Low-power Appliance Recognition using Recurrent Neural Networks

Rizky Pratama, Azkario; Simanjuntak, Frans Juanda; Lazovik, Aliaksandr; Aiello, Marco

Published in:
Frontiers in AI and Applications (FAIA)

Low-power Appliance Recognition using Recurrent Neural Networks

Azkario R. Pratama [a,1], Frans J. Simanjuntak [b], Alexander Lazovik [a] and Marco Aiello [a]

[a] *Distributed Systems Group, Johann Bernoulli Institute for Mathematics and Computer Science, University of Groningen, The Netherlands*

[b] *Master Student in Computing Science, Faculty of Science and Engineering, University of Groningen, The Netherlands*

Abstract. Indoor energy consumption can be understood by breaking overall power consumption down into individual components and appliance activations. The classification of components of energy usage is known as load disaggregation or appliance recognition. Most of the previous efforts address the separation of devices with high energy demands. In many contexts though, such as an office, the devices to separate are numerous, heterogeneous, and have low consumptions. The disaggregation problem becomes then more challenging and, at the same time, crucial for understanding the user context. In fact, from the disaggregation one can deduce the number of people in an office room, their activities, and current energy needs. In this paper, we review the characteristics of office appliances load disaggregation efforts. We then illustrate a proposal for a classification model based on Recurrent Neural Network (RNN). RNN is used to infer device activation from aggregated energy consumptions. The approach shows promising results in recognizing 14 classes of 5 different devices being operated in our office, reaching 99.4% of Cohen's Kappa measure.

Keywords. Appliance recognition, Load disaggregation, Deep neural network, Energy smart systems, Smart building

1. Introduction

The trend of electricity consumption monitoring has been shifting from the analog power meter to the digital one [1]. It then becomes easy to collect vast amounts of power consumption data and mine it to get better context insight into user behaviors, especially indoors. Most power meters which have been installed in buildings measure composite loads of devices at a single point of measurement, such as in a circuit breaker of a room, of a building floor, or even of on entire department building. Therefore, mining an aggregated power consumption is more realistic and challenging than mining per-device data. Extracting meaningful information from an aggregated data can be beneficial in cases such as: 1) breaking down the energy load per-appliance for giving insight to the building administrator of energy expense, consumption pattern, and thus potential energy saving; or 2) providing enriched contexts to a smart system to adapt the building behavior accordingly and, hence, satisfying predetermined building managers' goals.

[1] Corresponding Author: a.r.pratama@rug.nl

Previous works have shown successful recovery rates in identifying individual power consumptions from a composite load measured in a single point of measurement, e.g., [2,3,4]. However, the scope of these works mainly covers energy hungry devices in residential buildings, such as microwave ovens, fridges, front load washers, and heat pumps [2]. The electricity consumption is relatively easily separable in these cases (e.g., consuming different amounts of power, having regular temporal patterns, and different power consumption behavior due to various composing circuits). In the office setting, one cannot make use of these characteristics. Of the three types of electrical appliances, i.e., resistive, inductive, and capacitive [3], typical office appliances (e.g., monitor, CPU, beamer, laptop charger) fall into the same class, which is capacitive. Moreover, the office appliance consumption is relatively low compared to the residential appliances, for example, 14 Watts of power consumption for a 22 inch monitor screen compared to 80 Watts an hour for a home television. This rises of the complexity of disaggregation due to oscillations and masked low-power consumptions [5].

Despite these difficulties, we focus on monitoring an office. The reason is that the related loads contribute for more than 70% of the baseline electricity load, even though the individual loads consume low amounts of power [6]. We provide an analysis of appliance profiling in terms of cosphi, current, active-, reactive-, and apparent-power from a real measurement of office space equipment at the University of Groningen, The Netherlands. Furthermore, we explore the use of deep learning techniques, namely Recurrent Neural Network (RNN) in this context, due to its effectiveness in sequence modeling.

The key contributions of the present work can be summarized as follows: 1) Identifying the behavior of low-power appliances commonly present in offices; and 2) Presenting the implementation of Recurrent Neural Network to recognize active devices from a single point of measurement. The paper organization is as follows. We briefly introduce the typical type of office appliances and our recognition design in Section 2. The experimental setup and metrics are presented in Section 3. We report results and discussion in Section 4. Related works and concluding remarks are presented in Section 5 and Section 6, respectively.

2. Design and Implementation

To get a better understanding of devices' behavior, we analyze the typical types of appliances in terms of their power consumption. We thus discuss our effort to use machine learning algorithms in classifying appliance activation.

2.1. Appliance Profiling

According to their inner circuit, devices can be classified into three types, i.e., resistive, capacitive, and inductive [3]. A purely resistive device is the one that has the current waveform in the same phase as the voltage, such as an incandescent lamp or a heater. As for the capacitive and inductive devices, the inductor and capacitor components shift the current wave with respect to the voltage wave. The shifting can be leading (for capacitive loads) or lagging (for inductive loads). This shape can be observed through the active and reactive power components, as if voltage and current are shifted, the power transferred to the appliance is lower than when no shifts occur. Those appliances with motors such as pumps, generators, and washing machines are inductive loads; those with capacitors used in their circuits, such as laptops and TVs, are capacitive loads. The relation of active- and apparent-power is measured as power factor PF, shown in Equation 1. The closer the value to one, the more pure (resistive) the devices is.

$$PowerFactorPF = \frac{ActivePower(P)}{ApparentPower(S)} = cos(\phi) \tag{1}$$

Even though there are three types of common appliances, it does not mean that those devices that fall in the same category have the same characteristic. For example, capacitive devices may shift current wave with respect to voltage wave, but the degree of shifting might be different and thus could be the indication of a pattern. Moreover, if we combine different appliance loads measured at one point, we will form a combination load constructed from the interaction of individual loads that can be either strengthening or canceling out one another. For instance, the typical appliances in an office are a desktop PC and monitors. The characteristic of the measured power does not always follow the nature of the components. Hence, if the aforementioned appliances are activated at the same time, it is hard to predict the effect on the electric footprints. Appliance profiling provides prior insight of how the actual power consumption is drawn before dealing with device classification.

2.2. Appliance Recognition

We address the recognition problem by applying a supervised pattern recognition technique. The observation of devices' signatures is done through an instantaneous snapshot taken at any fixed interval or time windows, instead of capturing the switching ON/OFF state [7]. The reason is that the switching states of the low-power device are easily masked by the ripple or oscillation of the power consumption of operating devices, thus negatively affecting systems' robustness [8].

For each time interval, we capture several power-related signatures using a power meter, namely active-, reactive-, and apparent-power, power factor, and current. These signatures are described as follows.

Active Power (P), Watt, the actual amount of power being used;

Reactive power (Q), Volt-Amps-Reactive, the dissipated power resulting from inductive and capacitive loads;

Apparent power (S), Volt-Amps, the product of the root-mean-square voltage and the root-mean-square current;

Power factor or cosphi, percentage, representing the ratio of the real power flowing to a load divided by the apparent power; and

Current, in Ampere, the electron flow in an electrical conductor.

We experiment with three feature combinations to compare recognition performance based on various measurements, as shown in Table 1. The first combination consists of active- and reactive-power, as proposed by Hart *et al.* [9]. The second set includes active power, current, and cosphi. In the last feature set, we use five components measured by the power meter.

Table 1. Feature sets in the experiment

Feature Set	Signatures
FS-1	Active- and reactive-power
FS-2	Active power, current, and cosphi
FS-3	Active-, reactive-, and apparent-power, current, and cosphi

Further, we select a pattern recognition approach to infer the most likely class given the signature sequences. We choose Recurrent Neural Network (RNN) as it benefits in operating over sequences of vectors [10]. RNN is a type of neural network that takes the previous output as the next input value in a sequence. The intuition behind the chosen technique is that the activity of office related devices is relatively stable, i.e., once a monitor and CPU are being used, most likely they will remain in the same state for hours. Hence, we aim at discovering whether giving a notion of order in time provides better insight than only supplying one single point of measurement to the classifier. We utilize LSTM (Long Short-Term Memory) to deal with the vanishing gradient problem. That is, the problem that arises when data is flattened due to passing hidden layers (e.g., applying a logistic sigmoid function) multiple times. LSTM is a technique that is useful in deciding when to forget or keep the current input for the future output.

We design an RNN architecture such that we can input a sequence (e.g., 60 data points during 5 minutes) of n-dimensional features to the model. The taken sequences are from non-overlapped moving windows, as performance is better than the overlapping windows, according to our experiments [11]. For each sequence, we expect a class label to be learned or classified. We feed 20 mini-batches of the sequences in an iteration to reduce the number of looping needed to complete one epoch, thus, speeding up the learning process. We apply either one or two LSTM layers (we refer to LSTM or LSTM*, respectively, in the rest of this paper) stacked in the hidden layer of the network. Each LSTM layer consists of h_t hidden states, where $h_t = 20$. The reason of the chosen architectures is that the layer size is not more important than the layer depth [12], hence we stick in the fixed size of h_t and change the level of the hidden layer. We apply Adam optimizer [13] and determine the learning rate of 0.001 to optimize a cost function during iteration. The cost function is based on cross entropy [14].

We need to determine the number of epochs to be completed to make sure that the model has learned sufficiently without memorizing the training data (overfitting). To do so, we implement the early stopping strategy. We evaluate the model performance on a validation set and save the best model snapshot when it outperforms the previous best winner. We terminate the training when the network is not improving after the i-th epoch, where $i = 100$ when the maximum number of epochs is 400, and $i = 50$ when the maximum is set to 200.

To evaluate the RNN approach, we also consider k-NN (with $k = 7$) [15] and Support Vector Mechine (SVM) with linear- and polynomial-kernel [16] as a comparison. We consider k-NN as this is one of the simplest techniques. It works by assigning a label based on the highest votes of the nearest samples into a query. SVM is a widely used technique for classification due to its ability to generate nonlinear decision boundary. It was initially designed for binary classification problems. The multi-class problem is solved by combining several binary classifiers (i.e., 'one-against-one' approach) [17]. Both k-

Table 2. List of used devices

Id	Device	Rated power
1	Monitor 22' brand Iiyama	21 W (max)
2	Monitor 22' brand Philips	15.43 W (normal)
3	CPU brand HP Z230	400 W PSU
4	Laptop charger MacBook Pro	60W
5	Heater Philips	2000 W (max)

NN and SVM are designed to classify on the basis of one single point of measurement, instead of giving a sequence of data as input.

3. Experiment

We monitor office appliances in a workspace at the University of Groningen, The Netherlands, during several weeks between June 2017 and February 2018, that is in more seasons. The goal of the experiment is to evaluate the performance of Recurrent Neural Networks to recognize active devices from a single point measurement.

3.1. Setup

We consider several office-related appliances, listed in Table 2. The rated power information listed in the table is given by the manufacturer. For some devices (e.g., CPU), there is insufficient information about the averaged power consumption, as it indeed depends on the actual operation. Therefore, we provide the capacity of Power Supply Unit (PSU) as an indication of the maximum power that can be drawn by the PC.

The appliances are connected to room electrical sockets that supply electricity from a central electric network of the building. We attach a measurement kit consisting of a power meter node and a clip. The clip is clamped on the main line to measure the total power consumed by the appliances. For simplicity in deployment, the kit can be placed in the fuse box that is commonly installed in particular area (e.g., in a room or zone).

We use Smappee[2] energy monitor device to collect the data. To collect denser data (i.e. in 5 seconds interval), we use a local hub library[3] and build our gateway to receive and store data from the sensor node. We refer to [18] for the data flow diagram coming from sensors (i.e., power meter) to an integrated server for further process.

We collected 92,580 data points (about 128.6 hours) from 14 classes, consisting of both individual appliances (i.e., laptop, CPU, monitor-1, monitor-2, and electric heater) and some combinations of these devices. The raw data is normalized to the scale between 0 and 1, depending on the minimum and maximum value of each feature. We utilize k-fold cross-validation, with $k = 5$, to assess model generalization to the dataset. It is worth noting that as we analyze sequence data using RNN, we randomly partition the data based on the group of data points during 5 minutes interval.

[2]https://www.smappee.com/nl/home
[3]https://github.com/NMichas/smappee-local-mqtt

3.2. Metrics

In order to evaluate the classification, we provide total accuracy and confusion matrix to give better insight of classification performance per-class. We also consider Cohen's Kappa measure [19] to eliminate the accuracy bias due to imbalanced class distribution.

Accuracy defines the number of windows with correctly-predicted label divided by the total number of classifications made, i.e., $totalAccuracy = \frac{correctlypredictedclass}{N}$, where N is the total number of windows being classified.

Cohen's Kappa measures the agreement between accuracy of the system to the accuracy of a random system, as shown in Equation 2. The total accuracy is an observational probability of agreement while the random accuracy is a hypothetical expected probability of agreement under an appropriate set of baseline constraints [19].

$$kappa = \frac{totalAccuracy - randomAccuracy(RA)}{1 - RA} \tag{2}$$

where RA is the sum of the products of reference likelihood and result likelihood for each class. Mathematically,

$$randomAccuracy = \frac{\sum_{c \in C} actualclass_c * predictedclass_c}{N^2}$$

4. Results and Discussion

4.1. Appliance profiling

The properties of five different devices is shown in Table 3. As for the individual devices, the electric heater consumes the highest active power, up to 1,015 Watt. For the other office-related devices, they consume less than 30 Watt on average. In spite of the highest power consumption, the heater draws the lowest reactive power among the others, about 15.8 VAR, giving a very high cosphi percentage, 99.0%. Monitor-1 and monitor-2 have one-third active power difference, but their cosphi percentages are comparable, reaching about 58%. The laptop draws 48.7% cosphi value and is the lowest cosphi percentage among the individual appliances. The laptop and CPU are the most fluctuating loads, as shown by a high standard deviation value, up to 4.94 Watt, of their averaged power consumption, i.e., 20 and 29.6 Watts, respectively.

As for composite loads, when we combine CPU with monitors (either monitor-1, 2, or both), the cosphi value will be slightly higher than CPU's cosphi itself, ranging between 72-74% on average. Interestingly, when we add one more device, such as a heater or a laptop charger, the cosphi value can either increase or decrease, reaching 99% and 68.9%, respectively. By adding both devices, the cosphi reaches about 98.98%.

While the cosphi values do not follow the additive criterion, the active power values do. That is, the total power (in Watt) is the sum of the power consumption of its unified devices. For example, two monitors and a CPU in total consume an average power of

Table 3. Device consumption properties

Devices	Id	Apparent (S)	Reactive (Q)	Active (P)	Current (I)	Cosphi	Volt
Laptop	10	39.9944	34.4420	20.3187	0.1690	48.7352	235.9286
	std.	8.7177	7.2080	4.9454	0.0370	2.2287	0.5990
CPU	1	42.6769	30.6346	29.6370	0.1826	67.5016	233.0149
	std.	4.7489	1.7105	4.9080	0.0203	3.7528	1.4248
Mon-1	2	34.8440	27.6004	21.2403	0.1485	58.9423	233.7170
	std.	2.7060	2.1001	2.0144	0.0114	4.9012	1.7116
Mon-2	3	23.5081	18.8441	14.0503	0.1004	56.1919	232.8989
	std.	0.9884	0.7579	0.7092	0.0043	2.3941	1.2056
Heater	8	1015.6	15.8	1015.4	4.4	99.0	229.7
	std.	5.4440	4.8278	5.3955	0.0110	0	0.7838
CPU + mon-1	4	71.1456	46.5500	53.7291	0.3055	74.0030	232.4804
	std.	6.5571	2.1491	6.8051	0.0282	2.8612	1.0642
CPU + mon-2	5	63.9867	41.5065	48.5110	0.2742	73.8093	232.8909
	std.	9.6870	3.3673	10.0324	0.0418	5.6097	0.9138
CPU + 2 monitors	7	92.8489	62.3019	68.7935	0.3981	72.8101	232.9445
	std.	5.6886	2.2540	5.8374	0.0251	2.7407	1.0220
Mon-1 + mon-2	6	58.7115	46.6721	35.5946	0.2514	59.0784	233.0153
	std.	4.3197	3.2825	3.1019	0.0186	4.0328	0.9249
Laptop + 2 monitrs	13	94.1747	76.9155	54.3128	0.4024	56.5481	233.7015
	std.	6.8031	5.0313	4.8889	0.0295	1.7659	0.8893
CPU + montrs + laptop	11	114.6322	81.8831	80.1831	0.4925	68.9402	232.4662
	std.	7.4370	4.3592	6.5265	0.0318	1.9694	0.9942
CPU + montrs + heater	9	1062.9	44.2	1061.9	4.7	99.0	227.9
	std.	32.0644	7.6395	33.5706	0.1365	1.0007	1.1147
All 5 devices	12	1089.3334	70.8284	1086.9892	4.7691	98.9819	228.3417
	std.	24.5299	7.4510	25.1245	0.1020	0.8412	1.1263

68.79 Watts, that is the amount of the sum of the power of each device. Current (in Ampere) also behaves in the same way. The total current flow through the electric extension socket is equal to the sum of the currents through each appliance's circuit. It does not happen to the apparent- and reactive-power.

4.2. Appliance recognition

The classification results are shown in Table 4. In general, the performance indicated by the accuracy and Kappa measure is comparable. The reported measures are the average of the completion of 5-folds cross-validation. The RNN with LSTM based classification delivers Kappa between 60-97%, depending on the network configuration, the number of iteration on the training phase and the feature set taken into account. LSTM and LSTM* differ in terms of cells number. The former uses a single LSTM cell while the latter uses two. Using LSTM, we can achieve a Kappa measure of up to 90.1%, while based on LSTM*, 96.8% of the same measure can be achieved. These results can be improved up to 97.6-99.4% by increasing the number of training epochs. The higher number of epoch iterations are allowed, the better results are obtained.

Table 4. Cohen's Kappa measure of classifiers with different feature sets

Method	Accuracy			Kappa measure			Remarks
	FS-1	FS-2	FS-3	FS-1	FS-2	FS-3	
LSTM	0.73	0.629	0.909	0.709	0.599	0.901	iteration max 200 epochs
LSTM*	0.779	0.719	0.97	0.761	0.696	0.968	
LSTM	0.819	0.782	0.978	0.804	0.765	0.976	iteration max 400 epochs
LSTM*	0.848	0.841	0.995	0.836	0.829	0.994	
k-NN	0.999	0.996	0.999	0.999	0.996	0.999	$k = 7$
SVM (lin)	0.84	0.92	0.934	0.827	0.914	0.928	$C = 1$
SVM (poly)	0.31	0.444	0.456	0.24	0.396	0.408	degree=3, $C = 1$

In general, the classifications using the feature set-3 achieve higher performance than the other sets, reaching at least 90%. SVM with polynomial kernel is an exception. While SVM with linear kernel can achieve Kappa measure of about 82-93%, the same classifier with polynomial kernel (degree 3) reaches roughly 40% for feature set-2 and 3. The k-NN based inference results up to 99.9% Kappa measure on the all set of predictors.

4.3. Discussion

The utilization of power meter supports the appliances' behavior observation from different measurements, e.g., power consumption, current flow, and the shifting between current- and voltage-wave. From the observation, we clarify that the four of five of the office-related appliances fall in the capacitive type. It is indicated by measured cosphi. Office-related loads, such as CPU, monitors, and laptop, shift the current waveforms with respect to the voltage waveforms. They deliver higher reactive power, and finally affect the cosphi value. Even though these devices fall in the same capacitive category, they have a different degree of shifting. Two different types of monitors have a comparable behavior regarding phase shifting, while CPU has a higher one. Among four appliances, the laptop is the most capacitive. It has the smallest cosphi percentage, thus the highest phase shifting. We can also see from the dissipated reactive power that is the highest among the others. The shifting does not happen for the heater, where the cosphi is measured at 99%. It means that the heater does not generate any phase difference, or in the other word, the active- and apparent-power are in the same direction. It is because the heater circuit is resistive where the delivered active power is directly converted to heat.

The behavior of the composite load changes when several devices are activated simultaneously. The response behavior depends on the comprising devices. The more resistive devices being attached, the higher cosphi being measured (or, the more vanishing current-voltage shifts). For instance, when we add a big resistive load such as a heater, the cosphi is climbing, reaching 99%. In contrast, the more capacitive the device we add, the higher dissipated reactive power will be measured. Thus, higher apparent power will be obtained, and cosphi value will decrease. One example is when we combine the laptop to the monitors, where the cosphi is reduced of about 3%.

The appliances classification, given the sequence data of aggregated power consumption, shows promise when we supply the right predictors or features to the RNN classifier. In our setting, feature set-3 provides the best classification outputs among the other two sets. This indicates that the classifier recognizes patterns, resulting the promis-

Table 5. Confusion matrix of the predicted classes. Left-side: single LSTM cell (epoch-171); Right-side: double cells (epoch-174)

Pred. Act.	0	1	2	3	4	5	6	7	8	9	10	11	12	13	Pred. Act.	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	23	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	24	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	22	0	0	0	2	0	0	0	0	0	0	0	0	**1**	0	21	0	0	1	2	0	0	0	0	0	0	0	0
2	0	0	24	1	0	0	0	0	0	0	0	0	0	0	**2**	0	0	23	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	22	0	0	0	0	0	0	0	0	0	0	**3**	0	0	0	25	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	11	13	0	0	0	0	0	0	0	0	**4**	0	0	0	0	24	0	0	0	0	0	0	0	0	0
5	0	3	0	0	4	24	0	0	0	0	0	0	0	0	**5**	0	0	0	0	10	22	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	23	0	0	0	0	0	0	0	**6**	0	0	0	0	0	0	24	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	23	0	0	0	0	0	0	**7**	0	0	0	0	0	0	0	25	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	19	0	0	0	0	0	**8**	0	0	0	0	0	0	0	0	18	0	0	0	0	0
9	0	0	0	0	0	0	0	0	1	17	0	0	0	0	**9**	0	0	0	0	0	0	0	0	0	18	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	24	0	0	0	**10**	0	0	0	0	0	0	0	0	0	0	22	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	17	0	0	**11**	0	0	0	0	0	0	0	0	0	0	0	16	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	19	0	**12**	0	0	0	0	0	0	0	0	0	0	0	0	20	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	16	**13**	0	0	0	0	0	0	0	0	0	0	0	0	0	14

ing results (i.e., at least 90.1% Kappa measure), based on the feature set as a predictor. In this case, the pattern is the form of the waveforms during observation window (e.g., in 5 minutes interval). The experiments also reveal that the two cells LSTM* outperforms the single stack LSTM cell in all predictor combinations and iteration epochs. Such results support the finding that an RNN earns benefits in having deeper architecture [12].

The confusion matrix presented in Table 5 illustrates the classification outputs of one and two LSTM cells in the comparable epoch iteration (epoch 171 and 174). At first glance, both models are confused with class label 4 and 5 (i.e., CPU + mon-1 and CPU + mon-2). However, the model with two LSTM cells is better than the one with a single cell in predicting device CPU + mon-1 (label 4), indicating that the two cells model have faster convergence over the same epochs. Given similar results on the other folds, we believe that in this phase of iteration, the LSTM-based models may have not fully converged yet and more training steps are necessary.

The higher number of epoch iterations give better results both in one LSTM cell and two stacked LSTM cells. It is because the approach is based on iterative optimization (i.e., Adam optimizer [13]). The approach tries to adjust model's weights along the iterations to find the optimal network configuration.

LSTM mostly outperforms non-probabilistic binary linear classifier SVM providing the right predictors (i.e., feature set-3). When we apply SVM with polynomial degree-3 kernels, the classifier does not perform as expected. The reason is that, in SVM, decision boundary is decided by a hyperplane that is shaped by kernel functions. As for the polynomial kernel, the model tries to overfit the training data and fails to classify the rest of data, while the decision boundary formed by linear kernel works quite well in this work, reaching 92.8% Kappa measure. Interestingly, *k*-NN performs very well for all set of predictors. This might indicate that feature combinations are distinct and non-overlapping, where *k*-NN performs well in [15]. As a comparison, when we put only active power (P), *k*-NN will return worse result than the one reported in this work.

In the current work, we take into account five different devices by considering binary states. For laptop and PC, even though they have considerably numerous states, we utilize them only in two states (either ON or OFF). We consider ON when we activate CPU and leave it idle, or execute simple tasks (e.g., streaming and updating software). As for the laptop, it keeps on charging while in fully-charged condition, leaving relatively stable power consumption. As the approach is based on supervision, it requires the training phase before the operation one. Recognizing an un-trained condition is not possible and will result in misclassification. The scalability of the approach relies on the deployment of power meters. The more electric interconnection under the power meter, the more appliances being measured. Hence, the classification is even more challenging. Another limitation is that the experiments were conducted with one subject only. We plan to involve more people to deduce context from the power consumption.

5. Related work

Previous studies for recognizing appliances from electricity fingerprints have been proposed. For example, Reindhardt *et al.* observe the behavior of individual device through plug-based power meters [20]. By developing pattern recognition models (e.g., decision tree and Bayesian-based techniques), they classify unlabelled load signature to recognize appliances. They then identify the most relevant predictors on the basis of the amount of information gain. They report more than 90% True Positive for specific devices. However, they fail to distinguish LCD from CRT. Furthermore, they do not experiment with separating individual devices from a composite electricity load.

Balaji *et al.* made an open dataset of plug loads in the office environment [21]. The motivation is to support analysis of diverse appliance types and identical instances. Nevertheless, at the time of writing, the database has not been published.

Rogriguez *et al.* proposed an approach for individual device profiling [22]. They discover several properties, such as current (Ampere), settling time (ms), and phase shifting for binary state transition (ON/OFF). The measured data has considerably high sampling rate (1 KHz). By using decision trees, the authors identify unlabeled loads and evaluate the performance using 10-fold cross validation for individual loads and hold-out-validation method for aggregated loads. The accuracy is in the range of 50-80%, though it is not clear how much aggregated loads they have analyzed. Low-power workstation devices achieve worse results than kitchen appliances due to their lower power profile, i.e., $< 0.2A$. The thresholding-based classifier, such as decision tree, highly rely on the thresholds compatibility, hence, inappropriate threshold sets significantly affects classifier performance. In the present work, we utilize a clamp-based power meter that has lower sampling rate (5 seconds interval) and richer parameters (reactive-, active-, apparent-power, current, and cosphi). We focus on common appliances in an office where the achieved accuracy still lower than in the kitchen or household.

There are two signature forms in solving load disaggregation problems through pattern recognition, namely snapshot and delta forms [7]. While some works are based on delta form [22,8], the work of Zoha *et al.* [23] relies on the snapshot form taken every 15 seconds. They use Factorial HMM to perform load disaggregation of devices in relation to a work desk. Based on ten experiments, the F-measure performances range between 76-98% for recognizing 2-5 appliances. However, there is no further analysis of particular office devices, especially in terms of cosphi, reactive power, etc. From the results, it is hard to understand how appliances affect each other.

Using deep neural networks for solving the load disaggregation problem has been proposed by Kelly *et al.* [10]. The authors adapt three deep neural network architectures, including Recurrent Neural Network with Long Short-Term Memory (LSTM) networks. The results suggest that LSTM outperforms auto-encoder and “rectangle” neural network architectures for binary states appliances. However, the research scope is in homes (using an open dataset with 6 seconds sampling interval) that consist of several high-power appliances, i.e., 300-3100 Watts maximum power for five appliances (kettle, fridge, washing machine, microwave, and dishwasher). Furthermore, their work relies on the appliance activation, that is, the power drawn of a single appliance over one complete cycle instead of a defined time window. The evaluation is based on synthetic data. In contrast, we use actual measurements of various office-related appliances. We observe per-device behavior to give an insight of how our lower power devices behave and build an RNN model and infer active devices for each time window.

6. Concluding Remarks

We have presented a comprehensive review of the power consumption of five low-power appliances in an office. A readily available power meter was used to measure aggregated electricity footprints in a room as an impact of appliances utilization. The observable footprints include active-, reactive-, and apparent-power, current, and cosphi.

Some common office appliances (such as monitor, CPU, and laptop) fall in the capacitive category due to their inner component in nature. That is, these appliances have major capacitance components (such as an integrated circuit that consists of capacitors) to perform a specific function. Such a component makes current waveform leading with respect to voltage waveform. Nevertheless, the gap between the two varies among the appliances. In this work, we found that the laptop makes the highest lead between current and voltage waveform, followed by two different monitors and CPU. The different behavior of the observable appliances potentially reveals the pattern that can be recognized to provide the enriched context of the spaces.

We then proposed an appliance recognition based on RNN to the sequence of data. It is beneficial for understanding the context of the spaces, such as identifying occupancy, recognizing activity, and estimating the number of people [24]. We showed that RNN can be a viable option to solve the appliance recognition problem, with a Kappa measure of up to 99.4% of the detection for 14 different classes. Such results outperform the SVM approach. However, the results of *k*-NN seems to be better in terms of its robustness to the three different sets we have experimented.

Acknowledgement

Azkario Rizky Pratama and Frans Juanda Simanjuntak are supported by the Indonesia Endowment Fund for Education (LPDP). The research is also supported by the H2020 ERA-Net Smart Grids Plus project MatchIT, NWO contract number 651.001.011 and by the H2020-RISE FIRST project.

References

[1] M. Aiello and G. A. Pagani, "How energy distribution will change: an ICT perspective," in *Smart Grids from a Global Perspective*, pp. 11–25, Springer, 2016.

[2] S. Makonin, F. Popowich, L. Bartram, B. Gill, and I. V. Bajić, "AMPds: A public dataset for load disaggregation and eco-feedback research," in *2013 IEEE Electrical Power Energy Conf.*, 2013.

[3] A. G. Ruzzelli, C. Nicolas, A. Schoofs, and G. M. P. O'Hare, "Real-time recognition and profiling of appliances through a single electricity sensor," in *2010 7th Annual IEEE Communications Society Conf. on Sensor, Mesh and Ad Hoc Communications and Networks*, pp. 1–9, June 2010.

[4] A. Cominola, M. Giuliani, D. Piga, A. Castelletti, and A. Rizzoli, "A hybrid signature-based iterative disaggregation algorithm for non-intrusive load monitoring," *Applied Energy*, vol. 185, no. Part 1, pp. 331 – 344, 2017.

[5] M. Weiss, A. Helfenstein, F. Mattern, and T. Staake, "Leveraging smart meter data to recognize home appliances," in *IEEE Int. Conf. on Pervasive Computing and Communications*, pp. 190–197, March 2012.

[6] Y. Agarwal, T. Weng, and R. K. Gupta, "The energy dashboard: improving the visibility of energy consumption at a campus-wide scale," BuildSys'09, 2009.

[7] J. Liang, S. K. K. Ng, G. Kendall, and J. W. M. Cheng, "Load signature study 2014;part i: Basic concept, structure, and methodology," *IEEE Trans. on Power Delivery*, vol. 25, pp. 551–560, April 2010.

[8] A. R. Pratama, W. Widyawan, A. Lazovik, and M. Aiello, "Power-based device recognition for occupancy detection," in *Service-Oriented Computing — ICSOC 2017 Workshops*, vol. in press, Springer, Nov 2017.

[9] G. W. Hart, "Nonintrusive appliance load monitoring," *Proc. of the IEEE*, vol. 80, pp. 1870–1891, Dec 1992.

[10] J. Kelly and W. Knottenbelt, "Neural NILM: Deep neural networks applied to energy disaggregation," in *Proc. of the 2nd ACM Int. Conf. on Embedded Systems for Energy-Efficient Built Environments*, BuildSys '15, pp. 55–64, ACM, 2015.

[11] F. J. Simanjuntak, "Deep learning approach for electric appliances recognition." `http://ugm.id/fsimanjuntak`, August 2017.

[12] A. Graves, A. r. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *2013 IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pp. 6645–6649, May 2013.

[13] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint:1412.6980*, 2014.

[14] P. Golik, P. Doetsch, and H. Ney, "Cross-entropy vs. squared error training: a theoretical and experimental comparison.," in *Interspeech*, vol. 13, pp. 1756–1760, 2013.

[15] G. Shakhnarovich, T. Darrell, and P. Indyk, *Nearest-Neighbor Methods in Learning and Vision: Theory and Practice (Neural Information Processing)*. The MIT Press, 2006.

[16] A. Ben-Hur and J. Weston, "A user's guide to support vector machines," in *Data Mining Techniques for the Life Sciences*, pp. 223–239, Humana Press, 2010.

[17] S. Knerr, L. Personnaz, and G. Dreyfus, "Single-layer learning revisited: a stepwise procedure for building and training a neural network," in *Neurocomputing*, pp. 41–50, Springer Berlin Heidelberg, 1990.

[18] A. R. Pratama, W. Widyawan, A. Lazovik, and M. Aiello, "Multi-user low intrusive occupancy detection," *Sensors*, vol. 18, no. 3, 2018.

[19] J. Cohen, "A coefficient of agreement for nominal scales," *Educational and Psychological Measurement*, vol. 20, no. 1, pp. 37–46, 1960.

[20] A. Reinhardt, P. Baumann, D. Burgstahler, M. Hollick, H. Chonov, M. Werner, and R. Steinmetz, "On the accuracy of appliance identification based on distributed load metering data," in *2012 Sustainable Internet and ICT for Sustainability*, pp. 1–9, Oct 2012.

[21] B. Kalluri, S. Kondepudi, K. H. Wei, T. K. Wai, and A. Kamilaris, "OPLD: Towards improved non-intrusive office plug load disaggregation," in *2015 IEEE Int. Conf. on Building Efficiency and Sustainable Technologies*, pp. 56–61, Aug 2015.

[22] A. Rogriguez, S. T. Smith, A. Kiff, and B. Potter, "Small power load disaggregation in office buildings based on electrical signature classification," in *2016 IEEE Int. Energy Conf.*, pp. 1–6, April 2016.

[23] A. Zoha, A. Gluhak, M. Nati, and M. A. Imran, "Low-power appliance monitoring using factorial hidden markov models," in *2013 IEEE 8th Int. Conf. on Intelligent Sensors, Sensor Networks and Information Processing*, pp. 527–532, 2013.

[24] M. Milenkovic and O. Amft, "An opportunistic activity-sensing approach to save energy in office buildings," in *Proc. of the 4th Int. Conf. on Future Energy Systems*, e-Energy '13, pp. 247–258, 2013.