

Unseen Appliances Identification

Antonio Ridi^{1,2}, Christophe Gisler^{1,2}, and Jean Hennebert^{1,2}

¹ University of Applied Sciences Western Switzerland
College of Engineering and Architecture of Fribourg, ICT Institute
{antonio.ridi, christophe.gisler, jean.hennebert}@hefr.ch

² University of Fribourg
Department of Informatics, Fribourg, Switzerland
{antonio.ridi, christophe.gisler, jean.hennebert}@unifr.ch

Abstract. We assess the feasibility of unseen appliance recognition through the analysis of their electrical signatures recorded using low-cost smart plugs. By unseen, we stress that our approach focuses on the identification of appliances that are of different brands or models than the one in training phase. We follow a strictly defined protocol in order to provide comparable results to the scientific community. We first evaluate the drop of performance when going from seen to unseen appliances. We then analyze the results of different machine learning algorithms, as the k-Nearest Neighbor (k-NN) and Gaussian Mixture Models (GMMs). Several tunings allow us to achieve 74% correct accuracy using GMMs which is our current best system.

Keywords: Intrusive Load Monitoring (ILM), appliance recognition, electric signatures, load identification.

1 Introduction

The automatic recognition of appliances from their electric signatures has several applications such as energy consumption understanding and appliance management for energy consumption optimization [1]. Other applications can also be envisioned such as an indirect activity detection in houses or monitoring of elderly people [2].

Due to the rising price of energy and an increased sensitivity from people to environmental matters, the field of energy consumption understanding and management is nowadays rising interests. In US, 51% of the electricity consumption in homes is due to appliances and lighting [3]. In this context, a system able to recognize appliance would allow to know which appliance is consuming how much, giving an explanation on their contribution to the electricity bill. This will allow householders to optimize their energy consumption. Appliance identification could also be very useful for Building Management Systems (BMS), allowing to implement smarter rules and optimizing the local production and consumption of electric energy.

An electric signature represents the time evolution of the electricity consumption which is summarized by the active and reactive power on AC networks.

Appliances can be categorized into 4 classes [4]: two-states on/off appliances (e.g. lamps, toasters); multi-states appliances, when a finite number of operating states exist (e.g. fridges, dishwashers); continuously variable devices, when the consumption varies continuously (e.g. battery chargers); permanent consumer devices, when the consumption is constant over a long period of time (e.g. telephone sets, smoke detectors). According to this, a more complex task than appliance identification could consist of recognizing in which state a given appliance is at a given time, allowing for example to automatically detect stand-by. Recent studies are estimating this consumption at about 10% of the residential electricity use [5,6].

Appliance identification can be done using two approaches: Non-Intrusive Load Monitoring (NILM) and Intrusive Load Monitoring (ILM) [7]. NILM monitors the total house electricity consumption at the smart meter, while the ILM refers to a distributed sensing approach, using one or more sensor per appliance. In the first case, the signals have to be decomposed to identify single appliance, i.e. performing a *disaggregation* [4]. NILM approaches are less expensive but more difficult while ILM approaches are more expensive but more precise [7]. We focus in this paper on ILM approaches.

As detailed in Section 2, several modeling approaches have been proposed for appliance identification, often based on machine learning principles. Given the differences among brands and models, a challenge for such approaches is in the necessity to have large training databases that represent all types of appliances for a given type, including as many brands and models as possible. In this paper we address the problem of identification systems that are tested with unseen appliance, i.e. appliance brands that are not available in the training set. In other words, we evaluate the generalization capacity of such machine learning systems when dealing with new appliances that are not yet observed in the training set. In this direction, the availability of large databases is important (Section 3). System description, results and discussions are presented in Section 4 and 5.

2 Related Works

Several ILM approaches have been proposed. In the work of F. K. Adeel Abbas Zaidi and P. Palensky [8], machine learning approaches such as Dynamic Time Warping (DTW) and Hidden Markov Models (HMMs) are presented. Observations are sampled at 10^{-1} Hz on different appliances spread into 6 categories including fridges, microwaves, dishwashers, coffee machines, computers and printers. From the raw observations, features are extracted such as average energy consumption, edge counts, percentage energy consumption and discrete Fourier transform coefficients. The best feature sets are showing results up to 90% for five categories.

In the work of Reinhardt et al. [9], 33 appliance categories are used to build an identification system showing promising performance up to 95.5% accuracy. The system samples the current consumption at 1.6 kHz which brings much finer information on the time evolution. Their approach extracts numerous features

from the signal leading to 517 feature vector representing the electricity trace. Different classification algorithms were also analyzed showing the best results with random committee approaches. Due to a pretty large database of signals, they could analyze the impact of using different features and types of classifiers.

In the work of Zufferey et al. [10], the objective was to categorize appliances into 6 categories. The system is based on low-cost smart plugs measuring the electricity consumption parameters at low frequency every 10 seconds. k-Nearest Neighbors and Gaussian Mixture Models were compared, showing similar accuracies up to 85%. Interestingly, the raw observations were simply normalized and used directly as features. A continuation of this work was presented by Ridi et al. in [11] where the signature database ACS-F1 was used, increasing the number of categories to 10 and showing a tuned up system performance of 93.8%. In the next Section this database will be presented.

To the best of our knowledge, all these related works have been evaluated on appliance types and brands that were also seen in the training database, i.e., according to so-called *intersession* protocols where the same appliances are producing the training and testing signature materials. In this paper, we investigate the recognition of categories using *unseen appliance* protocols, i.e., where the testing signatures come from new appliances that are not observed in the training set. The task is expected to be more complex due to the extra inter-brand and inter-model appliance variability.

3 ACS-F1 Database

We based our work on the Appliance Consumption Signature Fribourg 1 (ACS-F1) database [12]. This database contains appliance signatures acquired using low-cost smart plugs capturing the electricity parameters at low frequency with a sampling rate of 10^{-1} Hz. A signature is a sequence of raw measurements $O = \{o_1, \dots, o_N\}$ where o_n is a vector of 6 coefficients including real power (W), reactive power (var), RMS current (A), RMS voltage (V), frequency (Hz) and phase of voltage relative to current (φ). The database contains for each appliance two acquisition sessions of one hour. 100 appliances are recorded and spread uniformly into 10 categories: mobile phone chargers, coffee machines, computer workstations with monitor, fridges and freezers, Hi-Fi systems, lamp (CFL), laptops chargers, microwave ovens, printers, and televisions (LCD or LED).

Two protocols are proposed with the database: the *intersession* and *unseen instances* protocols. In the first protocol all the instances of the first session constitute the train set, whereas those of second session are used for testing. With this protocol, all the testing signatures come from appliances already seen in the training phase. At the time of writing this article, the best performances on the *intersession* protocol are reported in [11]. In this work, two classifiers are compared, namely k-NN and GMM systems showing respectively 88% and 93.8% correct category identification.

The second protocol aims at evaluating *unseen instances* configurations as illustrated in Figure 1. The goal is here to classify instances that are not seen



Fig. 1. *Unseen instances* protocol of the ACS-F1 database

before by the classifiers. This protocol also proposes to use a 10-cross fold procedure to smooth the evaluation results. The fold partitions are made available by the providers of the database. This protocol evaluates the system capability of generalizing to new brands or models of appliances.

4 System Description

4.1 Feature Extraction

In our procedure and as proposed in [10, 11], we use as baseline coefficients the raw observation O as part of the features. We analyze here the impact of including information about the dynamics of the signal through the computation of the so-called *delta* and *delta-delta* or acceleration coefficients. These coefficients have been mainly used in speech recognition and have already been successfully used for appliance identification [11]. As explained in [13], the *delta coefficients* are computed with:

$$\Delta o_n = \sum_{w=-W}^W w \times o_{n-w} \quad (1)$$

where K represents the window length. The value $W = 2$ has been retained after some tests, which corresponds to a window of 50 seconds. We then perform a z-normalization of the features, after which the mean is equal to zero and the variance is equal to one. The normalization is mainly useful for classifiers based on distance computation such as k-NN with a side effect of balancing each feature contribution. Our feature sequence $X = \{x_1, \dots, x_N\}$ is therefore constituted of vectors composed of normalized observations and delta coefficients with $x_n = [c_{1n}, \dots, c_{6n}, \Delta c_{1n}, \dots, \Delta c_{6n}]$ and with c_{in} the normalized value of the corresponding o_{ik} observation. In a similar way, we also analyzed the extension of the features including the *acceleration* coefficients that are computed from the *delta* coefficients with:

$$\Delta \Delta o_n = \Delta o_{n+1} - \Delta o_{n-1} \quad (2)$$

In this work, we also evaluate the effect of applying power thresholding, eliminating from the sequence the observations where the value of the active power is below a given threshold T_P . Intuitively, this is related to the fact that appliances are difficult to discriminate when they are consuming a small quantity of energy, e.g. when they are off or in stand-by. After some pre-tests, the threshold T_P is set to $0.5W$.

4.2 Classification

Two machine learning algorithms are analyzed in this work: k-NN and Gaussian Mixture Models (GMM). A k-NN classifier computes the k closest features from the train set and then uses the labels of these features to perform the classification. In our case, we choose the winning class through a simple majority voting on the labels. In case of a tie, the class having the closest points is elected as winner. The normalized observation, *delta* and *acceleration* coefficients are representing different type of information. We then propose here to weight the distance computation with

$$dist(x_{ts}, x_{tr}) = \alpha \times d(c_{ts}, c_{tr}) + (1 - \alpha) \times d(\Delta c_{ts}, \Delta c_{tr}) \quad (3)$$

where d is the euclidean distance, x_{ts} a test feature vector and x_{tr} a train feature vector. The coefficient α is tuned between 0 and 1 to give more or less weight to the *delta* versus the plain coefficients.

A GMM is a parametric probability density function estimating the likelihood $p(x_n|M_j)$ of a feature vector x_n given a category M_j as a weighted sum of Gaussian component densities. The model can be configured with the number of mixtures I . In our configuration, we used GMM with diagonal covariance matrices making the hypothesis of uncorrelated coefficient. This hypothesis is not true in practice but allows to reduce the number of parameters to estimate and to speed up the computations. The model is computed using the classical Expectation-Maximization (EM) algorithm [14]. The initial values of the Gaussian distributions are computed using the k-means algorithm. For testing, the likelihood $p(X|M_j)$ of an observation sequence X given a model M_j is computed by multiplying the local likelihoods $p(x_n|M_j)$ by making the observation independence hypothesis.

5 Result and Discussion

Influence of the Delta Coefficients. We observe that the inclusion of *delta* coefficients is beneficial for both k-NN and GMM models. Accuracy rates increase from 45% to 52.5% for the k-NN model when including the deltas. Similarly, accuracy rates increase from 62% to 69% for the GMM when including the deltas¹. The dynamic information is bringing significant improvement to both systems.

¹ An optimization of k , the number of neighbors and I the number of Gaussian is systematically performed in all reported results.

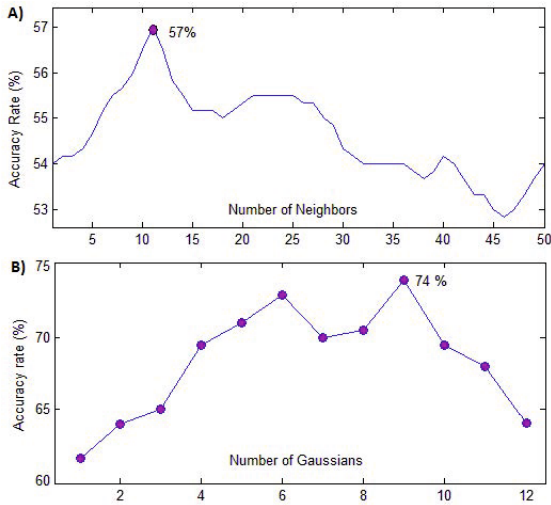


Fig. 2. Accuracy rate trend for a) k-NN varying the number of neighbors ($\alpha = 0.1$, with thresholding), B) GMMs varying the number of Gaussians (using delta and delta-delta coefficients, without thresholding)

Influence of the Thresholding. We observe that eliminating the feature vectors that show an active power below the threshold T_P is beneficial for the k-NN system. Accuracy rates increase from 52.5% to 54.5% when applying the thresholding. This can be intuitively explained considering that close-to-zero power features are present in most signatures, corresponding to stretches of time where the appliances are not used. The training features corresponding to these stretches are independent to the categories and lead to noisy neighbors in the k-NN procedure. Also, as expected, we do not observe a benefit of the thresholding for the GMM models where the zero power stretches bring equivalent score contributions in all categories.

Influence of Weighted Distance Computation. We observe the benefit of applying a weighted distance computation as explained in Eq. 3. The performance improved from 54.5% to 57% with the k-NN system using thresholding and a value of $\alpha = 0.1$. As illustrated in the top part of Figure 2, we obtain this performance for an optimal value of $k = 11$.

Influence of the Delta-Delta Coefficients. Including further the acceleration coefficients, we could achieve an improvement of the GMM system from 69% to 74%. As illustrated on the bottom part of Figure 2, we also observe the effect of tuning the number of mixtures I , with the best performance obtained with $I = 9$ mixtures in the model. A slight improvement of 1.5% is also observed for the k-NN system by including the delta-delta coefficients.

Table 1 provides more details with the confusion matrix for our best GMM system. The categories printer, hifi and lamp are showing the largest error rates. Categories fridge and battery charger are showing the best performances.

Table 1. GMM Confusion Matrix with $I = 9$, without thresholding, using delta and delta-delta coefficients

	Hifi	Television	Battery C.	Coffee M.	Computer	Fridge	Lamp	Laptop	Microwave	Printer
Hifi	.6	.05	0	0	0	.15	.2	0	0	0
Television	0	.7	0	0	.05	.05	.05	.05	0	.1
Battery C.	0	0	.9	0	0	.05	.05	0	0	0
Coffee M.	0	0	0	.75	0	0	0	0	.25	0
Computer	0	.15	0	0	.7	0	0	.15	0	0
Fridge	0	0	0	0	0	.9	0	.05	0	.05
Lamp	0	.1	.1	0	.15	.05	.55	0	.05	0
Laptop	0	.05	0	0	0	0	.05	.85	0	.05
Microwave	0	0	0	.15	0	0	0	0	.85	0
Printer	.05	.05	0	0	0	0	.3	0	0	.6

6 Conclusions

A first objective of this paper was to evaluate the feasibility of equipment identification using simple machine learning algorithms fed by low-frequency electricity consumption measurements. The answer to this question seems positive. We analyzed the performance of different algorithms for the task of identifying unseen appliances. A large database of electrical signatures was used with a total of 200 appliances. Our first conclusion is about the complexity of recognizing *unseen* appliances. When going from a *seen* appliance protocol to an *unseen* appliance protocol using the same database, we observe a drop of performance from 93.8% to 74% correct classification using the best GMM system for both protocols. The *unseen* task still shows acceptable performance but is much more difficult. Improving the performance could probably be reached by increasing the training data set which is still limited in the case of the experiments carried on here. A second conclusion is about the benefit to include dynamic coefficient that are, in our proposal, computed through simple delta and delta-delta coefficients. A third conclusion is about the tuning of some parameters including the weight α used to emphasize the information brought by the delta coefficient in k-NN systems and the number of Gaussians in the GMM model. Finally, as observed in previous works, we can also conclude on the superiority of the GMM over k-NN for signature modeling. Overall, our best accuracy has been raised up to 74% obtained with a GMM model using 9 Gaussians.

As future work, we plan to evaluate the use of state-based models such as HMMs, which should be particularly suitable for electrical signatures that intrinsically show a state nature. HMMs can also be seen as a generalization of GMMs. Comparison with discriminant approaches such as SVM and ANN will also be analyzed.

Acknowledgment. This work was supported by the research grant of the Hasler Foundation project *Green-Mod*, by the HES-SO and by the University of Fribourg in Switzerland.

References

1. Chetty, M., Tran, D., Grinter, R.E.: Getting to green: understanding resource consumption in the home. In: Proc. UbiComp 2008, pp. 242–251 (2008)
2. Rahimi, S., Chan, A.D.C., Goubran, R.A.: Usage monitoring of electrical devices in a smart home. In: Proc. IEEE EMBS 2011 (2011)
3. U.S. Household Electricity Report, E.I.A. (2005), http://www.eia.doe.gov/emeu/reps/enduse/er01_us.html
4. Hart, G.W.: Nonintrusive appliance load monitoring. Proceedings of the IEEE 80, 1870–1891 (1992)
5. Clement, K., Pardon, I., Driesen, J.: Standby power consumption in belgium. In: Proc. EPQU 2007 (2007)
6. Guan, L., Berrill, T., Brown, R.J.: Measurement of standby power for selected electrical appliances in australia. Energy and Buildings 43, 485–490 (2011)
7. Zoha, A., Gluhak, A., Imran, M.A., Rajasegarar, S.: Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey. Sensors, 16838–16866 (2012)
8. Adeel Abbas Zaidi, F.K., Palensky, P.: Load recognition for automated demand response in microgrids. In: Proc. IECON 2010 (2010)
9. Reinhardt, A., Baumann, P., Burgstahler, D., Hollick, M., Chonov, H., Werner, M., Steinmetz, R.: On the accuracy of appliance identification based on distributed load metering data. In: Proc. SustainIT 2012 (2012)
10. Zufferey, D., Gisler, C., Khaled, O.A., Hennebert, J.: Machine learning approaches for electric appliance classification. In: Proc. ISSPA 2012 (2012)
11. Ridi, A., Gisler, C., Hennebert, J.: Automatic identification of electrical appliances using smart plugs. In: Proc. Wosspa 2013 (2013)
12. Gisler, C., Ridi, A., Zufferey, D., Khaled, O.A., Hennebert, J.: Appliance consumption signature database and recognition test protocols. In: Proc. Wosspa 2013 (2013)
13. Hennebert, J.: Hidden Markov models and artificial neural networks for speech and speaker recognition. PhD thesis (1998)
14. Dempster, A., Laird, N., Rubin, D.: Maximum likelihood from incomplete data via the em algorithm. Journal of Royal Statistical Society 39, 1–38 (1977)