



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.

© 2016 Elsevier B.V. All rights reserved.

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].

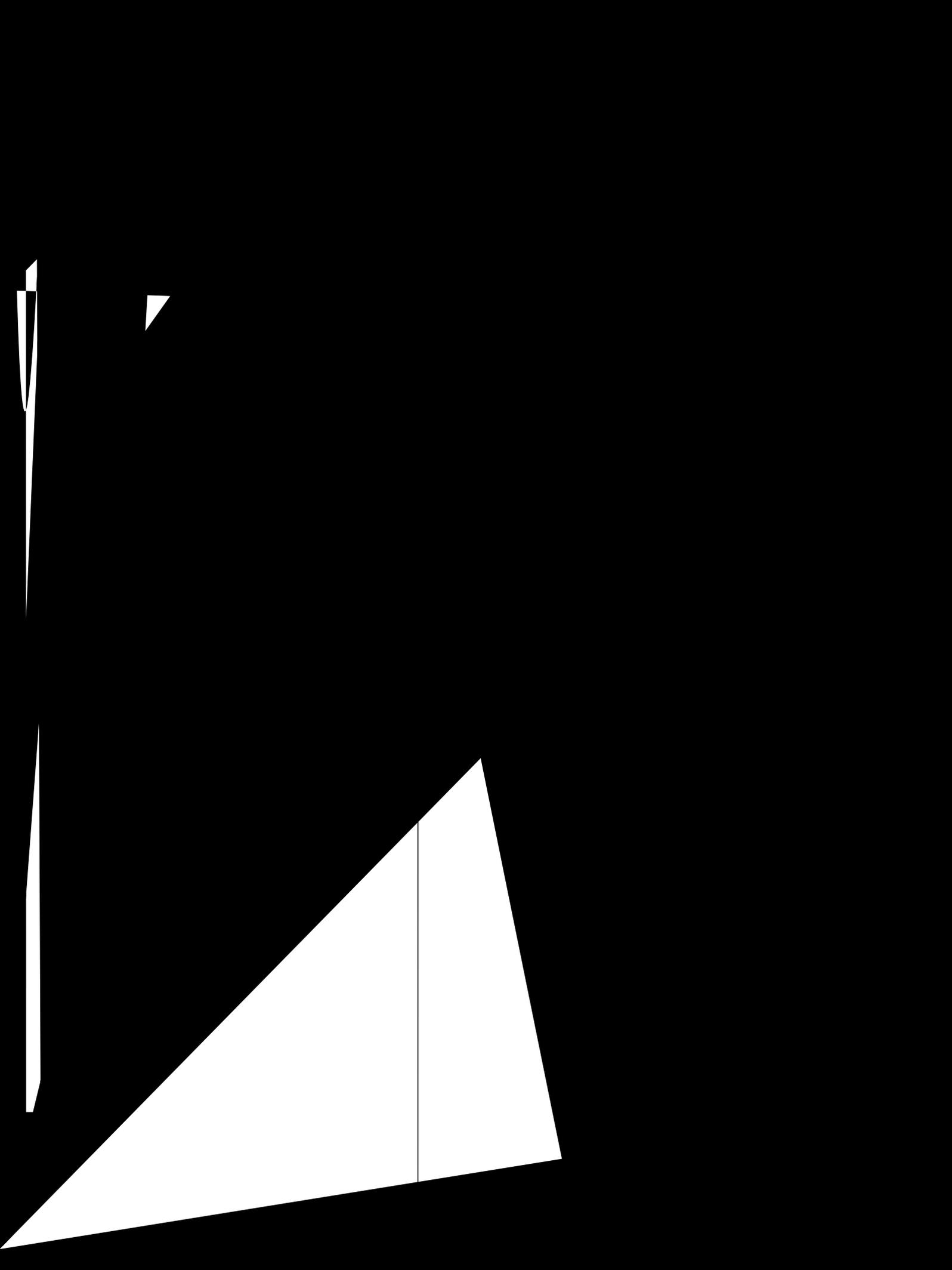
Energy disaggregation, also known as non-intrusive load monitoring (NILM), aims to identify major energy guzzlers by referring to the measurements only from a single meter of the household. As no extra metering cost is incurred, the technique is regarded as the most economical way to obtain appliance 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].

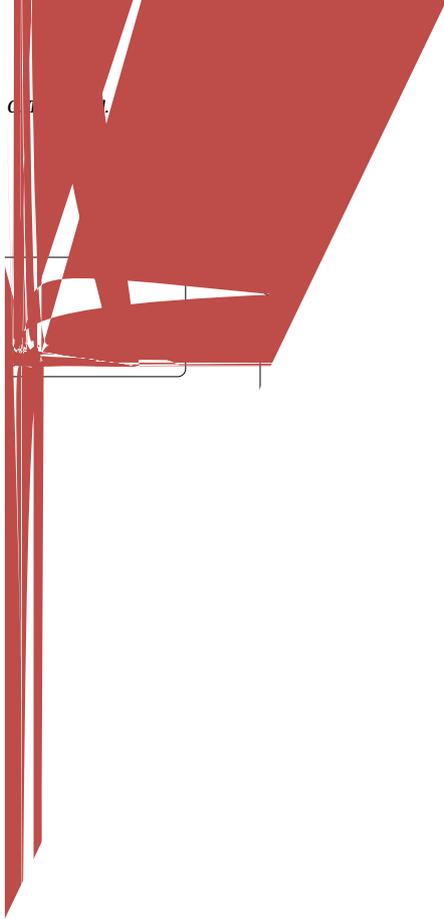
While broadly investigated, energy disaggregation is still challenging and has much room to improve. As one of the key problems, the computational complexity of energy disaggregation is usually high. For the disaggregation approaches based on appliances' electrical signatures, it has to traverse the 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).





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:

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

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 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 $\bar{\tau}$ as:

$$e_k = r_k \times |\bar{\tau}|, \quad (4)$$

where $|\bar{\tau}|$ 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:

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

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 i th appliance, $|\bar{\tau}|$ is the total length of unoccupied periods, and $s_t^{(i)}$ is the ON/OFF state of the i th appliances at time t .

After obtaining the ON/OFF states of appliances over the time period $\bar{\tau}$, 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:

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

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 j th 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:

$$\lambda := \{A, B, \pi\}, \quad (7)$$

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 i th 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/+/

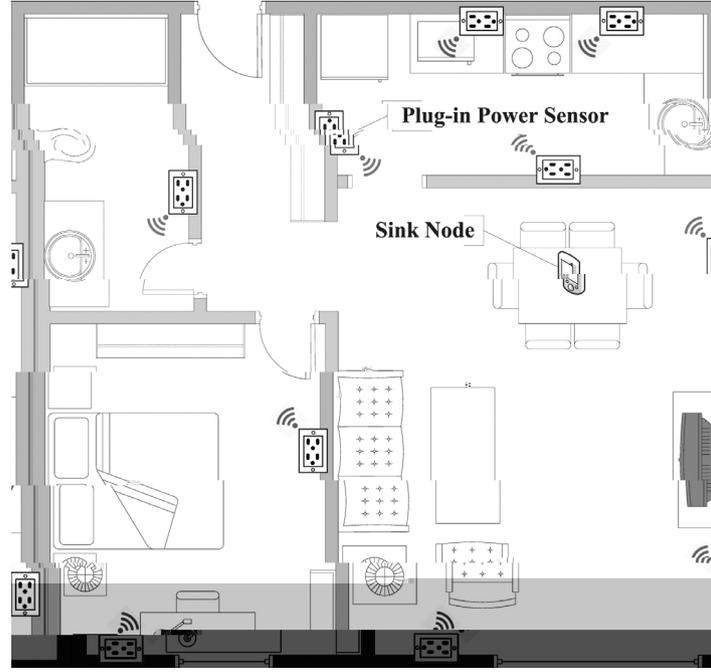


Fig. 3. Floor plan of an apartment and our smart home energy monitoring platform.

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:

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

where \mathbf{x} , $\mathbf{x}^{(i)}$, and $\hat{\mathbf{x}}^{(i)}$ 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%
-measure	80.2%	84.2%

Table 3

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

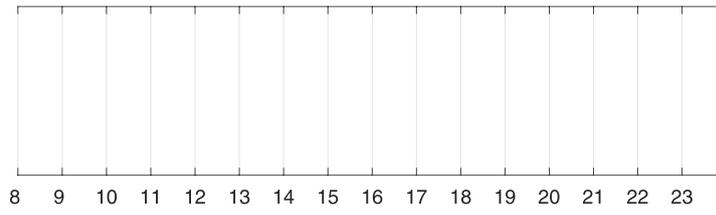
Disaggregation model	OAED (Unoccupied Per./Occupied Per./Overall)	Pure ED (Overall)
LSE model	86.33% / 62.17% / 68.20%	67.01%
HMM model	86.33% / 77.46% / 80.07%	78.87%

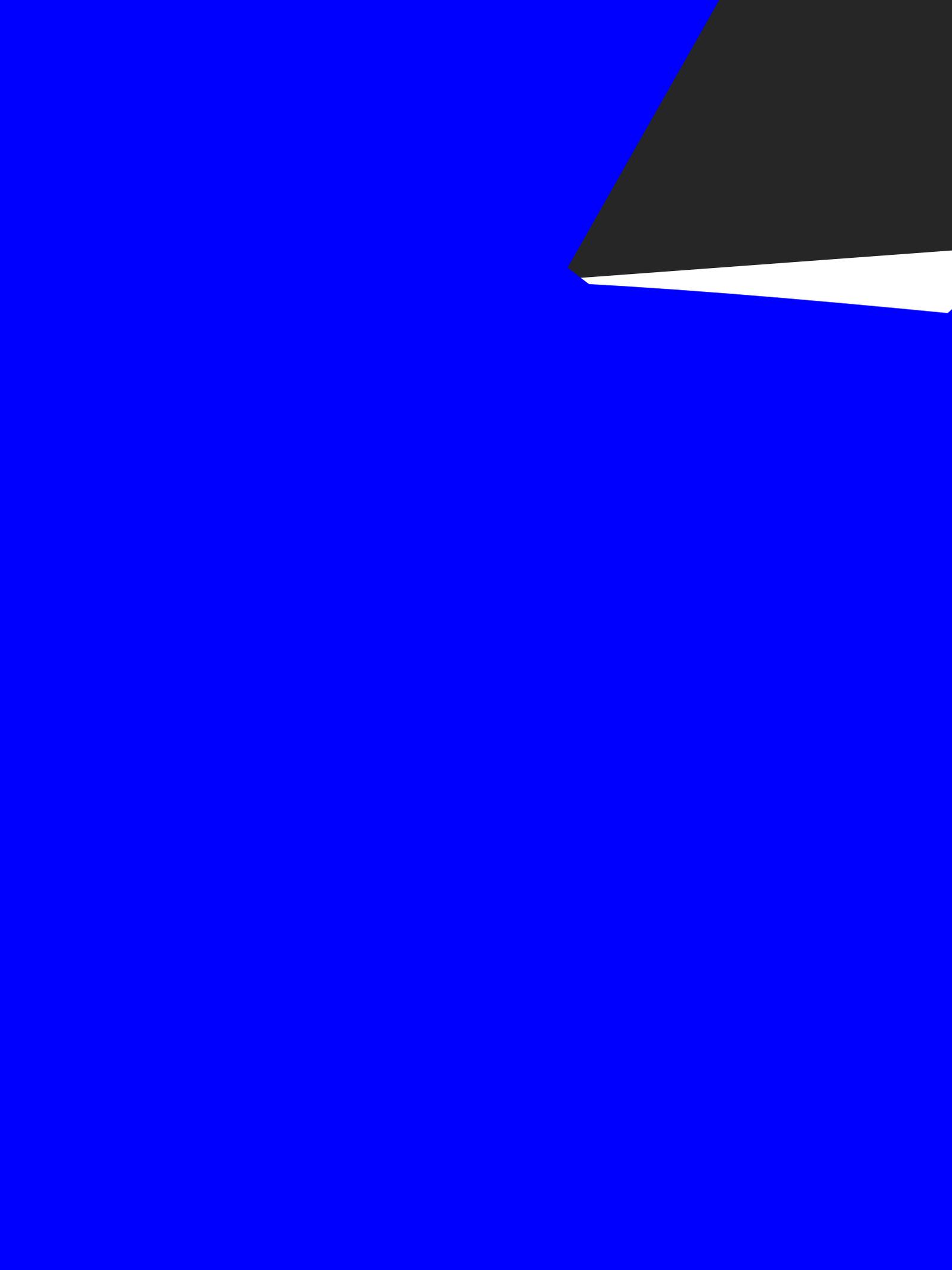
The occupancy inference accuracy for both approaches during the one-month time interval is summarized in Table 2. From the results, we can see that the occupancy inference is relatively accurate in both supervised NIOD and unsupervised NIOD, with an average *F-measure* value of 82.4%.

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.





- [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. Utama Nambi, A. Reyes Lua, V.R. Prasad, Loded: Location-aware energy disaggregation framework, in: *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments, ACM, 2015*, pp. .



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.