Computer Networks 117 (2017) 42–51

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Occupancy-aided energy disaggregation

Guoming Tang[a,*], Zhen Ling[b], Fengyong Li[c], Daquan Tang[d], Jiuyang Tang[d]

[a] Computer Science Department, University of Victoria, Victoria, B.C, Canada
[b] School of Computer Science and Engineering, Southeast University, Nanjing, China
[c] College of Computer Science and Technology, Shanghai University of Electric Power, Shanghai, China
[d] Science & Technology on Information Systems Engineering Lab, National University of Defense Technology, Changsha, China

ARTICLE INFO

Article history:
Received 26 August 2016
Revised 11 November 2016
Accepted 28 November 2016
Available online 30 November 2016

Keywords:
Energy disaggregation
NILM
Occupancy detection
Computational complexity

ABSTRACT

Energy disaggregation helps to identify major energy guzzlers in the house without introducing extra metering cost. It motivates users to take proper actions for energy saving and facilitates demand response programs. To reduce the computational complexity of pure energy disaggregation, we propose an occupancy-aided energy disaggregation (OAED) approach in this paper. Specifically, we make use of the occupancy information (whether or not the house/room is occupied by users) and classify the whole time interval into occupied and unoccupied periods. In unoccupied periods, we perform lightweight energy approximation; in occupied periods, we apply energy disaggregation with existing methods. Real-world experiments are conducted in an apartment hosting typical household appliances. Comparing with energy disaggregation without considering occupancy information, our occupancy-aided approach can significantly reduce the computational overhead while ensuring the accuracy of energy disaggregation.

1. Introduction

Based on the energy flow in the United States, the commercial and residential buildings consumed as much as 74% of the total electricity in 2013, and the residential buildings alone consumed over 38% of the overall usage [1]. To effectively cut down the electricity bill for residential customers as well as facilitate demand response (DR) programs for the utilities, it is meaningful to monitor the energy consumption of individual appliances in residential houses [2]. With the feedback of appliance level energy consumption, the residential consumers can not only be motivated to cut down energy usage automatically [3], but also make more intelligent decisions towards energy saving [4,5].

Energy disaggregation, also known as non-intrusive load monitoring (NILM), aims to identify major energy guzzlers by referring to the measurements only from a single meter of the household. As no extra metering cost is incurred, the technique is regarded as the most economical way to obtain appliance level energy information and has been well explored since 1980s [6]. Because of the cost saving, energy disaggregation has drawn tremendous efforts and investments from both academia and industry, and a broad spectrum of approaches have been attempted [7,8].

While broadly investigated, energy disaggregation is still challenging and has much room to improve. As one of the key problems, the computational complexity of energy disaggregation is usually high. For the disaggregation approaches based on appliances' electrical signatures, it has to traverse the whole load curve to search for the appliance signature one by one [9,10]. Other disaggregation approaches based on state transition of appliances, such as hidden Markov model (HMM) as well as its variants, are NP-hard when they discover the most likely state sequences of appliances [11,12]. Consequently, approximations and heuristics were developed to reduce the complexity, leading to less accurate results.

Is there any way that can reduce the computational complexity while still ensuring the accuracy for energy disaggregation? In a typical household, the occupancy states (whether someone is at home) play a significant role in energy consumption. As a real-world case shown in Fig. 1, we can observe that: i) most appliances' activities are triggered during occupied periods of the house; ii) there are quite few appliances (only one in our case) running in unoccupied periods of the house. Therefore, when performing energy disaggregation, we can focus on the occupied periods while roughly estimating the energy consumption of certain appliances running in unoccupied periods. By cutting out the unoccupied periods from the whole time interval, we can significantly reduce the computational complexity, especially when the unoccupied periods are dominated.

* Corresponding author. Fax: +12504725708.
E-mail address: guoming@uvic.ca (G. Tang).

http://dx.doi.org/10.1016/j.comnet.2016.11.019

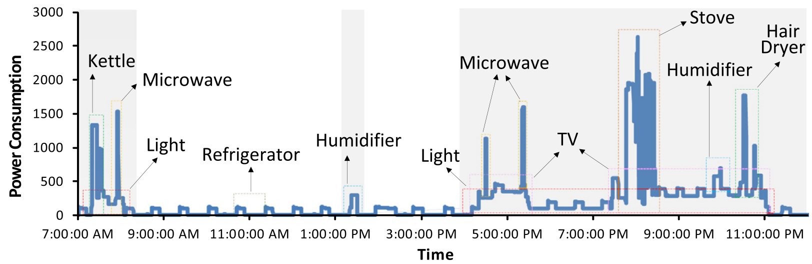

Fig. 1. Correlation between appliances' activities and the occupancy states of the house. Occupied periods of the house are illustrated by shaded areas.

To reduce the computational complexity of energy disaggregation while ensuring its accuracy, we make the following contributions in this work:

- We propose an occupancy-aided energy disaggregation (OAED) framework based on non-intrusive occupancy inference for the residential house. The framework draws a general picture of how occupancy information can be leveraged in the energy disaggregation process and can be easily adopted by any pure energy disaggregation approaches.
- To reduce the computational complexity of energy disaggregation by OAED, we first infer the occupancy states of the house based on the analysis of collected load curve data; then by applying occupancy inference, we provide energy approximation for the appliances running in the unoccupied periods, and perform energy disaggregation techniques for the appliances working in occupied periods.
- We conduct extensive experimental evaluations over a real-world residential energy monitoring platform and validate the effectiveness and robustness of the occupancy-aided approach. The performance results demonstrate that our occupancy-aided approach can much reduce the computational overhead with ensured accuracy of energy disaggregation.

The rest of our paper is organized as follow: In Section 2, we review current energy disaggregation approaches based on their methodology and utilized information. In Section 3, we draw the framework of occupancy-aided energy disaggregation (OAED). The technical details about occupancy inference, energy approximation and energy disaggregation are introduced in Section 4, Section 5 and Section 6, respectively. Section 7 and Section 8 show how we implement the OAED idea over a real-world energy monitoring platform and how we evaluate the effectiveness and robustness of the OAED approach. The paper is concluded in Section 10.

2. Related work

In this section, we discuss the related work in two aspects: (i) pure energy disaggregation in which no context information is utilized, and (ii) context-aware energy disaggregation where the contextual knowledge e.g., appliance location and related information of occupants' activities, is taken into consideration.

2.1. Pure energy disaggregation

2.1.1. Signature based approaches

Signature based energy disaggregation approaches exploit various appliance running features, e.g., real or reactive power, current, and voltage, to classify and identify various household appliances. In detail, the aforementioned features are regarded as the signatures of different appliances in load curve and treated as the target events with certain searching strategies. Then, the detected appliance activities (events) are assigned with (estimated) energy values to obtain the disaggregation result. As one of the initial tries in [6], the two-dimensional signature space with respect to both real and reactive powers was explored for energy disaggregation. Recently, the authors in [10] investigated the high potential of high frequency electromagnetic interference (EMI) signals in distinguishing ON/OFF states of appliances, and the authors in [13] explored the feasibility of V-I trajectory (i.e., the mutual locus of instantaneous voltage and current waveforms) in distinguishing difference appliances. In addition to the signatures extracted from the time domain, spectral features in frequency domain were also studied for energy disaggregation [9,14–16]. Generally speaking, the signature based approaches went through two procedures: appliance ON/OFF state switching event detection using the load curve data [17–19] and specific appliance identification in the household through these detected events [9,14,15].

2.1.2. State transition based approaches

The state transition patterns of appliances were investigated for energy disaggregation. As a typical model to imitate the activities of individual appliances, Hidden Markov Models (HMMs) were widely adopted. Specifically, the ON/OFF state transition probabilities of appliances are estimated through parameter learning methods (e.g., Maximum Likelihood or Maximum A Posteriori). Then the observed power emissions are feed into the inference algorithms (e.g., Viterbi) along with the state transition probabilities to infer the hidden states of individual appliances. The HMM based energy disaggregation approaches were utilized in [20–22], in which each appliance was treated as a single hidden Markov chain. In addition, the variants of HMM were proposed for energy disaggregation, e.g., FHMM [23,24], AFMAP [12], and CFHSMM [11], where particular generation patterns of the power emissions, characteristics of appliance activities, and even human factors/involvements were incorporated with the original HMM model. Recently, there were also methods making use of particular properties during appliances' state transition, such as the sparsity of state switching [25–27]. A comprehensive survey of state transition based methods can be found in [8].

2.2. Context-aware energy disaggregation

Besides appliance running features and state information, additional information can be used to enhance the accuracy of energy disaggregation, referred to as *context-aware energy disaggregation* in this paper. In [28], location information of appliances was taken into consideration when deriving appliance level energy consumption, and the accuracy of energy disaggregation was empirically validated across multiple datasets. In addition, an indirect approach for infer occupancy information was applied using WiFi/Bluetooth signals collected from the smartphones and wearable devices of occupants. A hybrid system called AARAP was proposed in [29] to exploit various mobile sensors so as to infer high-level activities information of residential customers and reduce the number of candidate appliances for energy disaggregation. To this end, the residential customers have to carry their smartphones all

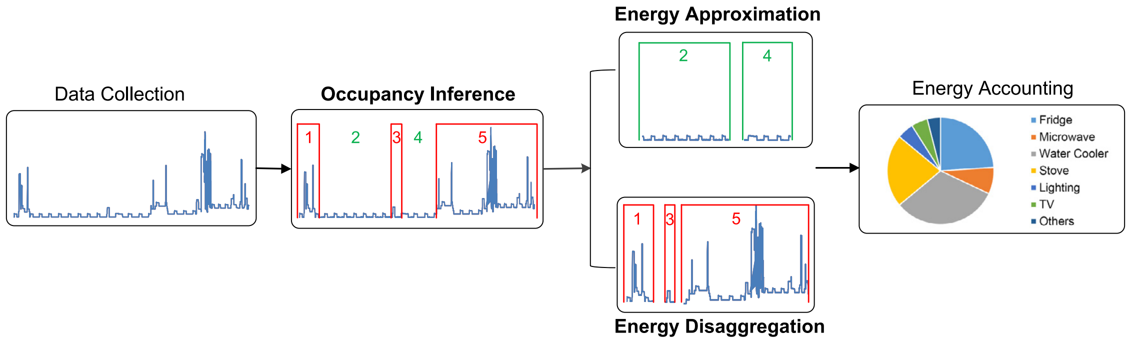

Fig. 2. Framework of occupancy-aided energy disaggregation.

the time for collecting persistent WiFi signals and smartphone accelerometer data. As extra examples employing information from other sensors, [30] used the electromagnetic field (EMF) sensing in the surrounding to determine the state changes of appliances; the authors of [31] placed low-cost sensors near the household appliances and estimated their power consumption referring to the ambient signals of the sensors.

Different from the above approaches in contextual information collection, we derive the occupancy information directly from the load curve data rather than relying on any other hardware devices. It significantly reduces the hardware cost incurred in context-aware energy disaggregation and eliminate the influence from inaccurate hardware measurements.

3. OAED framework

In this section, we introduce our framework for occupancy-aided energy disaggregation. The framework consists of three major components given the aggregated load curve data: occupancy inference, energy approximation, and energy disaggregation, as shown in Fig. 2. Specifically, based on the collected aggregated load curve data (power consumption signal in this paper) of a house, we first infer the occupancy states of the house by analyzing load curve variations or recovering human actions (Section 4). According to the inferred occupancy states, we estimate the energy consumption for the appliances working during the unoccupied periods by using either approximation based on coarse-grained power information or a lightweight energy disaggregation scheme as described in Section 5. Meanwhile, we perform transitional energy disaggregation for the appliances running in the occupied periods as illustrated in Section 6. Finally, we derive the appliance level energy consumption during the whole time interval.

4. Occupancy inference

There are diverse approaches to inferring occupancy based on either intrusive or non-intrusive sensing. For intrusive sensing based occupancy detection, additional sensors or dedicated devices are installed in the house, e.g., passive infrared/motion sensors [32], acceleration sensors [33], or cameras [34]. In this work, to avoid extra sensor/device deployment, we apply the non-intrusive occupancy detection (NIOD). Specifically, we focus on two types of NIOD approaches, i.e., supervised [35] and unsupervised [36] NIOD approaches using load curve data.

4.1. Supervised NIOD

Given that the ground-truth occupancy information can be collected and used as a training dataset, we can apply supervised learning for occupancy detection. Instead of applying sophisticated supervised machine learning methods [37], we adopt a simple yet effective NIOD approach based on the previous work [35].

We utilize statistic power features during a time period to detect occupancy state. We first evenly divide the whole time interval (e.g., one day) into smaller time windows. Considering a time window τ starting from time 1 to n, we represent the aggregated power readings[1] of a house by the following vector:

$$x := [x_1, x_2, \ldots, x_n]^\intercal. \tag{1}$$

Thus, the tth ($1 \le t \le n$) element of x denotes the aggregated power value of the house at time t. Then, three metrics are defined to infer the occupancy states of the house in τ:

- Average power value: $N_{avg} := avg(x)$;
- Standard power deviation: $N_{std} := std(x)$;
- IQR power range: $N_{rng} := Q_3(x) - Q_1(x)$.

Here, Q_1 and Q_3 denote the lower and upper quartiles, defined as the 25th and 75th percentiles of all samples, respectively. Note that in previous work of [35], the third metric was max power range and defined by the difference between maximum and minimum power values. We modify the metric to interquartile (IQR) range due to its robustness to the influence of outliers in load curve data.

Then, the occupancy state o (which is a binary variable) of the house in the time window τ is determined by the following conditions:

$$o_\tau = \begin{cases} 1, & N_{avg} \ge P_{avg} \text{ or } N_{std} \ge P_{std} \text{ or } N_{rng} \ge P_{rng} \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

where P_{avg}, P_{std}, and P_{rng} represent the thresholds for average power, power deviation, and power range, respectively.

A following key step is to perform extensive empirical experiments to collect the data and select appropriate values for the thresholds in (2) to accurately infer occupancy. Given the training dataset of ground-truth occupancy information, the cross-validation technique can be applied to choose the most effective thresholds in our context. Once the tuned threshold values are obtained, they can be used for occupancy inference for the specific house, based on the determination conditions in (2).

Nevertheless, there are situations where the training data is difficult to obtain, e.g., without collaboration of occupants. Therefore, in this case, an unsupervised NIOD approach without relying on training process may be more helpful.

[1] In this paper, we make use of the real power readings while the applied principles can be easily adapted to other signals.

4.2. Unsupervised NIOD

Our unsupervised NIOD approach in this work is based on [36] that relies on the load curve data and appliance power consumption information.

Preliminaries: Assume that a list of m (major) appliances appearing in the house is given by the appliance set M ($|M| = m$). Considering that most household appliances work under multiple operating modes, we further assume that appliance i can work in m_i different modes. Then, the rated power (or mean power) of appliance i working under mode j can be denoted as $\mu_j^{(i)}$, and corresponding power deviation can be estimated as $\delta_j^{(i)}$. Thus, the power consumption of appliance i working under mode j at any arbitrary instant falls into $[\mu_j^{(i)} - \delta_j^{(i)}, \mu_j^{(i)} + \delta_j^{(i)}]$ with a high probability. For the ON/OFF state of mode j of appliance i at time t, we denote it by $s_j^{(i)}(t)$, where $s_j^{(i)}(t) = 1$ if appliance i is running under mode j at time t; otherwise $s_j^{(i)}(t) = 0$.

With the notations in preliminaries, an optimization problem aiming at decoding the mode states of all appliances over the interested time interval τ, ($|\tau| = n$) can be formulated as:

$$\begin{aligned} \min_{s_j^{(i)}(t)} \quad & \sum_{t=1}^{n}\sum_{i=1}^{m}\sum_{j=1}^{m_i}\left|s_j^{(i)}(t+1) - s_j^{(i)}(t)\right| \\ s.t. \quad & (\mu_j^{(i)} - \delta_j^{(i)})s_j^{(i)}(t) \le x_t \le (\mu_j^{(i)} + \delta_j^{(i)})s_j^{(i)}(t), \\ & s_j^{(i)}(t) \in \{0, 1\}, \\ & \sum_{j=1}^{m_i} s_j^{(i)}(t) \le 1. \end{aligned} \tag{3}$$

In the above optimization problem, the objective function represents the *total variation norm* (or TV norm) of appliances' mode states between adjacent time instants. The TV norm was originally used as an approach to signal denoising, while we apply it here to decode the appliances' mode states. The first three constraints in the problem represent the facts that: (i) the aggregate power value at any time instant falls into the summation of all appliances' lower power bounds and upper power bounds, (ii) the mode state variables are binary values, and (iii) each appliance can only work under one mode at each time instant, respectively.

By solving the state decoding problem in (3), we can get the mode states of each appliance at each time instant. Thus, the mode switching events of each appliance over the time of τ can be easily derived by performing subtraction between the state values of adjacent time instants. Next, among the mode switching events, we find out those human-activated ones (i.e., the mode switching event has to be accomplished with the human action) with prior knowledge about appliances' mode switching. Then, the house is determined to be occupied at the moment when the human-activated switching events (named *recovered human actions*) are detected.

5. Energy approximation

We have observed that there are quite few appliances running during the periods when the house is unoccupied, as shown in Fig. 1. Those appliances that are left running in the unoccupied periods are usually "always-on" appliances, such as refrigerator and water cooler/heater. Due to the small number of running appliances in unoccupied periods, we can apply simple approximation methods to estimate their energy consumption. Here we provide two kinds of energy approximation.

5.1. ECR based approximation

To estimate the energy consumption of the "always-on" appliances, we can simply use the metric of energy consuming rate (ECR), e.g., hourly energy usage. This metric can be found in the user's manual or technique specifications of the appliance, and is usually evaluated by the ENERGY STAR agency [38]. In case that such information is not readily-available, we can also easily estimate it using extra devices, e.g., plug-in power meters.

In the unoccupied periods without interference from residential customers, the ECR metric of an appliance is relatively accuracy in evaluating its energy consumption. For example, in our evaluation, the accuracy of energy approximation in the unoccupied periods using hourly energy usage is as high as 87%.

Considering an "always-on" appliance (indexed by k) with energy consuming rate r_k, we can approximate the appliance's energy consumption e_k during the unoccupied periods $\bar{\mathcal{T}}$ as:

$$e_k = r_k \times |\bar{\mathcal{T}}|, \tag{4}$$

where $|\bar{\mathcal{T}}|$ denotes the total length of the unoccupied time periods.

Then, with the ECR information of the "always-on" appliances (either provided by the vendors or measured by the customers), their energy approximation during the unoccupied periods can be easily calculated.

5.2. CO based approximation

Considering that most "always-on" appliances just repeat the "ON-OFF" state transition pattern, we can apply the combinatorial optimization (CO) [6] to estimate their ON/OFF states and then calculate their individual energy consumption.

To estimate the ON/OFF states of the "always-on" appliances, the CO model can be formulated as:

$$\begin{aligned} \min_{s_t^{(i)}} \quad & \sum_{t=1}^{|\bar{\mathcal{T}}|}\left|x_t - \sum_{i=1}^{k}\mu^{(i)}s_t^{(i)}\right| \\ s.t. \quad & s_t^{(i)} \in \{0, 1\}, \end{aligned} \tag{5}$$

where x_t is the aggregated power consumption at time t, k is the number of ("always-on") appliances, $\mu^{(i)}$ is the rated power of the ith appliance, $|\bar{\mathcal{T}}|$ is the total length of unoccupied periods, and $s_t^{(i)}$ is the ON/OFF state of the ith appliances at time t.

After obtaining the ON/OFF states of appliances over the time period $\bar{\mathcal{T}}$, we can approximate their individual energy consumption in terms of their rated powers. Since the number of appliances is small and the multi-mode states of appliances are reduced to ON/OFF ones, the CO based energy approximation is, to some extent, lightweight energy disaggregation.

6. Energy disaggregation

To disaggregate energy from multiple appliances in the occupied periods, we need carefully designed disaggregation models/approaches. In this paper, since we focus on the contribution of occupancy information to energy disaggregation, we just adopt existing disaggregation approaches rather than developing new ones.

Two different energy disaggregation approaches are implemented for our testing: the signature based approach using the Least Square Estimation (LSE) model [15] and the state transition based approach applying an iterative HMM model [21].

6.1. Signature based disaggregation

The signature based approach applying LSE model was adopted in [15] for energy disaggregation. The current waveform of each appliance was extracted and stored beforehand, and then used as its signature for appliance identification. In this paper, we make use of the power signal instead of the current waveform, due to (i) the difficulty in obtaining the current waveform in our situation, and (ii) the suitability of power signal for energy disaggregation using our model.

The LSE model for energy disaggregation in the occupied periods $\mathcal{T}$ is formulated as:

$$\begin{aligned} \min_{s_j^{(i)}(t)} \quad & \sum_{t=1}^{|\mathcal{T}|} \left(x_t - \sum_{i=1}^{m} \sum_{j=1}^{m_i} \mu_j^{(i)} s_j^{(i)}(t) \right)^2 \\ s.t. \quad & s_j^{(i)}(t) \in \{0, 1\}, \\ & \sum_{j=1}^{m_i} s_j^{(i)}(t) \le 1, \end{aligned} \tag{6}$$

where x_t is the aggregated power consumption at time t, m is the number of all appliances in the house, m_i is the number of working modes of appliance i, $\mu_j^{(i)}$ is the rated power (or mean power) of appliance i working under the jth mode, $|\mathcal{T}|$ is the total time length of occupied periods, and $s_j^{(i)}(t)$ is the state value (either 1 or 0) of mode j of appliance i at time t.

6.2. State transition based disaggregation

As a state transition based method, the iterative HMM was proposed for energy disaggregation in [21]. We implement this model in three phases:

- Modeling phase: each appliance is modelled as a prior difference HMM, which is defined by:

  $$\lambda := \{A, B, \pi\}, \tag{7}$$

  where A is the prior state transition probability distribution, B is the emission probability distribution, and π is the starting state distribution of the appliance. In particular, (i) A is initialized with the transition probabilities proportional to the time spent in each state, and (ii) for any state change between modes j and k of the ith appliance, its corresponding emission probability in B is defined by a Gaussian distributed power consumption $\mathcal{N}(\mu_j^{(i)} - \mu_k^{(i)}, \delta_j^{(i)} + \delta_k^{(i)})$.
- Training phase: we apply the expectation maximization (EM) algorithm over the collected load curve data. The EM algorithm is initialized with the prior state transition matrix A and individual appliances' rated power. It terminates when a local optima in the log likelihood function is found or the maximum number of iterations (100 in our implementation) is reached.
- Inference phase: the extended Viterbi algorithm shown in [21] was applied to infer each appliance's mode state, considering the constraints of aggregated power and power changes at each time instant.

By applying the above energy disaggregation models, we can get the mode state of each appliance at each time instant and thus can estimate the energy consumption of each appliance by referring to its rated power. Since the length of occupied periods (i.e., $|\mathcal{T}|$) is expected to be much shorter than the whole time interval in consideration, the computational complexity in the disaggregation models can be much reduced.

7. Implementations

In this section, we show our implementation of a smart home energy monitoring platform, under which we collect real-world dataset for occupancy inference and energy disaggregation.

Table 1
Parameter setting for supervised NIOD.

Parameter	*Notations*	*Setting value*
Inference time window	τ	15 min
Average power threshold	P_{avg}	125 W
Standard power deviation threshold	P_{std}	72 W
IQR power range threshold	P_{rng}	108 W

7.1. Experimental platform

We established a smart home energy monitoring platform in an apartment. As illustrated in Fig. 3, the apartment is built with one bedroom, one living room, one kitchen and one bathroom. Note that although the platform is established in a small apartment, it is built with typical rooms and has most representative household appliances.

7.2. Data collection

We collected the power reading data from an apartment using off-the-shelf measuring devices from *CurrentCost*.[2] Two power sensor jaws were installed at the power entrance to measure the aggregated power consumption of the apartment. All the measurement data were sent to a sink node with frequency of 0.1 *Hz* and then forwarded to a data harvest computer. For evaluation purpose, we also record individual power consumption of ten major appliances using plug-in power meters (as shown in Fig. 3). By comparing to the monthly electricity bill, these major appliances under our consideration, including stove, refrigerator, microwave, etc., consume over 85% of the total energy.

We also collected and labeled the ground-truth occupancy information as training dataset for supervised occupancy inference. The Google mobile app named *Google+*[3] was installed on the mobile phones (with GPS module) of each occupant to gather the location information, from which we infer whether or not the occupant is at home. One-month power consumption and occupancy information were collected and used for the performance evaluation in Section 8.

7.3. Parameter setting

The detailed power information (rated power and power deviation) of all appliances under consideration was measured by the plug-in power meter, and used in the procedures of unsupervised NIOD, energy approximation and energy disaggregation. In specific, for supervised NIOD, we carefully trained the inference parameters introduced in Section 4.1 using the ground-truth occupancy information. Ten-fold cross validation was applied to find the most appropriate threshold values, as shown in Table 1. Note that in our occupancy inference, the load silence periods (e.g., when the occupants are sleeping) are recognized/treated as unoccupied periods.

8. Evaluations

In this section, using the real-world data collected from our energy monitoring platform, we perform extensive experiments to evaluate our occupancy-aided energy disaggregation approach and make comparison with the pure energy disaggregation ignoring occupancy states.

[2] www.currentcost.com
[3] www.google.com/mobile/+/

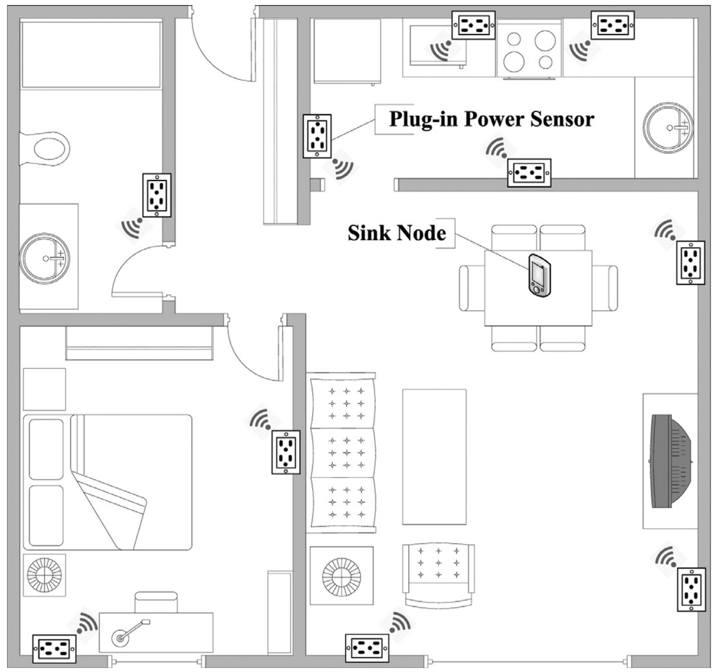

Fig. 3. Floor plan of an apartment and our smart home energy monitoring platform.

8.1. Performance metrics

Based on the results of occupancy inference, we first calculate the true/false positive (*TP*/*FP*), i.e., the number of points that are correctly/incorrectly identified as occupied states, and true/false negative (*TN*/*FN*), i.e., the number of points that are correctly/incorrectly identified as unoccupied states. Then, using *TP*, *FP*, *TN*, and *FN*, we evaluate the accuracy of occupancy inference by the following broadly-used metrics:

- $Precision = \frac{TP}{TP+FP}$ as a measure of exactness,
- $Recall = \frac{TP}{TP+FN}$ as a measure of completeness,
- $F\text{-}measure = \frac{2 \cdot Precision \cdot Recall}{Precision+Recall}$ as a harmonic mean between precision and recall.

To evaluate results of energy disaggregation, we use the performance metric of *energy disaggregation accuracy (EDA)* defined by:

$$EDA := 1 - \frac{\sum_{i=1}^{m} \left\| x^{(i)} - \hat{x}^{(i)} \right\|_1}{\|x\|_1}, \tag{8}$$

where x, $x^{(i)}$, and $\hat{x}^{(i)}$ represent the aggregated power readings, ground-truth power readings of appliance i, and estimated power readings of appliance i, respectively. Thus, the EDA metric indicates the accuracy of assigning correct power values to corresponding appliances along the time line.

8.2. Performance evaluation

8.2.1. Accuracy of occupancy inference

Supervised NIOD: We first validate the effectiveness of supervised non-intrusive occupancy detection introduced in Section 4.1. Using the tuned threshold values given in Table 1, we infer the occupancy state of the apartment during the one-month time interval. Fig. 4 shows an example of the occupancy inference result for a period of one day.

Unsupervised NIOD: Without using the ground-truth occupancy information for training, we also test the unsupervised NIOD approach introduced in Section 4.2. Fig. 5 illustrated the occupancy inference result for the same day shown in Fig. 4.

Table 2
Performance results of occupancy inference: supervised NIOD vs. unsupervised NIOD.

	Supervised	Unsupervised
Precision	87.7%	90.3%
Recall	73.9%	78.8%
F-measure	80.2%	84.2%

Table 3
Accuracy comparison of occupancy-aided energy disaggregation (OAED) and pure energy disaggregation (Pure ED).

Disaggregation model	OAED (Unoccupied Per./Occupied Per./Overall)	Pure ED (Overall)
LSE model	86.33% / 62.17% / **68.20%**	67.01%
HMM model	86.33% / 77.46% / **80.07%**	78.87%

The occupancy inference accuracy for both approaches during the one-month time interval is summarized in Table 2. From the results, we can see that the occupancy inference is relatively accurate in both supervised NIOD and unsupervised NIOD, with an average *F-measure* value of 82.4%.

8.2.2. Accuracy of energy approximation & disaggregation

With the inferred occupancy states from the previous step, we perform an energy approximation for unoccupied periods and energy disaggregation for occupied periods, respectively. Specifically, the ECR based approximation is used for appliance level energy estimation during the unoccupied periods, and both LSE and HMM models are adopted for appliance level energy estimation during the occupied periods. Furthermore, the overall (average) accuracy of both energy approximation and disaggregation during the whole time interval is calculated.

As a comparison, we also perform pure energy disaggregation (using LSE and HMM models) for the same time interval and record the resulted accuracy. The performance results from our occupancy-aided energy disaggregation (OAED) and the pure energy disaggregation (pure ED) are summarized in Table 3.

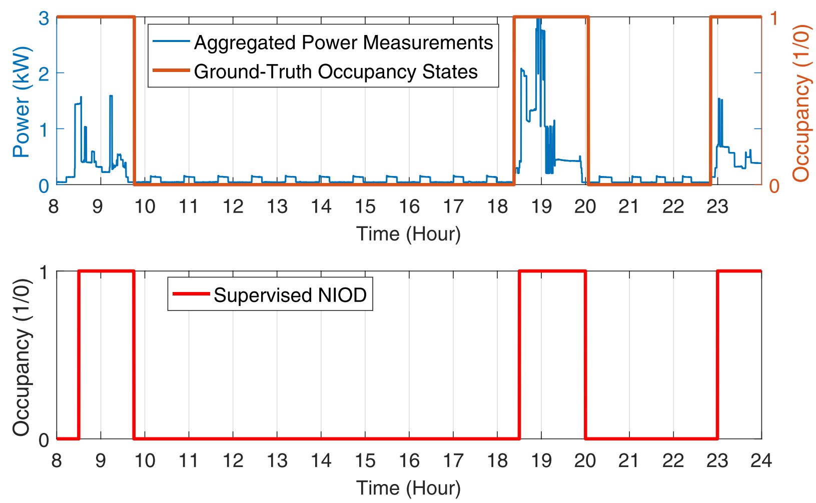

Fig. 4. Ground-truth occupancy states and estimated occupancy states from supervised NIOD.

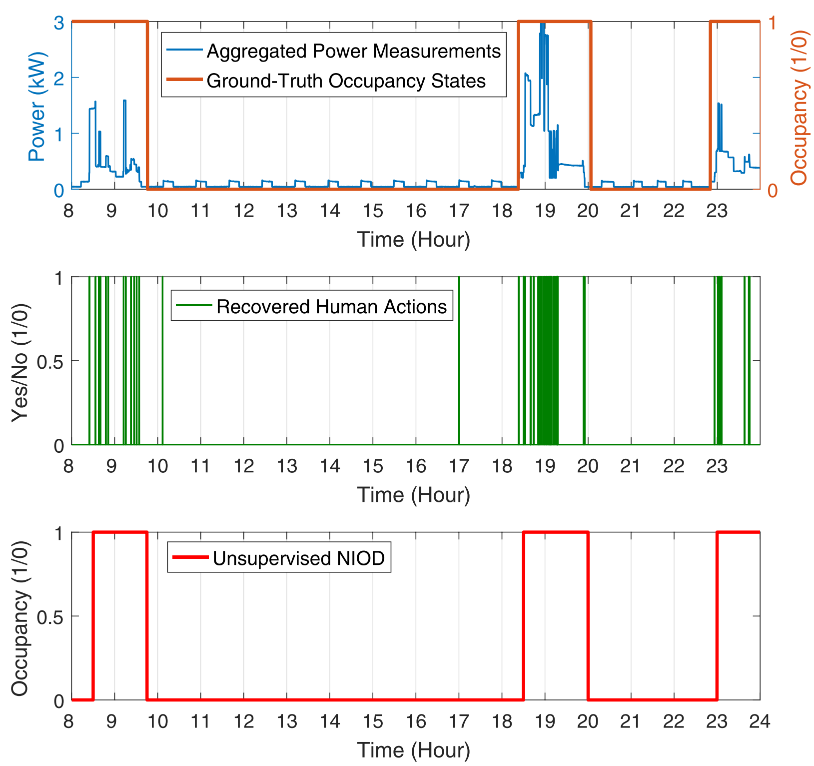

Fig. 5. Ground-truth occupancy states, recovered human actions, and estimated occupancy states from unsupervised NIOD. Note that the human actions recovered from false positive results have been filtered before being used for occupancy inference.

From the comparison, we can find that the two energy disaggregation routines result in comparable accuracy for appliance level energy estimation. Furthermore, in our situation, the occupancy-aided approach is slightly more accurate than the raw one, as the energy approximation in unoccupied periods is quite accurate.

8.2.3. Computational complexity

To measure the computational complexity of the two energy disaggregation routines, we refer to running time as overhead when solving the disaggregation models (i.e., LSE and HMM models). Both energy disaggregation models were implemented and run under MATLAB 8.5, with PC configuration of 32-bit Windows OS, 3.4 GHz CPU and 4 GB RAM. Corresponding elapse time is recorded and shown in Table 4.

Table 4
Overhead comparison of occupancy-aided energy disaggregation (OAED) and pure energy disaggregation (Pure ED).

Disaggregation model	OAED (Elapsed time)	Pure ED (Elapsed time)
LSE model	379.66 s	588.09 s
HMM model	2097.74 s	3051.10 s

According to the overhead shown in Table 4, we can see that: by cutting out the unoccupied periods from the whole time interval, the OAED is much faster than the pure ED, i.e., the computational complexity of energy disaggregation is much reduced. In our case, the running time of each model is shortened by over 30%.

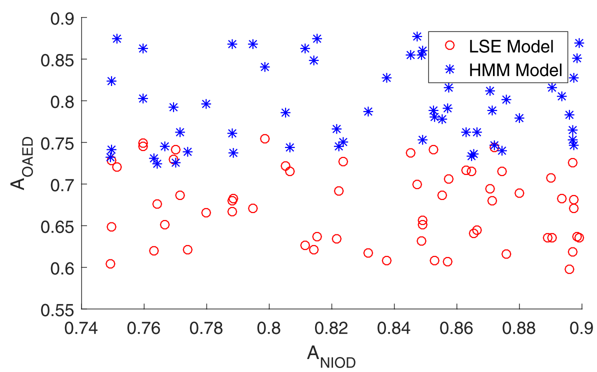

Fig. 6. Correlation between the accuracy of OAED and accuracy of occupancy inference.

8.2.4. Fault-tolerance testing

We further analyze the fault tolerance of our OAED approach to occupancy inference results. In specific, we answer questions like "*will inaccurate occupancy inference degrade the accuracy of energy disaggregation?*" and "*how sensitive the OAED accuracy is to the occupancy inference accuracy?*".

To obtain more sample cases, we enlarge our experimental dataset and extend the time interval from one month to two months. Then, the fault-tolerance testing is performed as the following three steps:

- For the load curve data in each day, we infer the occupancy states of the apartment using the techniques introduced in Section 4 and calculate the metric value of inference accuracy (denoted by A_{NIOD}).
- Then, we perform energy approximation and disaggregation based on the inferred occupancy result of each day and calculate the overall accuracy of energy disaggregation (denoted by A_{OAED}).
- Last, we draw a scatter plot with the values of A_{NIOD} and A_{OAED} (i.e., each point in the scatter plot is denoted by (A_{NIOD}, A_{OAED})) and analyze the correlation between the two metrics.

The scatter plot is shown in Fig. 6. By correlation analysis between the accuracy of OAED and the accuracy of occupancy inference, we can find that there is no apparent correlation between these two metrics, nor is the former sensitive to the latter.

Actually, since inference of non-occupancy always indicates low power consumption where only very few ("always-on") appliances are running, whether or not the inference is correct does not make big a difference to the overall energy disaggregation accuracy. Thus, this makes our OAED approach fault-tolerant to occupancy inference.

9. Further discussions

9.1. Parallelization of computation

By dividing the whole time interval into (occupied/unoccupied) periods, we can perform energy disaggregation for all periods in parallel. Such parallel computing is expected to speed up the energy disaggregation process, which is the bonus when applying our OAED approach.

9.2. Energy disaggregation in unoccupied periods

We can perform sophisticated energy disaggregation approaches instead of approximation for the appliances running in unoccupied periods. This can be adopted under the situation that (i) the number of "always-on" appliances are large, and (ii) the energy consuming rate is either inaccurate or difficult to measure.

9.3. Impact of household types

Our approach is usually effective since most households have periods when there is no one at home or the load curve is "in silence". Nevertheless, for those households that the occupancy periods are dominated (e.g., the house is always occupied), the effectiveness of our approach may not be that obvious.

10. Conclusions and future work

In this paper, we developed an occupancy-aided approach to cut down the computational complexity of energy disaggregation. A three-step routine was proposed for occupancy-aided energy disaggregation: (i) occupancy inference using load curve data, (ii) energy approximation for appliances working in unoccupied periods, and (iii) energy disaggregation for appliances working in occupied periods. We evaluated our approach using real-world datasets collected in an apartment. To validate the effectiveness of our approach, we compare it with existing energy disaggregation methods without utilizing occupancy information. The results showed that the occupancy-aided approach significantly reduces the computational overhead of energy disaggregation without sacrificing accuracy. One possible direction for our future work is looking into the deep learning framework in system modeling [39] and applying it for energy disaggregation purpose.

Acknowledgment

This work was supported by the Natural Science Foundation of China (No. 61373152, 61602295), the Natural Sciences and Engineering Research Council of Canada under the grants (No. 261409-2013), and the Natural Science Foundation of Shanghai (No. 16ZR1413100).

References

[1] Lawrence Livermore National Laboratory, Energy flow: Estimated energy use in 2013, 2014. https://flowcharts.llnl.gov/commodities/energy (accessed 29.05.16).
[2] B. Neenan, J. Robinson, R. Boisvert, Residential electricity use feedback: a research synthesis and economic framework, Electric Power Research Institute (2009) 3. http://www.epri.com/abstracts/Pages/ProductAbstract.aspx?ProductId=000000000001016844 (accessed 11.11.15).
[3] S. Darby, The effectiveness of feedback on energy consumption, a review for DEFRA of the literature on metering, billing and direct displays, 486, Environmental Change Institute, University of Oxford, 2006. http://www.usclcorp.com/news/DEFRA-report-with-appendix.pdf (accessed 11.11.15).
[4] J. Tang, G. Tang, K. Wu, Minnet: toward more intelligent smart home energy management systems with fewer sensors, Int. J. Sens. Netw. 20 (4) (2016) 252–263.
[5] R. Malekian, D.C. Bogatinoska, A. Karadimce, N. Ye, J. Trengoska, W.A. Nyako, A novel smart eco model for energy consumption optimization, Elektronika ir Elektrotechnika 21 (6) (2015) 75–80.
[6] G.W. Hart, Nonintrusive appliance load monitoring, Proc. IEEE 80 (12) (1992) 1870–1891.
[7] M. Zeifman, K. Roth, Nonintrusive appliance load monitoring: review and outlook, IEEE Trans. Consum. Electron. 57 (1) (2011) 76–84.
[8] A. Zoha, A. Gluhak, M.A. Imran, S. Rajasegarar, Non-intrusive load monitoring approaches for disaggregated energy sensing: a survey, Sensors 12 (12) (2012) 16838–16866.
[9] M. Dong, P.C. Meira, W. Xu, W. Freitas, An event window based load monitoring technique for smart meters, IEEE Trans. Smart Grid 3 (2) (2012) 787–796.
[10] S. Gupta, M.S. Reynolds, 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, ACM, 2010, pp. 139–148.
[11] H. Kim, M. Marwah, M.F. Arlitt, G. Lyon, J. Han, Unsupervised disaggregation of low frequency power measurements, in: SIAM International Conference on Data Mining (SDM), SIAM, 2011, pp. 747–758.
[12] J.Z. Kolter, T. Jaakkola, Approximate inference in additive factorial hmms with application to energy disaggregation, in: International Conference on Artificial Intelligence and Statistics, 2012, pp. 1472–1482.

[13] T. Hassan, F. Javed, N. Arshad, An empirical investigation of vi trajectory based load signatures for non-intrusive load monitoring, IEEE Trans. Smart Grid 5 (2) (2014) 870–878.

[14] H. Lam, G. Fung, W. Lee, A novel method to construct taxonomy electrical appliances based on load signatures, IEEE Trans. Consum. Electron. 53 (2) (2007) 653–660.

[15] K. Suzuki, S. Inagaki, T. Suzuki, H. Nakamura, K. Ito, Nonintrusive appliance load monitoring based on integer programming, in: SICE Annual Conference, 2008, IEEE, 2008, pp. 2742–2747.

[16] H.-H. Chang, K.-L. Lian, Y.-C. Su, W.-J. Lee, Power-spectrum-based wavelet transform for nonintrusive demand monitoring and load identification, IEEE Trans. Ind. Appl. 50 (3) (2014) 2081–2089.

[17] C. Laughman, K. Lee, R. Cox, S. Shaw, S. Leeb, L. Norford, P. Armstrong, Power signature analysis, Power Energy Mag. IEEE 1 (2) (2003) 56–63.

[18] L.K. Norford, S.B. Leeb, Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms, Energy Build. 24 (1) (1996) 51–64.

[19] H. Shao, M. Marwah, N. Ramakrishnan, A temporal motif mining approach to unsupervised energy disaggregation: Applications to residential and commercial buildings, Association for the Advancement of Artificial Intelligence (AAAI), AAAI, 2013.

[20] S. Makonin, I.V. Bajic, F. Popowich, Efficient sparse matrix processing for nonintrusive load monitoring (nilm), The Second International Workshop on Non-Intrusive Load Monitoring, 2014.

[21] O. Parson, S. Ghosh, M. Weal, A. Rogers, Non-intrusive load monitoring using prior models of general appliance types., Association for the Advancement of Artificial Intelligence (AAAI), AAAI, 2012.

[22] T. Zia, D. Bruckner, A. Zaidi, A hidden markov model based procedure for identifying household electric loads, in: IEEE Industrial Electronics Society, 2011, pp. 3218–3223.

[23] J.Z. Kolter, M.J. Johnson, Redd: a public data set for energy disaggregation research, in: proceedings of the SustKDD workshop on Data Mining Applications in Sustainability, 2011, pp. 1–6.

[24] M. Al Abassi, K. Johnsson, J. Karlsson, T.B. Schön, J. Wågberg, Ssgmi greenely report energy disaggregation with low resolution data, Energy 3 (2015) 5.

[25] S. Makonin, F. Popowich, I.V. Bajic, B. Gill, L. Bartram, Exploiting hmm sparsity to perform online real-time nonintrusive load monitoring, IEEE Trans Smart Grid PP (99) (2015) 1–11.

[26] G. Tang, K. Wu, J. Lei, J. Tang, A simple model-driven approach to energy disaggregation, in: Proceedings of IEEE International Conference on Smart Grid Communications (SmartGridComm), IEEE, 2014, pp. 566–571.

[27] G. Tang, K. Wu, J. Lei, A distributed and scalable approach to semi-intrusive load monitoring, IEEE Trans. Parallel Distrib. Syst. 27 (6) (2016) 1553–1565.

[28] A.S. Uttama Nambi, A. Reyes Lua, V.R. Prasad, Loced: Location-aware energy disaggregation framework, in: Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments, ACM, 2015, pp. 45–54.

[29] N. Roy, N. Pathak, A. Misra, Fine-grained appliance usage and energy monitoring through mobile and power-line sensing, Pervasive Mob. Comput. 30 (2016) 132–150.

[30] A. Rowe, M. Berges, R. Rajkumar, Contactless sensing of appliance state transitions through variations in electromagnetic fields, in: Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-efficiency in Building, ACM, 2010, pp. 19–24.

[31] Y. Kim, T. Schmid, Z.M. Charbiwala, M.B. Srivastava, Viridiscope: design and implementation of a fine grained power monitoring system for homes, in: Proceedings of the 11th International Conference on Ubiquitous Computing (UbiComp), ACM, 2009, pp. 245–254.

[32] Y. Agarwal, B. Balaji, R. Gupta, J. Lyles, M. Wei, T. Weng, Occupancy-driven energy management for smart building automation, in: Proceedings of the 2nd Workshop on Embedded Sensing Systems for Energy-Efficiency Buildings, ACM, 2010, pp. 1–6.

[33] M. Jin, R. Jia, Z. Kang, et al., Presencesense: zero-training algorithm for individual presence detection based on power monitoring, in: Proceedings of the 1st Conference on Embedded Systems for Energy-Efficient Buildings, ACM, 2014, pp. 1–10.

[34] A. Kamthe, L. Jiang, M. Dudys, A. Cerpa, Scopes: Smart Cameras Object Position Estimation System, in: European Conference on Wireless Sensor Networks, Springer, Cork, Ireland, 2009, pp. 279–295.

[35] D. Chen, S. Barker, A. Subbaswamy, D. Irwin, P. Shenoy, Non-intrusive occupancy monitoring using smart meters, in: Proceedings of the 5th Workshop on Embedded Systems For Energy-Efficient Buildings, ACM, 2013, pp. 1–8.

[36] G. Tang, K. Wu, J. Lei, W. Xiao, The meter tells you are at home! non-intrusive occupancy detection via load curve data, in: Proceedings of the IEEE International Conference on Smart Grid Communications (SmartGridComm), IEEE, 2015, pp. 897–902.

[37] W. Kleiminger, C. Beckel, T. Staake, S. Santini, Occupancy detection from electricity consumption data, in: Proceedings of the 5th Workshop on Embedded Systems For Energy-Efficient Buildings, ACM, 2013, pp. 1–8.

[38] Environmental Protection Agency (EPA), Product finder: Find and compare products, 2016 (http://www.energystar.gov/productfinder/) (accessed 10.07.16).

[39] X. Jin, J. Shao, X. Zhang, W. An, R. Malekian, Modeling of nonlinear system based on deep learning framework, Nonlinear Dyn. 84 (3) (2016) 1327–1340.

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

Zhen Ling is an assistant professor at the School of Computer Science and Engineering at the Southeast University, Nanjing, China. He received the BS degree (2005) and PhD degree (2014) in Computer Science from Nanjing Institute of Technology and Southeast University, China, respectively. He joined Department of Computer Science at the City University of Hong Kong from 2008 to 2009 as a research associate, and then joined Department of Computer Science at the University of Victoria from 2011 to 2013 as a visiting scholar. His research interests include network security, privacy, and smartphone security.

Fengyong Li received the M.S. degree in School of Information and Engineering in 2010 from Zhengzhou University, and the Ph.D. degree in School of Communication and Information Engineering in 2014 from Shanghai University. He is currently a lecturer of the College of Computer Science and Technology, Shanghai University of Electric Power, Shanghai, China. His research interests include multimedia security, privacy preserving.

Daquan Tang received the bachelor degree (1994), master degree (2002) and PhD degree (2009) from the National University of Defense Technology (NUDT), China, respectively. He is currently a Professor at College of Information System and Management of NUDT. He is the senior member of the Chinese Institute of Electronics. His research interests include information management and system, big data analysis, and social computing.

Jiuyang Tang received the Ph.D. degree from the National University of Defense Technology (NUDT), Changsha, China. He is currently a Professor at the College of Information System and Management, NUDT. His current research interests include big data analysis and artificial intelligence.