Occupancy-aided Energy Disaggregation

Guoming Tang and Kui Wu
Dept. of Computer Science, University of Victoria, B.C., Canada
{guoming, wkui}@uvic.ca

***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 increase the accuracy of energy disaggregation results and reduce the computational complexity, we make use of the occupancy information (whether or not the house/room is occupied by users) and split the whole time interval into occupied and unoccupied periods. In unoccupied periods, we apply simple energy approximation; in occupied periods, we perform energy disaggregation with existing methods. Real-world experiments are conducted in an apartment hosting most 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.**

I. Introduction

To effectively cut down the electricity bill for the customers as well as facilitate demand response (DR) programs for the utilities, it is meaningful to monitor the energy consumption to appliance level in residential houses [1]. 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 cheapest way to obtain appliance level energy information and has been well explored since 1980s [2].

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 [3], [4]. Generally speaking, there are two major categories of approaches to energy disaggregation: i) signature based event identification [5], [6], [7], [8] and ii) state transition based likelihood estimation [9], [10], [11]. Recently, there are also methods utilizing particular features of appliances activities, such as the state transition sparsity [12], [13].

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 first category of approaches based on appliances' signatures, it has to traverse the whole load curve to search for the appliance signature one by one [5], [6]. Methods in the second category based on state transition, such as hidden Markov model (HMM) as well as its variants, are NP-hard when they discover the most likely state sequences of appliances [9], [10]. Consequently, approximations and heuristics were developed to reduce the complexity, leading to less accurate results.

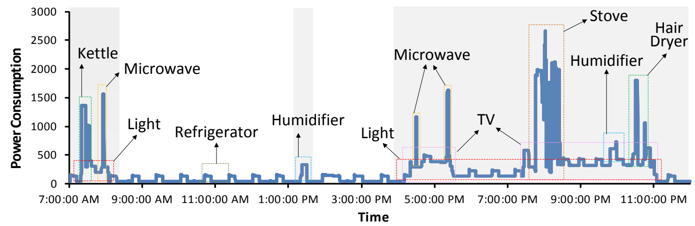

Fig. 1. Correlation between appliances' activities and the occupancy states of the house. Occupied periods of the house are illustrated by shaded areas.

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 estimate 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.

In this paper, to reduce the computational complexity of energy disaggregation while ensuring its accuracy, we develop an occupancy-aided energy disaggregation routine based on occupancy state inference for a house/room. Specifically, we first infer the occupancy states of the house/room 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 for the appliances working in occupied periods. Real-world experiments are conducted to validate the effectiveness of the occupancy-aided approach. With datasets collected from an apartment, we preliminarily demonstrate that our occupancy-aided approach can much reduce the computational overhead with ensured accuracy of energy disaggregation.

II. Occupancy-aided Energy Disaggregation

In this section, we first introduce our framework for occupancy-aided energy disaggregation, which consists of three major steps given the aggregated load curve data: occupancy inference, energy approximation, and energy disaggregation, as shown in Fig. 2.

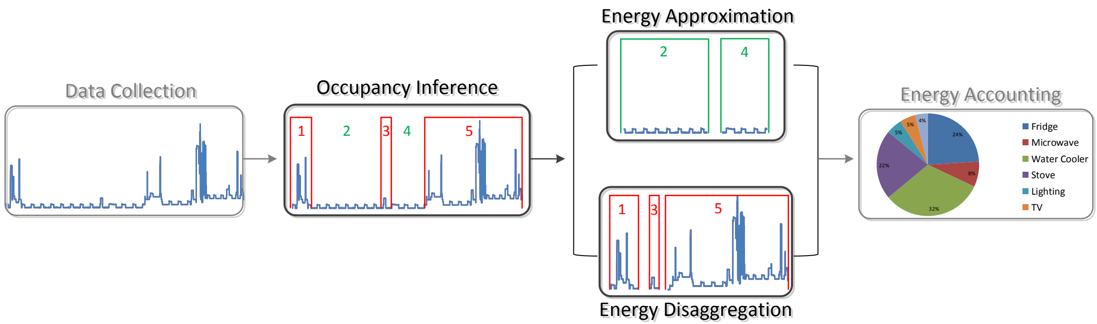

Fig. 2. Framework of occupancy-aided energy disaggregation.

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 (Sec. II-A). According to the inferred occupancy states, we estimate the energy consumption for the (quite few) appliances working during the unoccupied periods by referring to coarse-grained power information (Sec. II-B). Meanwhile, we perform energy disaggregation for the appliances running in the occupied periods (Sec. II-C). Eventually, we get the appliance-level energy consumption during the whole time interval.

A. *Occupancy Inference using Load Curve Data*

There are tremendous approaches to occupancy inference that are based on either intrusive or non-intrusive sensing. To avoid extra sensor deployment in intrusive sensing (e.g., PIR/motion sensor, acceleration sensor or camera), we apply the non-intrusive one based on load curve data analysis that was developed in [14].

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:

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

Thus, the t-th ($1 \leq t \leq n$) 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: $N_{std} := std(x)$;
- Power range: $N_{rng} := max(x) - min(x)$.

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

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

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

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

B. *Energy Approximation in Unoccupied Periods*

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.

To estimate the energy consumption of these appliances, we simply use the metric of *energy consuming rate*, such as hourly energy usage. This metric is provided by the user's manual or technique specifications, and is usually evaluated by the ENERGY STAR agency [15]. If such information is not readily-available, we can easily estimate it via simple devices such as the plug-in power meters. Without interference (appliances usage) of customers, the evaluated metric is actually quite accuracy. 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 $\mathcal{T}$ as:

$$e_k = r_k \times |\mathcal{T}|, \tag{3}$$

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

Then, with the priori knowledge of all 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.

C. *Energy Disaggregation in Occupied Periods*

To disaggregate energy for appliances in the occupied periods, we need disaggregation models/approaches. Since our focus in this paper is to investigate the contribution of occupancy information to energy disaggregation, we just adopt existing approaches instead of developing new ones.

Three different energy disaggregation approaches are implemented for our testing: the signature based approach using the Least Square Estimation (LSE) model [8], the state transition based approach applying iterative HMM model [11], and the sparse switching event recovery (SSER) model [13].

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 $s_j^{(i)}(t) = 0$.

1) LSE Model: The signature based approach based on LSE was adopted in [8] for energy disaggregation. The current waveform of each appliance was extracted and stored beforehand, and treated as its signature. In this paper, we make use of power signal instead of the current waveform. With aforementioned notations, the LSE model for energy disaggregation in the occupied periods $\mathcal{T}'$ is formulated as:

$$\begin{aligned} \min \quad & \sum_{t=1}^{|\mathcal{T}'|} \left(x_t - \sum_{i=1}^{m} \sum_{j=1}^{m_i} \mu_j^{(i)} s_j^{(i)}(t) \right)^2 \\ \textit{s.t.} \quad & s_j^{(i)}(t) \in \{0, 1\}, \\ & \sum_{j=1}^{m_i} s_j^{(i)}(t) \leq 1, \\ & 1 \leq i \leq m, 1 \leq j \leq m_i, 1 \leq t \leq |\mathcal{T}'|. \end{aligned} \tag{4}$$

2) HMM Model: As a state transition based method, the iterative hidden Markov model (HMM) was proposed for energy disaggregation in [11]. We implement this model in three phases:

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

$$\lambda := \{A, B, \pi\}, \tag{5}$$

  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 [11] was applied to infer each appliance's mode state, considering the constraints of aggregated power and power changes at each time instant.

3) SSER Model: SSER model was developed in [13] where the sparsity of appliances' switching events was applied. With the notations in preliminaries, the SSER model for energy disaggregation in the occupied periods $\mathcal{T}'$ is formulated as:

$$\begin{aligned} \min \quad & \sum_{t=1}^{|\mathcal{T}'|-1} \sum_{i=1}^{m} \sum_{j=1}^{m_i} \left| s_j^{(i)}(t+1) - s_j^{(i)}(t) \right| \\ \textit{s.t.} \quad & (\mu_j^{(i)} - \delta_j^{(i)}) s_j^{(i)}(t) \leq x_t \leq (\mu_j^{(i)} + \delta_j^{(i)}) s_j^{(i)}(t), \\ & s_j^{(i)}(t) \in \{0, 1\}, \\ & \sum_{j=1}^{m_i} s_j^{(i)}(t) \leq 1, \\ & 1 \leq i \leq m, 1 \leq j \leq m_i, 1 \leq t \leq |\mathcal{T}'|. \end{aligned} \tag{6}$$

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.

III. Implementations and Evaluations

In this section, we implement and evaluate our occupancy-aided approach for energy disaggregation using real-world datasets collected from an apartment.

A. Data Collection

We collected the power reading data from an apartment using off-the-shelf measuring devices from *Current-Cost* (www.currentcost.com). Two power sensor jaws were installed at the power entrance to measure the aggregated power consumption of the apartment. Then, the data were sent to a sink node and then forwarded to a computer with frequency of $0.1Hz$. For evaluation purpose, we also record individual power consumption of 10 major appliances across the apartment using plug-in power meters from *CurrentCost*. 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.

To find the best occupancy inference parameters, we also collected the ground-truth occupancy information as the training dataset. The Google mobile app named *Google+* (www.google.com/mobile/+/) 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-week power consumption and occupancy information were collected and used for the inference training and performance evaluation.

TABLE I
PERFORMANCE RESULTS OF OCCUPANCY-AIDED ENERGY DISAGGREGATION (E.D.) AND RAW ENERGY DISAGGREGATION

E.D. Model	**Occupancy-aided E.D.**		**Raw E.D.**	
	Accuracy (Unoccupied Per./Occupied Per./Overall)	*Overhead* (Elapsed time)	*Accuracy* (Overall)	*Overhead* (Elapsed time)
LSE Model	86.89% / 61.07% / **67.52%**	383.08 seconds	66.45%	593.71 seconds
HMM Model	86.89% / 78.19% / **80.36%**	2110.11 seconds	79.01%	3103.09 seconds
SSER Model	86.89% / 74.90% / **77.90%**	584.83 seconds	76.33%	866.30 seconds

TABLE II
PARAMETER SETTING FOR OCCUPANCY INFERENCE

parameter	*notations*	*setting value*
inference time window	τ	15 min
average power threshold	P_{avg}	125 watts
standard power deviation threshold	P_{std}	72 watts
power range threshold	P_{rng}	108 watts

B. Parameter Setting

The detailed power information (rated power and power deviation) of all appliances under consideration is measured by the plug-in power meter. To obtain accurate occupancy inference, we train and tune the inference parameters in Sec. II-A using the ground-truth occupancy information. The parameter setting in our experiments is shown in Table II. Note that in our occupancy inference, the load silence periods (e.g., when the occupants are sleeping) are treated as unoccupied periods.

C. Performance Evaluation

We perform occupancy-aided energy disaggregation following the steps in Sec. II and calculate the accuracy in occupied periods with performance metric in [16]. Moreover, we also record the accuracy of energy approximation in unoccupied periods and the overall accuracy in the whole time interval, respectively. As a comparison, we calculate energy disaggregation accuracy without following the occupancy-aided process (which we call *raw energy disaggregation*). Corresponding results are summarized in Table I.

To measure the computational complexity of the two energy disaggregation routines, we refer to running time as overhead when solving the disaggregation models. All 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 shown in Table I.

From the results, we have the following observations.

- **Comparable Accuracy:** The occupancy-aided approach is slightly more accurate than the raw one, as the energy approximation in unoccupied periods is quite accurate.
- **Much Reduced Overhead:** By cutting out the unoccupied periods from the whole time interval, the occupancy-aided approach is much faster than the raw one. Due to long unoccupied period in our case, the running time of each model is reduced by over 30%.

IV. CONCLUSIONS

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 the 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.

REFERENCES

[1] B. Neenan, J. Robinson, and R. Boisvert, "Residential electricity use feedback: A research synthesis and economic framework," *Retrieved October*, vol. 26, no. 2011, p. 3, 2009.

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

[3] M. Zeifman and K. Roth, "Nonintrusive appliance load monitoring: Review and outlook," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 1, pp. 76–84, 2011.

[4] 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.

[5] M. Dong, P. C. Meira, W. Xu, and W. Freitas, "An event window based load monitoring technique for smart meters," *IEEE Transactions on Smart Grid*, vol. 3, no. 2, pp. 787–796, 2012.

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

[7] S. R. Shaw, S. B. Leeb, L. K. Norford, and R. W. Cox, "Nonintrusive load monitoring and diagnostics in power systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 7, pp. 1445–1454, 2008.

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

[9] H. Kim, M. Marwah, M. F. Arlitt, G. Lyon, and J. Han, "Unsupervised disaggregation of low frequency power measurements," in *SIAM International Conference on Data Mining (SDM)*. SIAM, 2011, pp. 747–758.

[10] J. Z. Kolter and 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.

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

[12] S. Makonin, F. Popowich, I. V. Bajic, B. Gill, and L. Bartram, "Exploiting hmm sparsity to perform online real-time nonintrusive load monitoring," *IEEE Transactions on Smart Grid*, vol. PP, no. 99, pp. 1–11, 2015.

[13] G. Tang, K. Wu, J. Lei, and 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.
[14] D. Chen, S. Barker, A. Subbaswamy, D. Irwin, and 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.
[15] Environmental Protection Agency (EPA), "Product finder: Find and compare products," http://www.energystar.gov/productfinder/, 2016.
[16] J. Z. Kolter and M. J. Johnson, "Redd: A public data set for energy disaggregation research," in *proceedings of the SustKDD workshop on Data Mining Applications in Sustainability*, 2011, pp. 1–6.