	EDA	SRA	EDA	SRA	EDA	SRA
LSE		83.69%	82.70%	84.97%	83.39%	84.20%	82.71%	80.63%	79.20%	84.11%	82.09%
IHMM		91.43%	93.39%	93.40%	96.01%	93.55%	95.10%	91.90%	92.82%	95.47%	96.88%
SSER		**97.91%**	**100%**	**98.55%**	**100%**	**98.70%**	**100%**	**96.63%**	**100%**	**99.28%**	**100%**

Note: As shown in Table 4, there are seven sub-groups resulted from the deployment optimization of 30 appliances. Here we only show the first five ones for space limitation, and the performance values from the other two are quite similar to those from the first five.

7.3.2 Effectiveness of Sufficient Conditions

By looking into the accuracy of state recovery from each sub-groups, we can validate the effectiveness of sufficient conditions for unambiguous state recovery. To fairly compare our SSER model with others, we calculate the performance metrics of different models (LSE, IHMM and SSER) within each sub-group of appliances using only one meter. The performance results for the first five sub-groups of 30 appliances are shown in Table 5. From the comparison, we can see that using the same number of meter for the same group of appliances, the accuracy of state recovery and energy disaggregation from our SSER model is higher than that from the other two disaggregation models. Furthermore, as shown in the table, since the power models of each sub-group of appliances follow the sufficient conditions, the accuracy of state recovery can reach one hundred percent, while for the other two models, it is not.

7.3.3 Effectiveness of Deployment Optimization

To validate the effectiveness of SIALM with respect to meter deployment, we first compare the optimal deployment with the random deployment using the same number of meters. The performance of two deployment strategies is shown in the last two rows of Table 4. Compared with the random meter deployment, the optimal deployment strategy is considerably better than the random deployment.

To further validate the effectiveness of the deployment strategies under sufficient conditions, we compare the number of meters required in the SIALM approach with that

TABLE 6
Parameter Settings for Power Model and Load Data Generation

Parameter	*Setting*
Number of Appliances (N)	$30, 50, 100$, or 200
Number of Operating Modes (Δ)	3
Number of Incompatible Pairs (K)	$2, 4, 8$, or 10
Sampling Rate of Meters	0.1 Hz
Total Simulation length	24 hours
Lowest Appliance Power (p_{min})	$20\,w$
Highest Appliance Power (p_{max})	$2,000\,w$
Power Range Ratio (r)	$[0.05, 0.15]$
Poisson Parameter (τ)	180

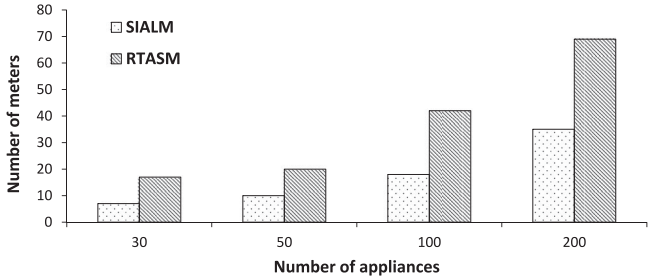

Fig. 9. Number of meters needed for unambiguous state recovery in SIALM and RTASM, respectively.

from a real-time appliance state monitoring method in [24] that can also provide unambiguous state recovery. The numbers of meters needed for unambiguous state recovery with SIALM and RTASM are shown in Fig. 9. We can see that, for the same number of appliances, SIALM needs much fewer meters than RTASM. Since RTASM only uses current information (real-time load data) to achieve unambiguous state monitoring, it does not allow any power overlap between any pair of appliances in a sub-group. In contrast, by utilizing all historical load data, SIALM can tolerate power overlaps of appliances to some degree (referring to ***C-1*** and ***C-2***), and thus can reduce the number of meters while still ensuring unambiguous state recovery.

7.4 Robustness Test

In practice, our approach might be subject to the following concerns: 1) the estimated power deviation of an appliance may not be accurate, 2) the network topology may change, e.g., some new appliances may be introduced or appliances may be moved from one place to another. It is also interesting to explore the performance of our approach when the sufficient conditions do not hold.

7.4.1 Impact of Inaccurate Power Deviation

We replace θ with $\rho \cdot \theta$ in model (13) and model (23), so that the estimated power deviations can be narrowed down or widened up by regulating ρ. The value of ρ is changed from 1.0 to 1.5, causing parametric errors of power deviation up to 50 percent. We do not consider the situation where appliances' power ranges are narrowed, since intuitively a user can always widen an appliance's power range if she is not sure about the real values.

From the results shown in Table 7, we can see that the accuracy does not change much when the parameter error varies, indicating that our method is robust to parameter estimation.

TABLE 7
Performance of SIALM Approach with Inaccurate Estimation of Power Deviation

ρ	30 Appliances			100 Appliances		
	#meters	EDA	SRA	#meters	EDA	SRA
1.0	7	98.60%	100%	18	87.48%	85.45%
1.1	8	94.49%	91.77%	20	84.59%	76.48%
1.2	9	90.82%	88.60%	24	82.45%	75.06%
1.3	9	89.20%	86.77%	25	81.01%	75.31%
1.4	10	86.03%	84.25%	26	79.93%	75.12%
1.5	10	84.61%	82.44%	29	77.90%	73.15%

TABLE 8
Performance of SIALM Approach with Newly Added Appliances

#New Appliances	30 Appliances		100 Appliances	
	EDA	SRA	EDA	SRA
0	98.60%	100%	87.48%	85.45%
1	97.77%	98.20%	86.92%	84.31%
2	95.15%	93.96%	87.04%	85.10%
3	95.45%	94.13%	86.42%	83.87%
4	93.70%	93.87%	84.93%	84.20%
5	92.12%	93.30%	84.01%	84.61%

7.4.2 Impact of Network Topology Changes

For the situation where new appliances are added into the network topology, if the power meters are plug-and-play as the ones used in our case, we can easily re-deploy them according to the newly computed deployment solution. In cases where the power meters are hard to change, we further test the performance of our approach when new appliances are added while the meters remain the same. Specifically, with the optimal meter deployment, we add a small number of new appliances into the existing network without changing the meters. In particular, one to five new appliances generated with parameters in Table 6 are randomly added into existing appliance sub-groups. We then test the accuracy of energy disaggregation with the existing meters. Note that we do not consider the situation where some existing appliances are removed, because under such situation the sufficient conditions for unambiguous state recovery still hold and it does not bring any impact on the accuracy.

Part of the results with our SIALM approach are shown in Table 8. We can see that the accuracy of energy disaggregation and the accuracy of state recovery only slightly deteriorate. Similarly, we also test the case when a small number of appliances move from one place to another and did not observe a large performance degradation.

7.4.3 Impact of Violating Sufficient Conditions

To explore the performance of our approach when the sufficient conditions do not hold, we further test its performance with various ambiguous degrees. In the tests, after obtaining the optimal deployment solution, we arbitrarily choose $3 \sim 10$ sub-groups of appliances and adjust their ambiguous degree to specified values by manually changing appliances' power models.

The results of accuracy versus average ambiguous degree ($\bar{d}$) from the tests are summarized in Table 9. According to the results, our approach can tolerate the violation of

TABLE 9
Performance of SIALM Approach with Sufficient Condition Violations

$\bar{d}$	30 Appliances		100 Appliances	
	EDA	SRA	EDA	SRA
0	98.60%	100%	87.48%	85.45%
1	97.30%	97.69%	87.32%	84.97%
2	96.06%	95.71%	86.88%	85.51%
5	85.45%	86.13%	85.12%	85.04%
10	−−	−−	69.01%	71.70%

sufficient conditions to a certain degree. For example, when $\bar{d} = 1, 2, 5$, the accuracy of energy disaggregation and state recovery does not decrease too much. Nevertheless, our approach suffers and cannot return a result when the ambiguous degree reaches 10 in a group of 30 appliances. Note that this extreme case is unlikely in practice, and it implies that the appliances are nearly indistinguishable in our model, making energy disaggregation extremely hard.

In summary, our SIALM approach is robust in the presence of inaccurate estimation of power deviation and small number of appliance changes. When the sufficient conditions do not hold, our approach can tolerate the ambiguous degree to a modest level.

8 Further Clarification

To better understand the nature of NIALM, several misconceptions deserve particular clarification.

- *Granularity:* While NIALM is normally referred to as "inferring energy consumption of individual appliances based on their aggregated values", the practical meaning of "individual appliances" should not be explained as each single device. For example, it is neither realistic nor necessary to estimate the energy consumption of every light bulb when multiple bulbs are controlled by the same switch. In this case, they should be considered as a single "virtual" device for the purpose of NIALM.
- *Completeness:* NIALM needs to track the energy consumption of all appliances. This expectation is unrealistic for most mobile, ad-hoc devices, such as smart phones, which may be plugged anytime to any available sockets. It is extremely hard, if not impossible, to track the energy consumption of those low-power devices.
- *Sparsity:* Sparsity does not suggest that every appliance' state switching is sparse. In fact, some heating appliances with thermostat can have a change every several seconds. Instead, sparsity refers to a *system-wide* phenomenon, i.e., the total number of state switching events in the whole system is significantly smaller than the total number of samples over a long time period. This phenomenon has been used as the optimization objective to solve many practical problems [27].

9 Conclusions

Monitoring the energy consumption of major appliances in commercial buildings is critical in energy management and demand response programs. Nevertheless, prohibiting metering cost has become the main barrier for fine-grained energy monitoring. To tackle this challenge, energy disaggregation becomes a promising technique and has recently attracted heavy investment in industry.

In this paper, we proposed a semi-intrusive appliance load monitoring approach to energy disaggregation for large-scale appliances. Instead of using only one meter, multiple meters were distributed in the power network to collect the aggregated load data from sub-groups of appliances. Based on a simple power model, we established a sparse switching event recovering model and proposed a parallel optimization algorithm to recovery appliance states from the aggregated load data. We also provided the sufficient conditions for unambiguous state recovery of multiple appliances. Furthermore, under the sufficient conditions and network topology constraints, a minimum number of meters was searched for via a greedy clique-covering algorithm.

To evaluate the performance of our SIALM approach, we not only used the real-world trace data from household appliances, but also applied Monte Carlo simulation to generate load data from large-scale appliances. The energy disaggregation accuracy and state recovery accuracy were compared with those from two benchmark NIALM approaches. The results showed that the SIALM approach can provide high-precision accuracy for appliance state recovery and improve the accuracy of energy disaggregation with a small number of extra meters.

Acknowledgments

This work was supported by the Natural Sciences and Engineering Research Council of Canada (No. 195819339), the Natural Science Foundation of China (Nos. 61373152, 61272437, 61472236), Innovation Program of Shanghai Municipal Education Commission (Nos. 13ZZ131, 14ZZ150), Foundation Key Project of Shanghai Science and Technology Committee (Nos. 12JC1404500, 13JC1403503), Project of Shanghai Science and Technology Committee (No. 14110500800).

References

[1] L. L. N. Laboratory. (2014). Energy flow chart of U.S. [Online]. Available: https://flowcharts.llnl.gov/
[2] S. Darby, "The effectiveness of feedback on energy consumption," *A Rev. DEFRA Literature Metering, Billing Direct Displays*, vol. 486, p. 2006, 2006.
[3] B. Neenan, J. Robinson, and R. Boisvert, "Residential electricity use feedback: A research synthesis and economic framework," *Retrieved Oct.*, vol. 26, no. 2011, p. 3, 2009.
[4] DOE. (2012). 2015 U.S. buildings energy end-use splits. [Online]. Available: http://buildingsdatabook.eren.doe.gov/TableView.aspx?table=1.1.5
[5] CurrentCost. (2015). Products introduction. [Online]. Available: http://www.currentcost.com/products.html
[6] EPA. (2015). A tool of home product finder by energy star. [Online]. Available: http://www.energystar.gov/productfinder/
[7] 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 *Proc. 12th ACM Int. Conf. Ubiquitous Comput.*, 2010, pp. 139–148.
[8] G. W. Hart, "Nonintrusive appliance load monitoring," *Proc. IEEE*, vol. 80, no. 12, pp. 1870–1891, Dec. 1992.
[9] M. Zeifman and K. Roth, "Nonintrusive appliance load monitoring: Review and outlook," *IEEE Trans. Consum. Electron.*, vol. 57, no. 1, pp. 76–84, Feb. 2011.
[10] A. Zoha, A. Gluhak, M. A. Imran, and S. Rajasegarar, "Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey," *Sensors*, vol. 12, no. 12, pp. 16 838–16 866, 2012.
[11] C. Laughman, K. Lee, R. Cox, S. Shaw, S. Leeb, L. Norford, and P. Armstrong, "Power signature analysis," *IEEE Power Energy Mag.*, vol. 1, no. 2, pp. 56–63, Mar.-Apr. 2003.
[12] M. Dong, P. C. Meira, W. Xu, and W. Freitas, "An event window based load monitoring technique for smart meters," *IEEE Trans. Smart Grid*, vol. 3, no. 2, pp. 787–796, Jun. 2012.
[13] K. Suzuki, S. Inagaki, T. Suzuki, H. Nakamura, and K. Ito, "Nonintrusive appliance load monitoring based on integer programming," in *Proc. SICE Annu. Conf.*, 2008, pp. 2742–2747.
[14] H. Shao, M. Marwah, and N. Ramakrishnan, "A temporal motif mining approach to unsupervised energy disaggregation," in *Proc. 1st Int. Workshop Non-Intrusive Load Monitoring*, 2012.

[15] S. B. Leeb, S. R. Shaw, and J. L. Kirtley Jr, "Transient event detection in spectral envelope estimates for nonintrusive load monitoring," *IEEE Trans. Power Del.*, vol. 10, no. 3, pp. 1200–1210, Jul. 1995.
[16] S. R. Shaw, S. B. Leeb, L. K. Norford, and R. W. Cox, "Nonintrusive load monitoring and diagnostics in power systems," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 7, pp. 1445–1454, Jul. 2008.
[17] S. Makonin, I. V. Bajic, and F. Popowich, "Efficient sparse matrix processing for nonintrusive load monitoring (nilm)," in *Proc. 2nd Int. Workshop Non-Intrusive Load Monitoring*, 2014.
[18] O. Parson, S. Ghosh, M. Weal, and A. Rogers, "Non-intrusive load monitoring using prior models of general appliance types," in *Proc. 26th Conf. Artif. Intell.*, 2012, pp. 356–362.
[19] J. Z. Kolter and M. J. Johnson, "Redd: A public data set for energy disaggregation research," in *Proc. SustKDD Workshop Data Mining Appl. Sustainability*, 2011, pp. 1–6.
[20] J. Z. Kolter and T. Jaakkola, "Approximate inference in additive factorial hmms with application to energy disaggregation," in *Proc. Int. Conf. Artif. Intell. Statist.*, 2012, pp. 1472–1482.
[21] H. Kim, M. Marwah, M. F. Arlitt, G. Lyon, and J. Han, "Unsupervised disaggregation of low frequency power measurements," in *Proc. SIAM Int. Conf. Data Mining*, 2011, pp. 747–758.
[22] J. Z. Kolter, S. Batra, and A. Ng, "Energy disaggregation via discriminative sparse coding," in *Proc. Adv. Neural Inf. Process. Syst.*, 2010, pp. 1153–1161.
[23] M. Figueiredo, B. Ribeiro, and A. M. de Almeida, "On the regularization parameter selection for sparse code learning in electrical source separation," in *Proc. Adaptive Natural Comput. Algorithms*, 2013, pp. 277–286.
[24] 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 *Proc. IEEE 3rd Int. Conf. Smart Grid Commun.*, 2012, pp. 25–30.
[25] 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 *Proc. 6th Int. Workshop Knowl. Discovery Sens. Data*, 2012, pp. 34–42.
[26] Y. Wang, X. Hao, and et. al., "Monitoring massive appliances by a minimal number of smart meters," *ACM Trans. Embedded Comput. Syst.*, vol. 13, no. 2s, p. 56, 2014.
[27] S. Osher, A. Solé, and L. Vese, "Image decomposition and restoration using total variation minimization and the h 1," *Multiscale Modeling Simul.*, vol. 1, no. 3, pp. 349–370, 2003.
[28] CVX. (2015). Matlab software for disciplined convex programming. [Online]. Available: http://cvxr.com/cvx/
[29] C. Lund and M. Yannakakis, "On the hardness of approximating minimization problems," *J. ACM*, vol. 41, no. 5, pp. 960–981, 1994.
[30] G. Tang, J. Chen, C. Chen, and K. Wu, "Smart saver: a consumer-oriented web service for energy disaggregation," in *Proc. IEEE Int. Conf. Data Mining Workshop*, 2014, pp. 1235–1238.
[31] A. Elisseeff and J. Weston, "A kernel method for multi-labelled classification," in *Proc. Adv. Neural Inf. Process. Syst.*, 2001, pp. 681–687.

Guoming Tang (S'12) received the bachelor's and master's degrees from the National University of Defense Technology, China in 2010 and 2012, respectively. He is currently working toward the PhD degree in the Department of Computer Science, University of Victoria, BC, Canada. His research interests include green computing, computational sustainability, machine learning, and data mining. He is a student member of the IEEE.

Kui Wu (S'98-M'02-SM'07) received the BSc and MSc degrees in computer science from Wuhan University, China in 1990 and 1993, respectively, and the PhD degree in computing science from the University of Alberta, Canada, in 2002. He joined the Department of Computer Science at the University of Victoria, Canada in 2002 and is currently a professor there. His research interests include cloud computing, mobile and wireless networks, distributed systems, and network performance evaluation. He is a senior member of the IEEE.

Jingsheng Lei received the BS degree in mathematics from the Shanxi Normal University in 1987, and the MS and PhD degree in computer science from Xinjiang University in 2000 and 2003, respectively. He is a professor and the dean of the College of Computer Science and Technology, Shanghai University of Electronic Power. His research interests include machine learning, data mining, pattern recognition, and Cloud computing. He has published more than 100 papers in international journals or conferences. He serves as an editor-in-chief of the *Journal of Computational Information Systems*. He is the member of the Artificial Intelligence and Pattern Recognition Technical Committee of the China Computer Federation (CCF), the member of the Machine Learning Technical Committee of the Chinese Association of Artificial Intelligence (CAAI), and the member of the Academic Committee of ACM Shanghai chapter.