International Conference on Applied Energy 2019
Aug 12-15, 2019, Västerås, Sweden
Paper ID: ~~xxxx~~0784

SOCKET NILM: A SOCKET-LEVEL NON-INTRUSIVE LOAD MONITORING METHOD FOR HOME APPLIANCE USAGES RECOGNITION

Zhuang ZHENG[1,2], Muhammad SHAFIQUE[1,2], Xiaowei LUO [1,2*]

1 Architecture and Civil Engineering Research Center, Shenzhen Research Institute of City University of Hong Kong, Shenzhen, China

2 Dept. of Architecture and Civil Engineering, City Univ. of Hong Kong, Hong Kong.

ABSTRACT

To support home energy management, users or operators prefer appliance-level energy consumption information than the house monthly electricity bill report. Two methods exist for appliance energy usages recognition: Non-intrusive Load Monitoring (NILM) and Intrusive Load Monitoring (ILM). Both have not been widely used due to either insufficient performance or high cost. This paper proposed a practical socket-level non-intrusive load monitoring method. First, through socket submeters, the load disaggregation accuracy can be improved by reducing occurrences of indistinguishable appliances when using simple power features; Second, by involving users' feedback, the load classification accuracy can be enhanced by feature registration and match. An unsupervised hierarchical clustering algorithm was used for load disaggregation, and the dynamic time wrapping algorithm was used for appliance feature match. This method was validated through a public dataset and showed a great promise.

Keywords: house NILM, socket NILM, load disaggregation, hierarchical clustering, load classification, feature registration

1. INTRODUCTION

Home energy management system (HEMS) is gaining more and more attention from the building sector to reduce energy consumption and carbon emissions in buildings. Appliance-level energy consumption information is of great significance for HEMS. Compared to the house monthly electricity bill, the energy consumption details of individual appliances can benefit efficient operations of both grid and building systems. For example, the improved energy saving awareness of consumers [1], the more accurate household short-term load forecast for demand response [2], the building occupancy inference through appliance energy usages [3].

There exist two kinds of load monitoring method for appliance energy usages - Intrusive Load Monitoring (ILM) and Non-Intrusive Load Monitoring (NILM). ILM mainly refers to install submeters for each appliance. Other forms of ILM were also proposed to reduce costs, like installing one submeter for a zone or a plug [4]. Katol et al. [5] installed sensors at each home outlet for appliance monitoring and control. Reinhardt et al. [6] presented a distributed load monitoring system by placing submeters in the mains connection of appliances. Instead, NILM uses a single house-level meter to disaggregate the house electricity consumption into appliance-level consumptions. Since NILM proposed by Hart in 1992 [7], researchers probed various features and algorithms. Zheng et al. [8] used ANN to build a multi-class classifier with current harmonics as features input. Wang et al. [9] used SVM to build a multi-class classifier using U-I trajectories. Researchers also tried some deep learning techniques for load disaggregation [10–13]. Kelly and colleagues applied three types of deep neural networks (DNN) for load disaggregation [10]. Lukas and Yang utilized the LSTM network to extract appliance level data in 2015 [11] and improved it by combining HMM and DNN in 2016 [12]. Bonfigli et al. treated load disaggregation as a denoising problem and proposed an encoder-decoder deep convolution network [13]. However, the above NILMs face two barriers: 1) the lack of plenty of labeled appliance-level curves or features

Selection and peer-review under responsibility of the scientific committee of the 11th Int. Conf. on Applied Energy (ICAE2019).

and 2) insufficient generalization performance of load classifiers. Currently, the submeter installations for each appliance is too costly to use and limited to research purposes.

Therefore, in this paper, the authors proposed a practical socket-level non-intrusive load monitoring method – Socket NILM – by involving socket-level submeters and users' feedbacks. Simple power features (active and reactive power, P&Q) and hierarchical clustering algorithm are used to disaggregate the power curves. Followed by feature registration and match, these disaggregated appliance power curves are further classified. This method was evaluated over the public dataset and showed an improved performance than house-level NILM. Though both socket-level NILM and P-Q features are not new, the main contribution of this paper is to combine socket-level NILM, P-Q diagram features, users' feedbacks together and formulate a practical framework to recognize appliance energy usages in homes.

2. METHODOLOGY

2.1 Architecture

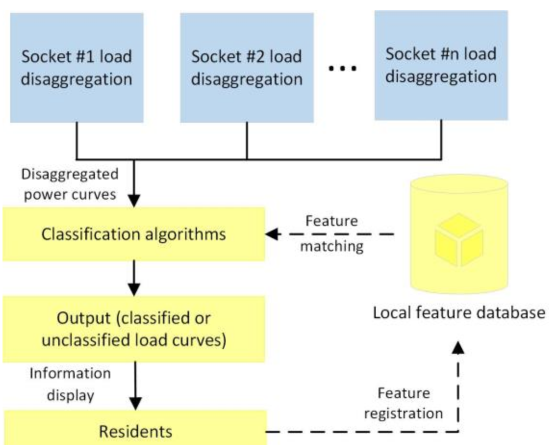

Fig 1 Socket NILM framework comprised of load disaggregation (blue part) and load classification (yellow part) (dashed lines denote the user's involvement in load classification stage)

As shown in fig 1, the Socket NILM framework consists of two tasks: load disaggregation (blue part) and load classification (yellow part). Suppose each socket is installed with a meter recording the active and reactive power values. For each socket, the load disaggregation is conducted to generate unlabeled 'appliance load curve'. Then, these collected load curves from all sockets are input into the preliminary classification algorithms. The output classified and unclassified 'appliance load curve' are sent to users. The users could give revisions or confirmations for the correct class tags of appliances based on their memory or prior knowledge. The confirmed labeled load curves are stored in a local feature database for feature registration and match. The detailed algorithms are described in the following.

2.2 Algorithm overview

2.2.1 Load disaggregation

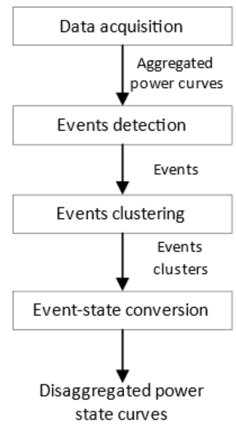

Fig 2 Flowchart of NILM load disaggregation

Fig 2 illustrates the main steps for load disaggregation in each socket. First, the aggregated power curves are metered and transferred from distributed socket-level submeters to data processing unit like PC through a communication network. Second, the event detector detects the network state changes by making 1^{st} order difference between any two neighboring values by equation (1). A power threshold value of 30w is selected empirically to filter ineffective events and reduce the false event detections.

$$\begin{aligned}\Delta P_{i+1} &= P_{i+1} - P_i \\ \Delta Q_{i+1} &= Q_{i+1} - Q_i\end{aligned} \quad (1)$$

where ΔP_{i+1}, ΔQ_{i+1} is the difference of active, reactive power between two adjacent recorded values. If $|\Delta P_{i+1}| > 30W$, then one event is detected and the active and reactive power changes $(\Delta P_{i+1}, \Delta Q_{i+1})$ are stored as an event feature for future use. The event detector here uses a constant absolute threshold and an adaptive threshold may provide better results. However, the threshold methods always face problems to discriminate the true events from normal load variations. The authors intend to detect events by detecting adjacent states first through clustering techniques in future.

Then, unsupervised hierarchical clustering algorithm clusters these event-based features (ΔP, ΔQ) for disaggregation. Positive ΔP events correspond to power boost events like *turn-on* TV, while negative ΔP events correspond to power buck events like *turn-off* TV. To avoid a repeat, positive ΔP-ΔQ features of power boost events are clustered. Ideally, each event cluster corresponds to a set of power boost events for a type of appliance.

Algorithm: event-state conversion by searching and matching

```
Input: one boost event cluster C with timestamp index list Tc, all buck events set N with
timestamp index list Tn, search window length T
Output: The matched boost&buck event timestamp list Tm for cluster C, the estimated
power states during these periods Pm.
    Tm, Pm are empty lists, search_pointer:=0
    For p in C:
        n = N[search_pointer]
        For n in N:
            win = Tn[n] - Tc[p]
            If win >= T then
            pass
            print ("exceeding search window length")
              elif: abs(C[p]+N[n]) < C[p]/3:
            Tm.append([Tc[p], Tn[n]])
Pm.append([abs(C[p])+abs(N[n])]/2)
                        search_pointer = search_pointer + 1
            break
              else:
            pass
    return Tm, Pm
```

Fig 3 Pseudocode of event-state conversion algorithm

Besides, in practice, consumers care about how much power consumed by what appliance at what time slot. Therefore, the event-state conversion algorithm in fig 3 converts appliance usage time sequence to appliance power state sequence. Given the disaggregated boost event clusters and all negative buck events, the algorithm pairs boost-buck events with the interval state estimated by the average magnitude of the paired events. A search window length *T=3600s* is set beforehand, considering the duration of the home appliance is usually less than one hour. Finally, each boost event cluster is paired with buck events and converted into a load power curve. Note that though some appliances operate constantly during the day like refrigerator and water heater, they can still be disaggregated through event-based NILM and get paired within the search window length. As these appliances stay idle with small latent current during the most time of day and their states change either periodically (refrigerator) or change when the associated temperature is crossing the comfort band (heater). The time interval between consistent state changes are within the search window length.

2.2.2 Load classification

These disaggregated load curves are further classified in this section. Load classification includes two steps: Preliminary Global Classification and Feature Registration & Match. First, the preliminary global classifier gives us an initial coarse guess of appliance types by utilizing prior knowledge. For example, according to the appliance usage time, usage frequency, and power level, the preliminary classifiers may identify what kind of appliance it may be or not be. Second, to enhance the classification accuracy of the preliminary global classifier, the feedbacks by users are involved by three phases: information display, feature registration, and feature match. These preliminarily classified or unclassified curves are displayed to users through the web or mobile devices. Then users could give revisions or confirmations to these load curves based on their memory. After receiving the feedbacks, the corresponding appliance features are registered to a local appliance feature database. When next round comes, the feature database is searched and matched with the disaggregated power curves through Dynamic Time Warping (DTW) algorithms. DTW algorithms could find the global alignment between two time series with different lengths and temporal distortions. In deed, the users' feedbacks may be wrong and misleading. This could be remedied by warning the users through preliminary global classification and by accepting the later revisions made by users.

3. EXPERIMENT EVALUATION

3.1 Experiment setup

The authors used one public dataset for evaluation– Almanac of Minutely Power dataset (AMPds) [14]. As the lack of socket-level meters, the authors grouped appliances with the possibility to work at the same socket (e.g., the clothes washer and clothes dryer, the oven and fridge) into "virtual" socket submeters. The "virtual" house-level meter was also constructed by adding all these socket submeters' values together. To illustrate the advantages of the proposed framework, the authors compared two sceneries regarding disaggregation accuracy: House NILM and Socket NILM. Table 1 shows the virtual meter configurations of two sceneries.

Table 1. Virtual meter configurations for two sceneries: Socket NILM and House NILM

Scenery	Virtual meter	App. Abbreviation	App.name
Socket NILM	Socket 1	CDE	Clothes Dryer
		CWE	Clothes Washer
	Socket 2	DWE	Dishwasher
		FGE	Kitchen Fridge
	Socket 3	EQE	Security/Network
		FRE	HVAC/Furnace
		HPE	Heat Pump

		WOE	Wall Oven
	Socket 4	HTE	Instant Hot Water Unit
		TVE	entertainment equipment
House NILM	global meter	CDE, CWE, DWE, FGE, EQE, FRE, HPE, WOE, THE, TVE	——

The first month's data was disaggregated following the steps in section 2.2.1. These disaggregated load curves were manually labeled with actual type names of the appliance to simulate the process of users' correct feedbacks. Then, these labeled curves were saved to the PC database for feature registration. Second, the second month's data was also disaggregated. The DTW distance between each disaggregated curve and the registered curves in the database were calculated. The pair of curves with minimum DTW distance was considered to match.

3.2 Metrics

A well-known metric measures the load disaggregation accuracy – the *proportion of total energy correctly assigned (TECA)* (2). A larger *TECA* value represents a better disaggregation performance.

$$\text{TECA} = 1 - \frac{\sum_{t=1}^{T}\sum_{i=1}^{K}\left|\hat{y}_t^{(i)} - y_t^{(i)}\right|}{2\sum_{t=1}^{T}\bar{y}_t} \qquad (2)$$

where $\hat{y}_t^{(i)}$, $y_t^{(i)}$ is the disaggregated signal and original appliance signal for appliance i, respectively; K denotes the number of disaggregated load curves, T denotes the total samples in time axis; $\bar{y}_t$ denotes the observed aggregated signal.

4. RESULTS AND DISCUSSION

4.1 Load disaggregation

Table 2 gives the load disaggregation results for House NILM and Socket NILM, respectively. It is shown that only three loads were correctly disaggregated from the aggregated power curves by House NILM while eight loads were disaggregated by Socket NILM.

Table 2. Number of disaggregated loads for House NILM and Socket NILM

Scenery	Number of disaggregated loads				
	Socket 1	Socket 2	Socket 3	Socket 4	Global
Socket NILM	2	2	2	2	8
House NILM	—	—	—	—	3

Table 3 illustrates the *minutely*, *hourly*, *daily* and *weekly* proportion of total energy correctly assigned (TECA) for two sceneries. It is observed that Socket NILM could always achieve nearly 5% TECA improvement than House NILM. Another interesting finding is that the more coarse-grained data gave slightly higher TECA (minutely< hourly< daily< weekly).

Table 3. Evaluation of load disaggregation under two sceneries: House NILM and Socket NILM

Sceneries	TECA (minutely)	TECA (hourly)	TECA (daily)	TECA (weekly)
House NILM	73.33%	75.22%	75.28%	75.42%
Socket NILM	77.42%	79.86%	81.14%	81.58%

4.2 Load classification

Table 4 illustrates the appliance feature matching results. The first row denotes the actual name of registered curves after disaggregating and labeling the first month's data. The first column denotes the unclassified disaggregated curves of second month's data (the *Abbreviations* in brackets mean the actual appliance types). According to the rule - *the matching with minimum DTW distance is correctly matched*, six curves were correctly labeled out of seven disaggregated curves. Only the *Socket 1 cluster 2 (CWE)* was misclassified into HTE with slightly smaller DTW distance. It validated the promise of the DTW algorithm for feature match. In future, the authors intend to apply DTW algorithm for transient feature matching.

Table 4. DTW distance matrix between disaggregated curves and registered appliance curves

disaggregated appliance curves	Registrated appliance curves						
	CDE	CWE	FGE	DWE	HPE	HTE	TVE
(CDE) Socket 1 cluster 1	15461	167400	180676	203918	365227	177704	167937
(CWE) Socket 1 cluster 2	357531	8302	14665	44794	190622	5817	10282
(FGE) Socket 2 cluster 1	4612461	11257	9525	120163	1539882	48458	11169
(DWE) Socket 2 cluster 2	360552	29199	385623	18903	114513	139836	102323
(HPE) Socket 3 cluster 3	286150	134182	1154413	78600	7661	368637	310083
(HTE) Socket 4 cluster 1	953724	7435	45191	84995	506485	3792	11358
(TVE) Socket 4 cluster 2	878485	9144	9855	72210	457833	9740	8861

5. CONCLUSIONS

This paper proposed a socket-level non-intrusive load monitoring method for home appliance usages recognition with practical concerns. Compared to house-level NILM, the low-frequency features' divisibility is improved by hardware (socket-level submeters) for Socket NILM. Compared to appliance-level metering, the developed method could be a scalable solution with relatively low cost, especially for buildings or sites with frequent changes of connected devices (e.g. office buildings). Two main conclusions are obtained: 1) the disaggregation performance can be improved through hardware pre-disaggregation of socket-level meters; 2) instead of building a powerful global classifier, one light preliminary global classifier enhanced by users' feedbacks are effective for local load classification.

ACKNOWLEDGEMENT

This work was supported by the Shenzhen Science and Technology Funding Programs (JCYJ20150902162946055). The conclusions herein are those of the authors and do not necessarily reflect the views of the sponsoring agency.

REFERENCE

[1] Carrie ArmelK, GuptaA, ShrimaliG, AlbertA. Is disaggregation the holy grail of energy efficiency? The case of electricity. Energy Policy 2013;52:213–34. doi:10.1016/j.enpol.2012.08.062.

[2] ZhengZ, ChenH, LuoX. A Kalman filter-based bottom-up approach for household short-term load forecast. Appl Energy 2019;250:882–94. doi:10.1016/j.apenergy.2019.05.102.

[3] HongT, LinH. Occupant Behavior : Impact on Energy Use of Private Offices. Asim IBSPA Asia Conf 2012.

[4] RidiA, GislerC, HennebertJ. A survey on intrusive load monitoring for appliance recognition. Proc - Int Conf Pattern Recognit 2014:3702–7. doi:10.1109/ICPR.2014.636.

[5] KatoT, ChoHS, LeeD, ToyomuraT. Appliance Recognition from Electric Current Signals for Information-Energy Integrated Network in Home Environments n.d.

[6] ReinhardtA, BurkhardtD, ZaheerM, SteinmetzR. Electric appliance classification based on distributed high resolution current sensing. Proc - Conf Local Comput Networks, LCN 2012:999–1005. doi:10.1109/LCNW.2012.6424093.

[7] HartGW. Nonintrusive Appliance Load Monitoring. Proc IEEE 1992;80:1870–91. doi:10.1109/5.192069.

[8] ZhengZ, ChenH, LuoX. A Supervised Event-Based Non-Intrusive Load Monitoring for Non-Linear Appliances. 2018.

[9] WangAL, ChenBX, WangCG, HuaD. Non-intrusive load monitoring algorithm based on features of V–I trajectory. Electr Power Syst Res 2018;157:134–44. doi:10.1016/j.epsr.2017.12.012.

[10] KellyJ, KnottenbeltW. Neural NILM: Deep Neural Networks Applied to Energy Disaggregation. Proc 2nd ACM Int Conf Embed Syst Energy-Efficient Built Environ Korea,4–5 Novemb 2015 n.d.:55–64.

[11] MauchL, YangB. A new approach for supervised power disaggregation by using a deep recurrent LSTM network. 2015 IEEE Glob Conf Signal Inf Process Glob 2015 2016:63–7. doi:10.1109/GlobalSIP.2015.7418157.

[12] MauchL, YangB. A novel DNN-HMM-based approach for extracting single loads from aggregate power signals. ICASSP, IEEE Int Conf Acoust Speech Signal Process - Proc 2016;2016–May:2384–8. doi:10.1109/ICASSP.2016.7472104.

[13] BonfigliR, FelicettiA, PrincipiE, FagianiM, SquartiniS, PiazzaF. Denoising autoencoders for Non-Intrusive Load Monitoring: Improvements and comparative evaluation. Energy Build 2018;158:1461–74. doi:10.1016/j.enbuild.2017.11.054.

[14] MakoninS, PopowichF, BartramL, GillB, BajićIV. AMPds: A public dataset for load disaggregation and eco-feedback research. 2013 IEEE Electr Power Energy Conf EPEC 2013 2013. doi:10.1109/EPEC.2013.6802949.