AUTOMATIC IDENTIFICATION OF ELECTRICAL APPLIANCES USING SMART PLUGS

Antonio Ridi, Christophe Gisler, Jean Hennebert

University of Fribourg, Department of Informatics
Boulevard de Pérolles 90, CH-1700 Fribourg, Switzerland

University of Applied Sciences Western Switzerland, HES-SO
EIA-FR, Boulevard de Pérolles 80, CH-1705 Fribourg, Switzerland

ABSTRACT

We report on the evaluation of signal processing and classification algorithms to automatically recognize electric appliances. The system is based on low-cost smart-plugs measuring periodically the electricity values and producing time series of measurements that are specific to the appliance consumptions. In a similar way as for biometric applications, such electric signatures can be used to identify the type of appliance in use. In this paper, we propose to use dynamic features based on time derivative and time second derivative features and we compare different classification algorithms including K-Nearest Neighbor and Gaussian Mixture Models. We use the recently recorded electric signature database ACS-F1 and its intersession protocol to evaluate our algorithm propositions. The best combination of features and classifiers shows 93.6% accuracy.

1. INTRODUCTION

We are interested in this paper by the automatic identification of household appliances from their electric consumption signatures. A first application would be to add information to the electricity bill which is nowadays most of the time a blind measurement at the meter of the house. Another application would also be in the optimization of load shedding knowing which appliance can be switched on or off according to global or local electric production. Finally, other type of applications could also emerge such as indirect elderly surveillance or intrusion detection.

From a general perspective, appliance category identification is a challenging task. First, there is a large number of categories with potential overlaps, such as, for example, laptops and tablets. Moreover, there is a large variety among appliances belonging to the same category due to their different functioning or technical differences existing amongst brands. Generally speaking, appliance recognition should guarantee a certain capacity of generalization without over fitting on a specific set of appliances.

In our research work, we are interested in the use of low-cost devices able to measure the electric power consumption at low frequency, in our case 10^{-1} Hz. Such devices are often called smart-plugs as they can be easily installed in front of or as a replacement of standard electric plugs. The devices are producing time series of measurements that are specific to the appliances and that we call here *electric signatures*.

In a previous work, we have shown through preliminary experiences that data-driven machine learning algorithms are providing promising appliance classification performances using electric signatures [1]. We are reporting here on several improvement of our algorithms and on the use of an evaluation database that was recently released [2]. This database, called Appliance Consumption Signature–Fribourg 1 or ACS-F1, contains two acquisition sessions of one hour for 100 appliances spread uniformly into 10 categories. The electrical consumption is measured at low frequency, every 10 seconds including the 6 measurements as described before. The ACS-F1 database comes also with two evaluation protocols that allows the scientific community to provide comparable results.

The paper is organized as follows. An overview of the related work is provided in Section 2. We give in Section 3 details about the algorithm proposals and in Section 4 an overview on the used protocol. In Section 5, we present and discuss the results of the appliance identification task. Finally, in Section 6, we conclude and describe future works.

2. RELATED WORKS

Several research teams have started working in the domain of electric load analysis. A main application is the so-called *energy disaggregation* aiming at identifying which devices are being used, from only the whole-home electric signal. This signal could typically be measured at the smart meter of the home and for this reason, the approach is also named *non-intrusive load monitoring*. A short review of energy disaggregation approaches is provided in [3]. In this direction, some works have concentrated on the high frequency analysis (MHz) of transient electricity signals appearing when appliances are switched on or off [4, 5]. While identification results are promising, such transient approaches have the drawback to require costly high frequency analyzers. Other teams are working

978-1-4673-5540-7/13/$31.00 ©2013 IEEE

on disaggregation using medium range frequency analysis of the electricity signal (up to 15 kHz). In this direction, several databases have been recently made available such as the Reference Energy Disaggregation Dataset (REDD) and the Building Level Fully labeled Electricity Disaggregation Dataset (BLUED) [3, 6].

Another approach is to use a network of distributed home energy aware sensors, typically embedded in low-cost smart-plugs able to measure the electricity consumption in a range of 0.1 to 10 Hz. In this direction, the Tracebase database, a collection of about 1'000 single appliance electric signatures has been recorded [7]. Another database, the Appliance Consumption Signature–Fribourg 1 or ACS-F1 is also available including two acquisition sessions of one hour for 100 appliances spread uniformly into 10 categories [2]. This database is used for the evaluations carried on in this paper.

The difficulty of appliance identification is also depending to the functioning of the considered appliances. Appliances can be divided in two categories, namely the *two-state appliances*, characterized by their on/off state (i.e. lamp), and the *multi-state appliances*, where a binary representation is inadequate (i.e. coffee machine, fridge, microwave) [8]. Some authors include another category, the *continuously variable appliances*, which have an infinite number of states [9]. Other authors propose a different categorization, such as *thermostatically controlled appliances* (i.e. fridge), *fixed operation appliances* (i.e. coffee machine) and *usage dependent appliances* (i.e. microwave), taking into account the human-machine interaction [10].

Various identification algorithms have been used to analyze the consumptions of electrical devices: K-Nearest Neighbors, Decision Trees, Naïve Bayes, Bayesian Networks, Gaussian Mixture Models [1, 7, 11, 12]. Modelling techniques able to handle the temporal and state-based nature of some signals have also been proposed such as Hidden Markov Models and Factorial Hidden Markov Models (FHMMs) [3, 5, 10, 13].

3. ALGORITHM PROPOSALS

The electricity sensors output a sequence of observations $O = \{o_1, \ldots, o_N\}$ with o_n a vector of components corresponding to the measurements done with a periodicity, in our case 10 s. The acquired data includes the real power (W), the reactive power (var), the RMS current (A), the RMS voltage (V), the electric frequency (Hz) and the phase (φ). Some of the values are dependent to other values according to electric principles. However, in our case, we have kept all the values to make our implementation independent to the specificities of the sensors that can expose some or all of the listed values.

3.1. Pre-processing

From the sequence of observations O, we compute a feature vector sequence X by injecting coefficients about the dynamics of the signal and by performing a normalization. The dynamic information is computed with the so-called *delta* and *delta-delta* coefficients. These coefficients are providing information about the temporal dynamics of the signal corresponding to variation and speed of variation that could potentially be useful for the appliance identification. Such coefficients are successfully used in speech recognition. They have also been proposed in the context of energy signal processing [14]. The *delta coefficients* are computed as explained in [15]:

$$\Delta o_n = \sum_{k=-K}^{K} k * o_{n-k} \tag{1}$$

where K determines the window length. We use here $K = 2$, corresponding to a window of 50 seconds. The *delta-delta* coefficients are computed in a similar way from the *delta* coefficients:

$$\Delta\Delta o_n = \Delta o_{n+1} - \Delta o_{n-1} \tag{2}$$

The feature vector is then composed of the original observation completed with the delta and delta-delta coefficients with $x_n = (o_n, \Delta o_n, \Delta\Delta o_n)$. In the next Section we discuss the classification results using both the original feature vector and the feature vector populated with the dynamic coefficients. Consequently, we will be able to directly evaluate the influence of the delta and delta-delta coefficients on the results of the classification algorithms.

The features are also z-normalized using the mean (μ) and sigma (σ) vectors computed on all the train data. The normalized feature vector x_n^z is computed with:

$$x_n^z = \frac{x_n - \mu}{\sigma} \tag{3}$$

As consequence of this operation, the normalized feature vector has zero mean and unit variance. Such normalization is usually applied for distance based classifiers in order to balance the score contribution of each coefficients in the feature vector. For sake of simplification in the rest of this paper, we use the notation x_n for the normalized values. The feature vector x_n is a $D \times 1$ vector with $D = 6$ or, when the dynamic coefficients are included, $D = 18$.

3.2. Gaussian Mixture Models

A GMM is a parametric probability density function computed as a weighted sum of Gaussian component densities:

$$p(x_n|M_i) = \sum_{i=1}^{I} w_i \mathcal{N}(x_n, \mu_i, \Sigma_i) \tag{4}$$

in which I is the number of mixtures, w_i is the weight for mixture i and M_i is the model for mixture i. The mixture weights w_i satisfy the constraint $\sum_{i=1}^{M} w_i = 1$. The Gaussian densities $\mathcal{N}$ are parameterized by a mean $D \times 1$ vector μ_i, and a $D \times D$ covariance matrix Σ_i.

In our case, we make the hypothesis that the features are uncorrelated so that we can simplify the Gaussian density computation using diagonal covariance matrices. We also make the additional hypothesis that the electricity measurements are independent, allowing us to compute the global *likelihood* score S_i for the sequence of feature vectors $X = \{x_1, ..., x_N\}$ given a model M_i with:

$$S_i = p(X|M_i) = \prod_{n=1}^{N} p(x_n|M_i) \quad (5)$$

In order to fit a mixture model to a set of training data, the classical Expectation-Maximization (EM) iterative algorithm is used in our implementation [16, 17]. For a given category i of appliance, all the training data of the category is used to compute the model M_i. The EM being iterative, there are different ways to compute the initial values of the Gaussian distributions $\mathcal{N}(x_n, \mu_i, \Sigma_i)$ such as building random point partitions or using a k-means algorithm. Making the hypothesis of equal priors, the likelihood can be used to perform the classification according to the Bayes rule:

$$P(M_i|X) = \frac{p(X|M_i)P(M_i)}{p(X)} \quad (6)$$

3.3. K-Nearest Neighbors

K-NN classifiers are based on distance computation in the feature space between the test object and the training examples. The object is classified by performing a majoring voting on the class labels of the k nearest neighbors. If $k = 1$, the test sample is simply assigned to the class of its nearest neighbor. The distance computation is usually dependent to the nature of the feature space using criteria such as euclidean, city-block, Chebychev, Minkowski, Mahalanobis, etc. In the case of our experiments, several criteria have been tested to determine the best distance metric. The majority voting principle may cause tie when multiple classes have the same number of nearest points amongst the k nearest neighbors. In this case, the vote is usually attributed to the class having the lowest overall distorsion. In a similar way as for GMMs, we can compute N independent decisions for the feature vector sequence $X = \{x_1, \ldots, x_N\}$. In our implementation, we perform again a majority voting on the N decisions to determine the category associated to the sequence. In other words, we compute a score S_i by accumulating the number of local decisions in favor of category i for a given test sequence.

4. ACS-F1 DATABASE

For our evaluation, we use the recently recorded electric signature database ACS-F1 [2] [1] . The database contains two acquisition sessions of one hour on 100 home appliances spread into 10 categories. Two evaluation protocols are also defined with the database, namely the *intersession* and *unseen instances* protocols. The intersession protocol consists in taking all the instances of the first session as train set and all the instances of the second session for testing. In this case, the train and test sets contains electricity measurements from the same appliances. As illustrated in Figure 1, the algorithms have to classify acquisition sequences coming from appliances already used for the training. The unseen instances protocol assumes that the train and test sets are disjoint in terms of appliance, evaluating the ability of the models to generalize on unseen instances. The intersession protocol should in practice lead to better accuracy than the unseen instance protocol.

In our experiments, we use the intersession protocol of the ACS-F1 database. With this protocol, a total of 100 train sequences are used to train the models with 10 sequences per model. Equivalently, 100 test sequences are available for the evaluation. A given sequence $X = \{x_1, \ldots, x_N\}$ contains one hour of acquisitions with $N = 360$[2]. The results of the intersession protocol are presented here in the form of confusion matrix as well as overall recognition rates.

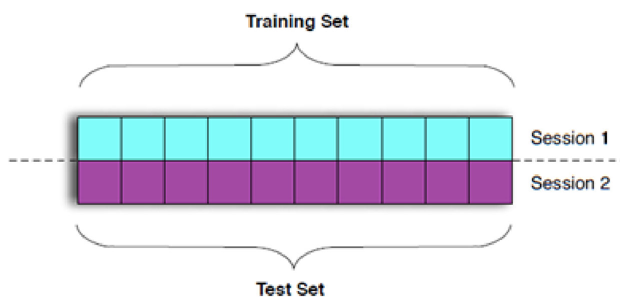

Fig. 1. Intersession protocol.

5. RESULTS AND DISCUSSIONS

We performed an initial set of experiments with GMMs to determine the optimal number of mixtures in the model. In our settings, we used a random partitioning of the training points to initialize the Gaussian mixtures and the EM procedure. The random initial attribution of the data points can make the EM converge to different final parameters of the Gaussian. For this reason, we performed the computation of the models 20 times, randomly changing the initial conditions, and averaging the performances. The evolution of the accuracy of the GMMs as a function of the number of mixtures is depicted in Figure 2. As expected, the accuracy rate is improving when increasing the number of Gaussians, showing the ability of higher order GMMs to better capture the details of the probability density functions. The performances are increasing significantly until 10 mixtures where the improvements start saturating. In this experiment, we also compared the accuracy achieved with the original feature vector (curve in red) with the feature vector populated with the dynamic coefficients (curve in blue). In most situations, the accuracy is higher when the dynamic coefficients are included,

[1] www.wattict.com

[2] We observe some variation around 360, probably due to loss of packets between the sensors and the gateway recording the values.

	Hifi	Television	Battery C.	Coffee M.	Computer	Fridge	Lamp	Laptop	Microwave	Printer
Hifi	92	0	0	0	0	2	6	0	0	0
Television	0	88	0	0	0	0	2	3	0	7
Battery C.	0	0	97	1	0	2	0	0	0	0
Coffee M.	0	0	0	95	0	0	0	0	5	0
Computer	0	2	0	0	90	0	0	8	0	0
Fridge	0	1	0	0	0	98	0	1	0	0
Lamp	0.5	0.5	8	1	0	5	84	0	0	1
Laptop	0	0	0	0	3	0	5	86	0	6
Microwave	0	0	0	1	0	1	0	0	97	1
Printer	0	0	1	0	0	5	1	0	1	92

Table 1. Confusion matrix showing per class accuracy (in percent) for the GMMs with 11 mixtures and with the dynamic coefficients. *Battery C.* stands for *Battery Charger* and *Coffee M.* for *Coffee Machine*.

	Hifi	Television	Battery C.	Coffee M.	Computer	Fridge	Lamp	Laptop	Microwave	Printer
Hifi	90	0	10	0	0	0	0	0	0	0
Television	0	80	0	0	0	0	10	10	0	0
Battery C.	0	0	80	0	0	10	0	0	10	0
Coffee M.	0	0	0	100	0	0	0	0	0	0
Computer	0	0	0	0	100	0	0	0	0	0
Fridge	0	0	0	0	0	100	0	0	0	0
Lamp	0	0	10	0	0	0	90	0	0	0
Laptop	0	0	0	0	10	0	0	70	0	20
Microwave	0	0	0	0	0	0	0	0	100	0
Printer	0	0	10	0	0	0	0	0	0	90

Table 2. Confusion matrix showing per class accuracy (in percent) for the 1-NN classifier using dynamic coefficients. *Battery C.* stands for *Battery Charger* and *Coffee M.* for *Coffee Machine*

showing the benefit of using such coefficients. The best overall performance is obtained using 40 mixtures and the dynamic coefficients with an accuracy of 93.6%. Table 1 provides more details with the confusion matrix for the GMM configuration using 11 mixtures. The category lamp has the worst performance with 84% accuracy and frequent confusions with category Battery Charger. We can also observe that category Computer is also frequently confused with Laptop which somehow makes sense considering the proximity of the internal equipments.

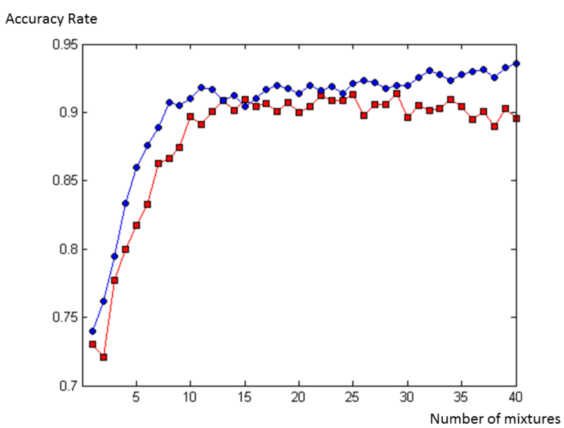

Fig. 2. GMMs accuracy rate evolution when increasing the number of Gaussians using the original feature vector (in red) and using the feature vector populated with the time-coefficients (in blue).

We performed similar tests to explore the different configurations of the k-NN classifier. Our conclusions are the following. The city-block distance computation is leading to better performance than the other distance criteria. Regarding the value of k, the best results are obtained with $k = 1$. Higher values of k seems to introduce more instability of the scores which can intuitively be explained considering the high inter-class overlap leading to more frequent wrongly labeled nearest neighbors. Such results have also been observed in similar context where the 1-NN configuration apparently led to stable and consistent performances [11]. For the configuration with $k = 1$, the inclusion of dynamic coefficients also improves the accuracy from 88% up to 90% which is the best performance obtained with k-NN. The confusion matrix shown in Table 2 provides more details on the per class accuracy for the 1-NN configuration. We observe that the results are more granular as we are not able, in the case of k-NN, to perform multiple runs of the experiments changing initial conditions, which was the case for the GMMs.

6. CONCLUSION

In this paper we report on the evaluation of an appliance recognition task using the recently released electrical signatures database ACS-F1. We followed the intersession evaluation protocol proposed with the database. We compared different configurations of two well-spread classifiers namely GMMs and k-NN. We also evaluated the benefits of injecting dynamic information in the feature extraction step by computing delta and delta-delta coefficients.

As expected, GMMs provide overall better results compared to k-NN, showing the benefit of using models able to estimate the probability density function of the features. The delta and delta-delta coefficients seems also to provide systematic gain that is useful for the classification with GMMs and k-NN.

As future work, we intend to analyze the opportunity to perform classifier fusion which is motivated by the observation that GMMs and k-NN perform different kind of confusions. We also intend to compare the results of this protocol with those of the second reference protocol proposed in [2], evaluating the ability to generalize on new appliances that are unseen in the training set. Finally, we will evaluate other types of feature extraction and feature reduction methodologies.

7. ACKNOWLEDGEMENTS

This work has been supported by the grant Smart Living Green-Mod from the Hasler Foundation in Switzerland.

8. REFERENCES

[1] D. Zufferey, C. Gisler, O. Abou Khaled, and J. Hennebert, "Machine learning approaches for electric appliance classification," in *Proceedings of the 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA'12)*, Montreal, Canada, July 2012, pp. 740–745.

[2] C. Gisler, A. Ridi, D. Zufferey, O. A. Khaled, and J. Hennebert, "Appliance consumption signature database and recognition test protocols," in *Proceedings of the 8th Internation Workshop on Systems, Signal Processing and their Applications (Wosspa '13)*, 2013.

[3] J. Z. Kolter and M. J. Johnson, "Redd: A public data set for energy disaggregation research," in *Proceedings of the ACM Workshop on Data Mining Applications in Sustainability (SustKDD 2011)*, San Diego, USA, Aug. 2011.

[4] S. N. Patel, T. Robertson, J. A. Kientz, M. S. Reynolds, and G. D. Abowd, "At the Flick of a Switch : Detecting and Classifying Unique Electrical Events on the Residential Power Line," in *UbiComp*, 2007, pp. 271–288.

[5] J. B. Durand, L. Bozzi, G. Celeux, and C. Derquenne, "Analyse de courbes de consommation électrique par chaînes de markov cachées," Tech. Rep. 4, 2004.

[6] A. Kyle, A. Filip, D. Bentez, D. Carlson, A. Rowe, and M. Bergs, "Blued: A fully labeled public dataset for event-based non-intrusive load monitoring research," in *Proceedings of the ACM Workshop on Data Mining Applications in Sustainability (SustKDD 2012)*, Beijing, China, Aug. 2012.

[7] 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 *Proceedings of the 2nd IFIP Conference on Sustainable Internet and ICT for Sustainability (SustainIT 2012)*, Pisa, Italy, Oct. 2012.

[8] M. N. V. Perez, "A non-intrusive appliance load monitoring system for identifying kitchen activities," M.S. thesis, Aalto University, Faculty of Electronics Communications and Automation, Espoo, Finland, 2011.

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

[10] F. K. Adeel Abbas Zaidi and P. Palensky, "Load recognition for automated demand response in microgrids," in *Proceedings of the IEEE 36th Annual Conference on IEEE Industrial Electronics Society (IECON 2010)*, Glendale, USA, Nov. 2010.

[11] M. Berges, E. Goldman, H. Scott Matthews, and L. Soibelman, "Learning systems for electric consumption of buildings," in *Proceedings of the 2009 ASCE International Workshop on Computing in Civil Engineering*, Austin, USA, 2009.

[12] A. Reinhardt, D. Burkhardt, M. Zaheer, and R. Steinmetz, "Electric appliance classification based on distributed high resolution current sensing," in *Proceedings of the 7th IEEE International Workshop on Practical Issues in Building Sensor Network Applications (SenseApp 2012)*, Clearwater, USA, Oct. 2012.

[13] J. Z. Kolter and T. Jaakkola, "Approximate inference in additive factorial hmms with application to energy disaggregation," in *Proceedings of the 15th International Conference on Artificial Intelligence and Statistics (AISTATS)*, La Palma, Canary Islands, 2012.

[14] J. Liang, S. K. K. Ng, G. Kendall, and J. W. M. Cheng, "Load signature studypart i: : Basic concept, structure, and methodology," *IEEE Transaction on power delivery*, vol. 25, no. 2, pp. 561–569, 2010.

[15] J. Hennebert, *Hidden Markov models and artificial neural networks for speech and speaker recognition*, Ph.D. thesis, EPFL, Lausanne, October 1998.

[16] A.P. Dempster, N.M. Laird, and D.B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of Royal Statistical Society*, vol. 39, no. 1, pp. 1–38, 1977.

[17] Y. Chen and M. R. Gupta, "Em demystified: An expectation-maximization tutorial," *Electrical Engineering*, vol. 206, 2010.