Electricity Usage Profile Disaggregation of Hourly Smart Meter Data

Bochao Zhao, Lina Stankovic, and Vladimir Stankovic
Department of Electronic and Electrical Engineering
University of Strathclyde, Glasgow, G1 1XW, UK
Email: {b.zhao, lina.stankovic, vladimir.stankovic}@strath.ac.uk

***Abstract*—This paper is motivated by the growing demand of disaggregating electricity consumption measured by smart meters, down to appliance level. The very low 15-min to 60-min granularity of energy measurements available for analysis, as is standard by the majority of nationwide smart metering programmes, is posing serious challenges. The non-intrusive load monitoring (NILM) solutions for these very low data rates cannot leverage on low (1-60sec) to high rates (in the order of kHz to MHz) NILM approaches, and so far have not received much attention in the literature. In this paper, we propose a novel electricity profile hourly disaggregation of energy consumed (kWh) based on K-nearest neighbours (K-NN), that relies on features such as statistical measures of the energy signal, time usage profile of appliances and reactive power consumption (if available). We propose relative standard deviation as a metric to assess the quality of each feature per appliance. For validation, three publicly accessible real-world datasets are used, namely the REDD, REFIT and AMPds (Version 2), for up to 3 months.**

I. Introduction

Non-intrusive load monitoring (NILM) is defined as estimating individual appliance energy usage from smart meter readings (e.g., voltage, current, power) of a whole household using purely software tools [1]. Though various machine learning methods including Hidden Markov Model (HMM) and its variants, Dynamic Time Warping (DTW) [2], K-nearest neighbours (K-NN) [3]-[5], sparse coding [6], Neural Network [7] and Graph Signal Processing [8] have been applied to the NILM problem, most approaches focus on either high sampling rates in the order of kHz or MHz or sampling rates between 1sec and 1min. Disaggregation in these cases is usually performed via feature extraction and state transition modelling on active and/or reactive power data. Very low-rate (10-60 mins) NILM is slowly gaining interest [3]-[6], [9], [10] since electricity meters deployed at scale in most countries tend to provide extremely low-rate measurements, at 15, 30 minutes or hourly granularity. Compared with power measurements of higher granularity, the energy consumption signal at very low granularity features limited state transitions, fewer low-consuming appliances' feature patterns and much higher probability of multiple appliances running simultaneously. Thus, lack of well-known features and increased appliance noise make very low NILM a challenging problem [4], which can be looked at as electricity usage profile disaggregation, since the input is the total energy use within fixed time intervals (e.g., in kiloWatt-hour), instead of active/reactive power readings collected at relatively high frequency in Watts or VArs.

Sparse coding is proposed for hourly measurements in [6], where appliance models built from weekly typical training sets are used to predict appliance-level power consumption for unseen houses. Plug-level disaggregation results for hourly power consumption data is shown in [6] but the aggregate power consumption used is artificially obtained as the sum of sub-metered components disregarding measurement noise and noise due to unknown appliances, typically present in aggregate smart meter data. In [9], piecewise functions of power consumption versus external temperature per house are modelled for hourly disaggregation of base-load, heating and cooling consumption but not specific to any appliances. Similarly, appliance-level disaggregation results are not demonstrated in [10], where contextual supervision is applied to the single-channel source separation problem for overall heating, cooling and base load.

K-NN, as a low-complexity time-series classification/clustering approach, has been attempted for electricity usage profile classification in [3], [5] and [4]. In [3], K-NN is applied as a tool to find houses most similar to the test house and then estimate appliance-level monthly energy consumption, using as features house size, occupancy and room number but not weather; monthly consumption is then estimated based on a very large training dataset. Our proposed method is not limited by training on a large dataset. In [5], Principal component analysis (PCA) is used for determining suitable houses for transfer learning, and K-NN is used as a time-series classifier for electricity usage profile classification with granularity of 10 min, but without exploiting time information as feature. The features utilised in [4] are derived from both magnitude and time, as in our proposed algorithm. However, unlike our proposed algorithm, daily total consumption for each appliance is not considered in [4]. We also add an additional feature for representing correlation between active and reactive power. Moreover, in our proposed algorithm, feature selection for each appliance is determined by our proposed metric for feature quality evaluation. Classification results in terms of classification accuracy are presented in [4] but not disaggregation results, e.g., estimated load-specific consumption.

In this paper, we propose a supervised K-NN based electricity usage profile disaggregation of energy measurements at 15 and 60 min granularity to identify a range of appliances. Relative standard deviation is proposed as a metric to determine which features are most useful for disaggregating particular appliances. Unlike [3], [4] and [5], we validate the disaggregation results using three open access datasets of true power measurements: REDD [11] (US houses), REFIT [12] (UK houses) and AMPds [13] (a Canadian house). For all datasets, we calculate the electricity energy profile of the aggregate load for 15-min and 1 hour in Watt-hour (Wh).

II. Electricity usage profile disaggregation: Notation and problem statement

We denote by $\mathbf{P}$ and $\mathbf{Q}$ active and reactive power signal, respectively. Let P_t be the active power collected at time instant t and W_{P_i} be the total household's electric energy consumption measurement within the i-th time interval. Namely,

$$W_{P_i} = \int_{T_i} P_t dt, \tag{1}$$

where T_i is the duration of the time interval i. Then the electricity usage profile disaggregation task is, $\forall i$, to estimate energy contributed by each individual appliance m, $W_{P{m_i}}$, towards the total energy W_{P_i}. That is,

$$W_{P_i} = \sum_{m \in \mathbf{M}} W_{P{m_i}} + n_i, \tag{2}$$

where $\mathbf{M}$ is set of appliances or loads contributing towards the aggregate load, and n_i refers to noise due to measurement error and unknown appliances. W_{Q_i} is calculated as in Eq. (1) by replacing P_t with Q_t, only as a feature category but contains no practical meaning.

III. Methodology

In this section we describe our proposed algorithm. We start with a brief background of K-NN and then move on to defining feature categories used in the proposed algorithm.

K-NN is a time-series classification method where test samples are classified by a majority vote of neighbours via distance calculation between samples' attributes and corresponding features of instances in the training database using a distance metric [5]. Popular distance metrics include Euclidean, Manhattan, Hamming, DTW, etc. Let y be a test sample. Then, we calculate distances between y and all samples in the training dataset $x_1, \ldots, x_K$, and find the minimum distance:

$$d_y = min\left\{d(y,x_1); d(y,x_2); ...; d(y,x_K)\right\}, \tag{3}$$

where $d(\cdot,\cdot)$ is a distance measure.

A. Feature extraction

We assume, as in [3], that Individual appliance monitoring (IAM) measurements of individual loads are available for training. Note that if sub-metering data is not available, as it is the case for smart metering nationwide rollouts, we can use time-wise features extracted from a time-of-use diary to estimate magnitude-wise features from the aggregate load like in [4] and [5].

In many countries, including Spain and Italy, smart meter measurements are collected every 24 hours. Therefore, we perform daily disaggregation on 24-hour long windows, where for hourly readings, each window contains $n = 24$ energy samples. Within each time sample, an appliance can be in either OFF state, if it was not running at all during that hour, or in the ON state, otherwise. Thus, for a duration of one window, there are 2^n possible daily combinations of ON and OFF states. To reduce complexity, we limit the number of candidate ON-OFF state patterns by filtering out invalid combinations based on appliance time usage profile, e.g., refrigerators are always ON, which is not the case with electric heater.

Features are extracted as $apriori$ inputs to K-NN, as in previous works [3], [4] and [5], depending on the datasets and algorithms used. From $\mathbf{W_P}$, using the training dataset, we intuitively design the categories of features as follows: 1. Average daily ON duration; 2. Maximum daily ON duration; 3. Minimum daily ON duration; 4. Average daily switched-ON time; 5. Average daily switched-OFF time; 6. Median time of day for daily running; 7. Average consumed energy per day; 8. Maximum consumed energy per day; 9. Minimum consumed energy per day; 10. Variance of consumed energy per day; 11. Average daily total energy consumed. When both energy and reactive power consumption measurements are available, for each appliance, an additional feature can be used $-$ the average ratio between active and reactive measurements used in the ON state. Since only a subset of features are useful for disaggregating individual loads, an adaptive feature refining step is proposed based on the assumption that the subset of useful features should be extracted from attributes with high precision and low variability. We use relative standard deviation (RSD), as in [8], to represent the quality of each feature, where features are selected based on constant threshold G for evaluating RSD values. Note that additional features extracted from other attributes, such as weather and occupancy information used in [3], are available in our proposed algorithm, only if they result in small RSD.

Table I lists the selected features for several appliances from the AMPds dataset using $G = 0.5$. The abbreviations used for domestic loads considered in this paper are as follows: HWU is hot water unit; CW is clothes washer; DW is dishwasher; SNE is security/network equipment; HP is heat pump; UT is utility room; EWB is electronics workbench; GR is garage; FZ is freezer; KO is kitchen outlet; F is fridge and EH is electrical heater. From Table I, the majority of features extracted from $\mathbf{W_P}$ for most appliances are of *high quality*, e.g., they have low RSD. On the other hand, the features extracted from $\mathbf{W_Q}$ for HWU and UT have low precision and high RSD, so are not used. Apart from daily total consumption, no feature can be extracted from $\mathbf{W_P}$ or $\mathbf{W_Q}$ for SNE due to its low-consumption, which is attributed to the base-load.

B. Feature matching

Fig. 1 illustrates how d_y in Eq.(3) is calculated for the proposed K-NN based electricity profile disaggregation algorithm. During training, we obtained a set of aforementioned daily

TABLE I

FEATURES CONSIDERED FOR APPLIANCES FROM THE AMPDS DATASET.

	Features	HWU	CW	DW	SNE	HVAC	HP	UT	EWB	GR
$\mathbf{W_P}$	Average daily ON duration	√	√	√		√		√	√	√
	Maximum daily ON duration	√	√	√		√		√	√	√
	Minimum daily ON duration	√	√	√		√		√	√	√
	Average daily switched-ON time	√						√		
	Average daily switched-OFF time	√				√	√	√		
	Median time of day for daily running					√	√	√		
	Average consumed energy per day	√	√	√		√	√	√	√	√
	Maximum consumed energy per day	√	√	√			√	√	√	√
	Minimum consumed energy per day	√	√	√		√	√	√	√	√
	Variance of consumed energy per day	√		√			√	√	√	√
	Average daily total consumption	√	√	√	√	√	√	√	√	√
$\mathbf{W_Q}$	Average daily ON duration			√					√	√
	Maximum daily ON duration			√					√	√
	Minimum daily ON duration			√					√	√
	Average daily switched-ON time									√
	Average daily switched-OFF time					√	√			√
	Median time of day for daily running					√	√			√
	Average consumed energy per day		√	√		√	√		√	√
	Maximum consumed energy per day		√	√		√	√			√
	Minimum consumed energy per day		√	√		√	√		√	√
	Variance of consumed energy per day			√			√			
	Average daily total consumption		√	√	√	√	√		√	√
$\mathbf{W_P}$ & $\mathbf{W_Q}$	Average ON state magnitude ratio between W_{Q_i} and W_{P_i}	√	√	√		√	√	√	√	√

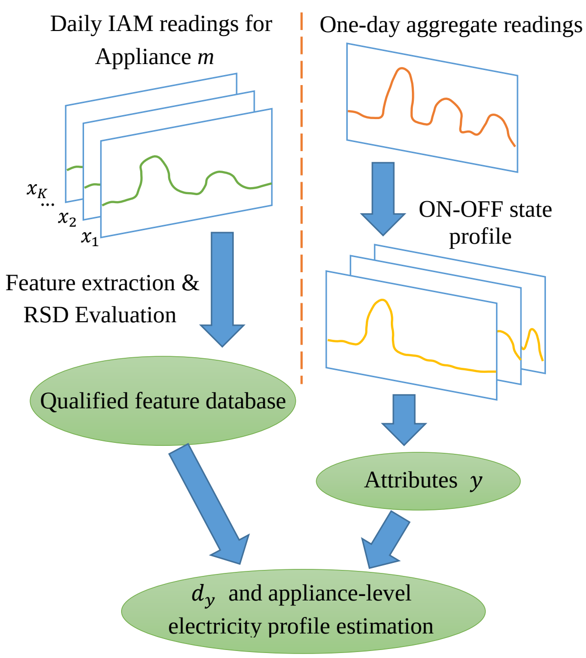

Fig. 1. Flow chart of the proposed algorithm.

features extracted and shortlisted (after RSD evaluation) for Appliance m, as discussed in Subsection III-A - left of Fig. 1. During testing (right of Fig. 1), we extract the same features as those in the qualified feature database. Then K nearest neighbours are defined as K daily readings, $x_1, \ldots, x_K$, from the training dataset, whose features have the shortest distances (calculated using Eq.(3)) to the testing candidate ON-OFF state pattern y for Appliance m. For daily disaggregation and hourly sampling rates, y and x_i's are all $n = 24$-length vectors. The resulting minimum distance d_y among all possible candidates y and selected neighbours x classifies appliances as per the training set.

IV. EXPERIMENTAL VALIDATION AND DISCUSSION

We demonstrate the performance of the proposed approach using (1) **P** of the US REDD dataset, (2) **P** of the UK REFIT dataset (Cleaned) [12] and (3) both **P** and **Q** of Canadian AMPds dataset (Version 2) [13]. The three datasets can be accessed via http://redd.csail.mit.edu/, DOI:10.15129/9ab14b0e-19ac-4279-938f-27f643078cec and DOI:10.7910/DVN/FIE0S4, respectively. REDD dataset offers the best sub-metering coverage but provides the fewest measurements. REFIT houses include fewer submetered loads than AMPds and aggregate load measurements are considerably noisier than in the other two datasets, due to unknown appliances. Since the datasets' original measurements' sampling rates are 1sec, 8sec, and 1min, we calculate the 15-min or hourly energy $\mathbf{W_P}$ via Eq.(1). As time of day is a key feature in our algorithm, only whole-day available measurements are used, e.g., if there are measurement samples missing within a 24-hour window of a day, that day is not included in our experiments. Based on the aforementioned data selection rules and individual appliance usage frequency, the following portions of data are empirically selected: 19/04/2013–01/05/2013 (13 days in total, 7 days for training and 6 days for testing) for REDD House 2; 28/10/2013–02/06/2014 (90 days in total, 30 days for training and 60 days for testing) for REFIT House 1; 28/10/2012–27/04/2013 (180 days in total, 90 days for training and 90 days for testing) for the AMPds house. Fridge and freezer measurements in REFIT House 1 are merged. Aggregate measurements from the AMPds dataset includes main house, garage and rental suite. As is common practice by utilities for summarising energy use, we use kWh to present our results. In all experiments, G for evaluating RSD is empirically set to 0.5.

Tables II and III show the disaggregation results for 15-min and 1-hour electricity usage profile, respectively, for REDD

TABLE II
PERFORMANCE OF THE PROPOSED METHOD FOR 15-MINUTE DATA.

	REDD House 2			REFIT House 1		
Appliance	FZ	KO	BL	F+FZ	EH	BL
Est. (kWh)	11.6	1.47	3.08	61.67	209.16	151.98
IAM (kWh)	12.32	1.47	2.88	63.08	253.71	101.96
Est./Total	36.68%	4.65%	9.74%	7.18%	24.34%	17.69%

House 2 Freezer (FZ), Kitchen Outlet (KO) and Base-load (BL), and REFIT House 1 Fridge&Freezers (F+FZ), Electrical Heater (EH) and BL. IAM row, used as ground truth, shows the actual submetered energy in kWh. Est./Total shows the percentage contribution of the estimated energy consumed by an individual load towards the aggregate (measured) load.

We can disaggregate 51% and 62.5%, for 15-min and 1-hour granularity, respectively, of REDD House 2 total load and about 49% and 54%, for 15-min and 1-hour granularity, respectively, of REFIT House 1 total load. The amount of energy consumed that can be accounted for due to individual loads is slightly lower for hourly granularity measurements. As the Est. and IAM rows for each appliance show, our disaggregated energy for each of the selected appliances is close to the actual energy consumed. The disaggregated energy for KO and BL for REDD House 2 is overestimated, which is inline with very low rate disaggregation results based on sparse coding [6] where performance of Fridge/Freezer is generally good but the overestimation problem generally exists for short-duration and low-energy loads.

TABLE III
PERFORMANCE OF THE PROPOSED METHOD FOR HOURLY DATA.

	REDD House 2			REFIT House 1		
Appliance	FZ	KO	BL	FZ	EH	BL
Est. (kWh)	9.52	1.6	8.67	43.53	235.62	197.42
IAM (kWh)	12.32	1.47	7.11	45.49	255.97	205.56
Est./Total	30.1%	5.07%	27.41%	4.96%	26.85%	22.5%

TABLE IV
PERFORMANCE OF THE PROPOSED METHOD FOR THE AMPDS DATASET FOR HOURLY DATA.

		HWU	CW	DW	SNE	HVAC
Est. (kWh)	$\mathbf{W_P}$	11.61	11.62	45.77	78.46	224.2
	$+\mathbf{W_Q}$	11.71	14.02	45.24	78.57	222.89
IAM (kWh)		16.85	10.23	36.78	86.97	252.11
Est./Total	$\mathbf{W_P}$	0.41%	0.41%	1.63%	2.79%	7.98%
	$+\mathbf{W_Q}$	0.42%	0.50%	1.61%	2.80%	7.94%

		HP	UT	EWB	GR	BL
Est. (kWh)	$\mathbf{W_P}$	456.77	108.95	57.34	3.29	212.74
	$+\mathbf{W_Q}$	477.74	108.95	57.34	3.29	212.74
IAM (kWh)		552.44	111.04	56.6	3.54	179.36
Est./Total	$\mathbf{W_P}$	16.26%	3.88%	2.04%	0.12%	7.58%
	$+\mathbf{W_Q}$	17.01%	3.88%	2.04%	0.12%	7.58%

Table IV shows the disaggregation results for hourly electricity usage profile for the AMPds house, where $+\mathbf{W_Q}$ refers to disaggregation with $\mathbf{W_Q}$ available as a feature. When considering both $\mathbf{W_P}$ and $\mathbf{W_Q}$ instead of $\mathbf{W_P}$ only, the algorithm performs worse for CW and better for HP. The features derived from $\mathbf{W_Q}$ of CW have low RSD resulting in overestimation. The proposed approach can disaggregate 43% of AMPds house total electricity consumption given $\mathbf{W_P}$ and 44% given both $\mathbf{W_P}$ and $\mathbf{W_Q}$. Overall, with hourly NILM, the inclusion of $\mathbf{W_Q}$ as a feature does not seem to improve results significantly.

V. CONCLUSION

In this paper, we propose a supervised K-NN based electricity usage profile disaggregation solution for daily appliance-level energy feedback. Unlike K-NN classifiers of [4], [5], appliance time usage profile is considered in our method to extract useful features. Furthermore, RSD is used to evaluate the quality of each feature and customise feature selection per appliance. After validation on three datasets for up to 3 months, we show that the proposed algorithm successfully disaggregates appliance energy consumption when compared to the individual, appliance-specific, energy measurements and can disaggregate up to 62% of the daily energy consumption from the total noisy electricity usage profile with 15-min and 60-min granularity. Future work includes weighting features based on RSD or other metrics; validation enhancement by adding state-of-the-art benchmarks, including sparse coding [6]; improvement of ON-OFF states prediction rules to trade-off efficiency; widening the set of loads that can be estimated reliably; transfer learning from similar houses and appliance ownership and usage profile.

VI. ACKNOWLEDGEMENT

The work was supported in part by the European Union Horizon 2020 research and innovation programme, Eco-Bot project under grant agreement No 767625.

REFERENCES

[1] G. W. Hart, "Nonintrusive Appliance Load Data Acquisition Method", *MIT Energy Laboratory Technical Report*, Sep. 1984.

[2] R. Bonfigli, S. Squartini, M. Fagiani, and F. Piazza, "Unsupervised algorithms for non-intrusive load monitoring: An up-to-date overview," in *IEEE Int. Conf. Envir. and El. Engineering,* pp. 1175–1180, 2015.

[3] N. Batra, A. Singh, and K. Whitehouse, "Gemello: Creating a detailed energy breakdown from just the monthly electricity bill," in *Proc. ACM SIGKDD,* pp. 431–440, Aug. 2016.

[4] J. Lines,A. Bagnall, P. C. Smith, and S. Anderson. "Classification of household devices by electricity usage profiles," in *Proc. IDEAL 2011,* pp. 403–412, Sep. 2011.

[5] K. Basu, V. Debusschere, A. Douzal-Chouakria, and S. Bacha, "Time series distance-based methods for non-intrusive load monitoring in residential buildings," *Energy and Buildings,* vol. 96, pp. 109–117, Mar. 2015.

[6] J. Kolter, S. Batra, and A.Y. Ng, "Energy disaggregation via discriminative sparse coding," *in Proc. NIPS*, Vancouver, Canada, Dec. 2010.

[7] J. Kelly and W. Knottenbelt, "Neural NILM: Deep neural networks applied to energy disaggregation," in *Proc. the 2nd ACM Int. Conf. Embedded Sys. for Energy-Eff. Built Env.*, pp. 55–64, Nov. 2015.

[8] B. Zhao, L. Stankovic, and V. Stankovic, "On a Training-Less Solution for Non-Intrusive Appliance Load Monitoring Using Graph Signal Processing," *IEEE Access*, vol. 4, pp. 1784–1799, Apr. 2016.

[9] B.J. Birt and et. al., "Disaggregating categories of electrical energy end-use from whole-house hourly data," *Energy and Buildings,* vol. 50, pp. 93–102, Jul. 2012.

[10] M. Wytock, J. Kolter, "Contextually Supervised Source Separation with Application to Energy Disaggregatio," *in AAAI*, pp.486–492, Jul. 2014.

[11] J. Kolter and M. Johnson, "REDD: A public data set for energy disaggregation research," in *SIGKDD*, San Diego, CA, Jan. 2011.

[12] D. Murray, L. Stankovic, and V. Stankovic, "An electrical load measurements dataset of United Kingdom households from a two-year longitudinal study," *Nature Scientific Data*, vol. 4, pp. 1–12, Jan. 2017.

[13] S. Makonin, B. Ellert, and I. V. Bajic, and F. Popowich "Electricity, water, and natural gas consumption of a residential house in Canada from 2012 to 2014," *Nature Scientific Data*, vol. 3, pp. 1–12, Jun. 2016.