ACS-F2 – A New Database of Appliance Consumption Signatures

Antonio Ridi*†, Christophe Gisler*†, Jean Hennebert*†
*University of Applied Sciences Western Switzerland,
College of Engineering and Architecture of Fribourg, IcoSys Institute
{antonio.ridi, christophe.gisler, jean.hennebert}@hefr.ch
†University of Fribourg, Department of Informatics, Fribourg, Switzerland
{antonio.ridi, christophe.gisler, jean.hennebert}@unifr.ch

***Abstract*—We present ACS-F2, a new electric consumption signature database acquired from domestic appliances. The scenario of use is appliance identification with emerging applications such as domestic electricity consumption understanding, load shedding management and indirect human activity monitoring. The novelty of our work is to use low-end electricity consumption sensors typically located at the plug. Our approach consists in acquiring signatures at a low frequency, which contrast with high frequency transient analysis approaches that are costlier and have been well studied in former research works. Electrical consumption signatures comprise real power, reactive power, RMS current, RMS voltage, frequency and phase of voltage relative to current. A total of 225 appliances were recorded over two sessions of one hour. The database is balanced with 15 different brands/models spread into 15 categories. Two realistic appliance recognition protocols are proposed and the database is made freely available to the scientific community for the experiment reproducibility. We also report on recognition results following these protocols and using baseline recognition algorithms like k-NN and GMM.**

***Keywords*-Intrusive Load Monitoring (ILM); Appliance Identification; Appliance Recognition**

I. Introduction

As time goes by, energy costs are rising and energy sources falling, while. People and governments need to find ecological solutions. From this perspective, a considerable progress could be made in the management of the various and numerous domestic appliances. Understanding and controlling the appliance electrical consumption and load shedding, depending indirectly on human activity, is a required step. Everyday, new smart meters are installed in buildings and plug-based energy monitoring devices appear on the market. However, existing systems are costly and complicated to install and use. In order to be more adaptive and efficient, they shall integrate automatic recognition of appliances running and potentially moved by people in a house. Such a recognition can be done using machine learning algorithms applied on electric consumption signatures of appliances [14]. Besides ecological aspects, appliance recognition has other utilities, e.g. the detection of defects, the localization of appliances in offices or hospitals, or the recognition of abnormal use of appliances (intrusion detection or elderly surveillance).

In a previous paper, we presented *ACS-F1* [4], a first database of electric consumption signatures acquired from domestic appliances. Low-end smart outlets were used [1] for measuring electric consumption signals at a low frequency (10^{-1} Hz). These devices allow to measure periodically the electricity consumption and communicate it to the computer wirelessly. Available for free to scientific community, this database has been already downloaded several times, pushing us going on with acquisitions. In this paper, we present the new database *ACS-F2*, which is an extension of *ACS-F1*. Two acquisition sessions of one hour were made on 225 domestic appliances divided into 15 categories. ACS-F2 can be freely downloaded for scientists to perform various machine learning experiments and performance comparison on those appliance consumption signatures.

This paper is organized as follows. First, Section II gives an overview of the related work. Section III describes the appliance consumption signature database *ACS-F2* in terms of data acquisition protocol, database content and data format. In Section IV, we propose and discuss new recognition test protocols elaborated to perform appliance recognition tasks and compare results. In Section V, we present the first experiments made and results obtained following these protocols on the database *ACS-F2*. Finally, in Section VI, we conclude the paper.

II. Related Work

Over the last few years, many research works were done on appliance recognition based on the analysis of the electricity consumption. There are two main approaches to the problem: *intrusive* and *non-intrusive load monitoring* [7].

Non-Intrusive Load Monitoring (NILM) consists in measuring the electricity consumption using a smart meter, typically placed at the house entry point. *Non-intrusive* means that no extra equipment is installed in the house. Relying on a single measure, it is also called *one-sensor metering*. Therefore, appliance signatures are superposed and have to be separated for comprehending the contribution of single appliances. This process is called *disaggregation* of the total electricity consumption. *Intrusive Load Monitoring* (ILM) consists in measuring the electricity consumption of one or several appliances separately and directly by means of low-end metering devices. *Intrusive* means that the meter is located in the living space of the home, typically close to appliances that are monitored and consequently to people using them. Scientific publications on NILM approaches outnumber those on ILM. Starting in the 90's, NILM is actually an elder research topic than ILM [16][7]. Databases used for NILM have also been made available for disaggregation tasks, contributing to its momentum [2][8][9]. NILM has also several advantages compared to ILM. Given its single-sensor based nature, installation of NILM systems is easier and data acquisition simpler. ILM is generally based on cheaper meters, however its cost scales linearly with the number of sensors. ILM also presents several advantages over NILM. Finer information is acquired as the sensors are more numerous. Finer details in consumption signatures are also available, which ease the appliance modelling. Typically, low power appliances or appliances in stand-by are not detected with NILM [11], while it is feasible with ILM. Recently, stand-by power became one of the largest source of domestic consumption [6], representing up to 26% of the total energy consumption [15]. Another drawback of NILM is the difficulty to detect appliances with multiple functioning states or showing continuously variable energy use [11].

However, research works in the task of automatic appliance recognition are not as numerous as in other machine learning tasks using similar algorithms. This can be explained by the lack of public data and the difficulty to get enough samples to train a performant appliance recognizer. Therefore we made the database *ACS-F2* freely available to the scientific community. Other existing databases for NILM and ILM are listed. The Reference Energy Disaggregation Dataset (REDD), provided by J. Zico Kolter and M. Johnson [8], contains data from six households, both high-frequency current/voltage waveform data and lower-frequency power data including labeled circuits in the house. The database purpose is to analyse the contribution of single appliances on the aggregated signal. The Building-Level Fully labeled Electricity Disaggregation Dataset (BLUED), provided by K. Anderson *et al.* [9], contains aggregated data from one household. The database is completed with a list of events indicating when appliances changed functioning state. A database of aggregated data from three households is provided by S. Barker *et al.* [2]. A part of the database is completed with appliance data coming from each circuit, and nearly every plug load, measured every few seconds. The *Tracebase* database, built by A. Reinhardt *et al.* [12], contains 31 different appliance categories for a total of more than 1000 traces recorded 24 hours a day. They extracted 517 features from signatures and in the machine learning phase they were able to identify most appliances with a high accuracy. Recently M. Maasoumy *et al.* [10] presented the BERkeley EneRgy Disaggregation Data Set (BERDS) containing measurements of energy and climate data. This data, recorded for a whole year, includes lighting power, HVAC power and receptacle power.

III. Database Description

The Appliance Consumption Signature - Fribourg 2 (ACS-F2) is the second version of our database of electric consumption signatures acquired from domestic appliances and it is publicly available on the site: www.wattict.com. It contains a total of 225 appliances of different brands and/or models. In the ACS-F1 the following categories were present: fridges & freezers, TVs (LCD), Hi-Fi systems (with CD players), laptops, computer stations (with monitors), compact fluorescent lamps (CFL), microwaves, coffee machines, mobile phones (via battery charger), and printers. In the ACS-F2 we moved from 10 to 15 categories, adding incandescent lamps, shavers, fans, kettles and monitors. We incremented as well the number of appliances per category, moving from 10 to 15 appliances per category. We recorded the signals using a low sampling frequency (10^{-1}Hz) for economizing the energy consumption and data storage but hindering the appliance identification task. Every appliance has been acquired over two separated instances (furthermore referred as *sessions*) of one hour.

A. Data Acquisition Protocol

A specific acquiring procedure for every appliance category has been established. The acquisition procedure of appliances in the ACS-F1 database can be found in [4]. Hereafter we list the acquisition procedure of the new five categories:

- *Fans (mechanical)*. They must be in operating mode, standby and turned off. All the possible velocities of the fan must be tested for a variable duration of time.
- *Lamps (incandescent)*. As the fluorescent lamp, they must be turned on and off during the acquisition.
- *Kettles*. They must be turned on and off during the acquisition. The level of water is variable.
- *Monitors*. They must be in operating mode, stand by and turned off. As for computer stations and laptops, different activities must be performed by the users, as surfing on the web or watching films. Changes of the brightness can be as well effectuated.
- *Shavers* (via chargers). They must be in operating mode, stand by and turned off. All the possible velocities of rotating or oscillating blades have to be tested.

B. Data Format

Two different data structures are present in the database: XML and MAT. The XML data structure contains meta-data information, stored in the header, and raw observations, stored in the body. The meta-data contains information about the version of the database ("1" for ACS-F1 and "2" for ACS-F2), the session ("1" or "2"), the author of

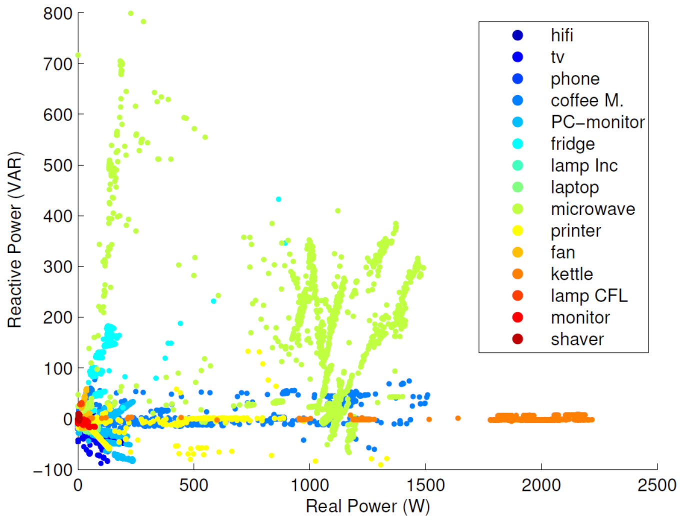

Figure 1. All the samples collected in the ACS-F2 database are plotted on the real power and reactive power (P-Q) plane.

the acquisition, the date, the acquisition place, the sensor device and finally the electrical parameters acquired. The body contains the values of the raw observations and their time-stamp.

The MAT data structure contains only the raw observations with the advantage to be easily readable by MATLAB and Octave software.

IV. Recognition Test Protocols

The ACS-F2 database is publicly available for the scientific community. However, one of the main problem when dealing with public databases is the difference between researchers' analysis procedures that complicates the results comparison. For this reason we present two different protocols in order to successfully compare machine learning algorithms. The protocols here proposed are conceptually the same presented with the ACS-F1 database.

A. Intersession Protocol

The first protocol, called *intersession* protocol, divides the database in two equal parts, one for the training and the other for the test. All the instances of the first session are in the training set and all the instances of the second session are in test set. The classifiers try to identify instances coming by appliances already seen during the training phase. In a real home environment this is equivalent to recognize the appliances after having manually labeled them "on-site". For this protocol it is allowed to used all the signal duration. The confusion matrix and the overall accuracy rate have to be presented.

B. Unseen Appliances Protocol

The second protocol, called *unseen appliances* protocol, performs the 15-fold cross validation on data. The folds are separated in order to have in every fold one appliance (both sessions) per category. The relation between appliances and folds is random. The classifiers try to classify instances coming by appliances never seen during the training phase. In a real home environment this is equivalent to use a priori information about appliance categories without knowing the labels of the appliances "on-site". For this protocol it is allowed to used all the signal duration. The confusion matrix and the overall accuracy rate have to be presented.

V. Recognition Experiments

A. Data Representation

In many cases appliances can be successfully discriminated considering simple electrical features, as real power (P) and reactive power (Q) [5]. In the P-Q plane appliances tend to form clusters depending on their circuital characteristics. Multiple clusters can be expected for a finite state-based machine depending on the differences between states. In Figure 1 all the samples contained in the ACS-F2 database have been plotted on the P-Q plane. A lot of points are concentrated near the (0,0) point, principally because of the stand-by states, off states and the charged states of battery devices.

The electrical appliances can have a resistive, capacitive or inductive behavior, depending on the phase of the voltage to current (φ) parameter. This parameter represents the shifting of the alternating voltage to the alternating current. P and Q can be represented on the imaginary plane with P on the real axis. The complex power S can be computed as $S = P + jQ$ (being j the imaginary unity) and φ is the argument of S. Appliances with heating components, as kettles, have a resistive behavior. Ideally they have φ equal to zero, so they tend to form clusters long the x-axis. Induction motors and fluorescent lamps have an inductive behavior. In this case φ is positive so they tend to form clusters over the x-axis. Rechargeable batteries and transformers have a capacitive behavior. In this case φ is negative so they tend to form clusters under the x-axis.

B. Pre-processing

An individual signature in the database can be represented by $O = \{o_1, \ldots, o_n, \ldots, o_N\}$ where N is the number of samples contained in the time series.

We extracted information about the signatures dynamic evolution through the computation of *velocity* and *acceleration* coefficients. These two parameters are also called respectively *delta* and *delta-delta* coefficients. In a work on the previous database version we demonstrated the usefulness of dynamic coefficients [13]. We computed the *velocity* coefficients as:

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

where W is the window length. We chose $W = 2$, corresponding to a window of 50 seconds, after some tests. We computed the *acceleration* coefficients as:

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

In such a way we are able to artificially increment the feature space by three times. We added the *delta* and *delta-delta* coefficients to the observations. The individual sample o_n is extended as follows: $o_n = \{o_n, \Delta o_n, \Delta\Delta o_n\}$

We normalize the features using the z-normalization. The mean equal to zero and the variance equal to one.

$$x_{kn} = \frac{o_{kn} - \mu_k}{\sigma_k} \quad (3)$$

where μ_k and σ_k are respectively the mean and variance vectors computed using all the signatures in the training set. In our test we will used the normalized signatures incremented by the *delta* and *delta-delta* coefficients.

C. Classification Methods

For the appliance identification task we applied two well know machine learning algorithms, namely k-NN and GMM. The k-NN is a non-parametric method giving as output the class membership for every sample in the test set. The majority voting algorithm is used for deciding the winner class: the sample is assigned to the most common class among its neighbors. In the case of tie, the label of the closest group of neighbors is selected as winner. In this work the distances are computed in the euclidean space. Generally, large values of k help in reducing the effect of noise on the classification. As a backdrop, when increasing k the boundaries of classes are less distinct.

GMM is a probabilistic model where mixture of Gaussians represents the probability distribution of observations. Depending on the desired fitting quality we can use a different number of Gaussians I. However the computational costs are proportionally to the number of Gaussians. We used a diagonal covariance matrix making the hypothesis of uncorrelated coefficients for reducing the computational costs. The model parameters are computed using the Expectation-Maximization (EM) algorithm [3]. For computing the initial values of the Gaussian distributions we used the k-means algorithm. K-means can converge to a local optimum, so we repeated the k-means 100 times and we chose the solution with the lowest total distance between centroids and points. Making the hypothesis of independence between samples in the same observation we computed the global likelihood $p(X|M_j)$ of a sequence $X = \{x_1, x_2, \dots, x_N\}$ by multiplying the local likelihoods $p(x_n|M_j)$ as reported in:

$$p(X|M_j) = \prod_{n=1}^{N} p(x_n|M_j) \quad (4)$$

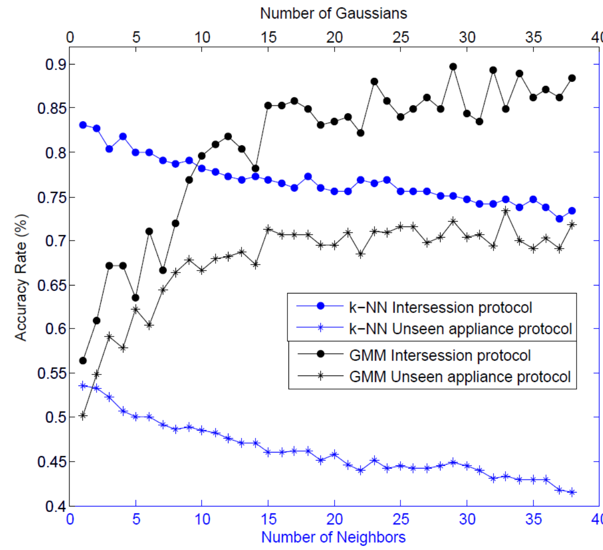

Figure 2. Trend of the accuracy rate using k-NN (blue lines) and GMM (black lines) when increasing respectively the number of neighbors and the number of Gaussians. The filled points are used for the intersession protocol, the asterisks for the unseen instance protocol.

given the category M_j. This procedure is repeated for all the categories and the model yielding the highest global likelihood is selected as winner.

D. Results and Discussion

Hereafter we report and discuss the results obtained using both machine learning algorithms and protocols on the ACS-F2 database. For k-NN we start from $k = 1$ and we increment progressively the number of neighbors. In Figure 2 (blue lines) the accuracy rate when varying the number of neighbors is depicted. The line with filled points represents the intersession protocol, while the line with asterisks represents the unseen instances protocol. As expected, the first protocol performs better than the second. The difference in performances between the two protocols is about 30% and the gap remains approximately constant while increasing the number of neighbors. Interestingly using both protocols the best results are obtained with $k = 1$. Increasing the number of neighbors degrades the system performances. This can be explained considering that close-to-zero power features are present in most signatures and they correspond to stretches of time where the appliances are not used, as stand-by or off state. Training features corresponding to these stretches are independent from appliance categories, leading noisy neighbors in the k-NN procedure. State-base modelling on the contrary is not influenced by those zero power stretches as their scores impact equally every category model.

For the *intersession* protocol the best overall accuracy rate is 83.11% when using k-NN. The confusion matrix is reported in Table I. With the exception of Mobile phone and Shaver, all the categories attain over 73% in terms

Table I
INTERSESSION PROTOCOL USING K-NN WITH $k = 1$

	Hi-Fi	Television	Mobile P.	Coffee M.	Computer	Fridge	Lamp Inc.	Laptop	Oven	Printer	Fan	Kettle	Lamp CFL	Monitor	Shaver
Hi-Fi	.87	0	.07	0	0	0	0	0	0	.07	0	0	0	0	0
Television	0	.87	0	.07	.07	0	0	0	0	0	0	0	0	0	0
Mobile P.	0	0	.6	0	0	0	.07	0	0	0	.07	.07	0	0	.2
Coffee M.	0	.07	0	.8	0	0	0	0	.07	0	0	.07	0	0	0
Computer	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Fridge	0	.07	0	0	0	.93	0	0	0	0	0	0	0	0	0
Lamp Inc.	0	0	0	0	0	0	.87	0	0	0	.07	.07	0	0	0
Laptop	0	0	0	0	0	0	0	.73	0	.13	.07	0	0	.07	0
Oven	0	0	0	0	0	0	0	0	.87	0	0	.13	0	0	0
Printer	0	0	.07	0	0	0	0	0	0	.73	.07	0	0	.07	.07
Fan	0	0	0	0	0	0	.13	0	0	0	.87	0	0	0	0
Kettle	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Lamp CFL	0	0	0	0	0	0	0	0	0	.07	0	0	.8	.07	.07
Monitor	0	0	0	0	0	0	0	.07	0	0	0	0	0	.93	0
Shaver	0	0	.13	0	0	0	.07	0	0	0	0	.2	0	0	.6

Table II
UNSEEN APPLIANCE PROTOCOL USING K-NN WITH $k = 1$

	Hi-Fi	Television	Mobile P.	Coffee M.	Computer	Fridge	Lamp Inc.	Laptop	Oven	Printer	Fan	Kettle	Lamp CFL	Monitor	Shaver
Hi-Fi	.67	0	.07	0	0	0	.03	0	.03	.07	.07	0	.07	0	0
Television	0	.33	0	0	.4	0	0	0	.13	0	0	0	.07	.07	0
Mobile P.	0	0	.63	0	0	0	.07	0	0	0	.03	.13	.03	0	.1
Coffee M.	.07	.03	0	.53	0	0	0	0	.03	.07	0	.13	0	.13	0
Computer	0	.2	0	0	.57	.07	0	.13	0	0	0	0	0	.03	0
Fridge	.03	.03	0	0	0	.67	.07	0	.07	0	0	.07	0	.07	0
Lamp Inc.	.1	0	0	0	.03	.03	.67	0	0	0	.07	.03	0	.07	0
Laptop	0	.03	0	0	.13	0	.03	.13	0	.13	.17	0	.03	.33	0
Oven	.07	.07	0	0	0	.07	0	0	.67	0	0	.1	0	0	.03
Printer	.07	.1	.1	.1	.03	.03	0	.07	0	.13	.1	0	.07	.13	.07
Fan	.03	.07	0	0	0	0	.1	.03	.03	0	.5	.1	.13	0	0
Kettle	0	0	0	0	0	0	0	0	0	0	.03	.97	0	0	0
Lamp CFL	.13	.03	0	0	0	0	0	0	0	0	0	.03	.67	.1	.03
Monitor	0	0	0	0	0	.03	.03	.13	0	.03	0	0	.07	.7	0
Shaver	.1	.03	.27	0	0	.07	.03	0	0	0	0	.23	.07	0	.2

Table III
INTERSESSION PROTOCOL USING GMM WITH $I = 29$

	Hi-Fi	Television	Mobile P.	Coffee M.	Computer	Fridge	Lamp Inc.	Laptop	Oven	Printer	Fan	Kettle	Lamp CFL	Monitor	Shaver
Hi-Fi	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Television	0	.93	0	0	.07	0	0	0	0	0	0	0	0	0	0
Mobile P.	0	0	.87	0	0	0	0	0	0	0	0	0	0	0	.13
Coffee M.	0	0	0	.67	0	0	0	0	.2	.13	0	0	0	0	0
Computer	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Fridge	0	0	0	0	0	.93	0	0	.07	0	0	0	0	0	0
Lamp Inc.	.07	0	0	0	0	0	.93	0	0	0	0	0	0	0	0
Laptop	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Oven	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Printer	0	0	0	0	0	0	0	.07	.07	.73	0	0	0	0	.13
Fan	0	.07	0	0	0	0	0	0	0	0	.93	0	0	0	0
Kettle	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Lamp CFL	0	.07	0	0	0	.07	0	0	0	0	0	0	.87	0	0
Monitor	0	0	0	0	0	0	0	.2	0	.13	0	0	0	.67	0
Shaver	.07	0	0	0	0	0	0	0	0	0	0	0	0	0	.93

Table IV
UNSEEN APPLIANCE PROTOCOL USING GMM WITH $I = 33$

	Hi-Fi	Television	Mobile P.	Coffee M.	Computer	Fridge	Lamp Inc.	Laptop	Oven	Printer	Fan	Kettle	Lamp CFL	Monitor	Shaver
Hi-Fi	.57	0	.07	0	0	.03	0	.03	0	.03	.17	0	.03	0	.07
Television	0	.73	0	0	.1	.03	0	.1	0	0	0	0	0	.03	0
Mobile P.	0	0	.73	0	0	0	0	0	0	0	0	0	0	0	.27
Coffee M.	0	0	0	.67	0	0	0	0	.07	.27	0	0	0	0	0
Computer	0	.13	0	0	.77	.03	0	.07	0	0	0	0	0	0	0
Fridge	0	0	0	0	.07	.8	0	.07	0	0	.07	0	0	0	0
Lamp Inc.	.07	.03	0	0	0	0	.77	0	0	.03	.1	0	0	0	0
Laptop	0	0	0	0	.07	0	0	.73	0	0	.03	0	.07	.1	0
Oven	0	0	0	.07	0	.1	0	0	.83	0	0	0	0	0	0
Printer	0	0	.07	0	.03	.1	0	.03	.03	.57	.07	0	0	.07	.03
Fan	0	.07	0	0	0	0	0	0	0	0	.93	0	0	0	0
Kettle	0	0	0	.03	0	0	0	0	.07	0	0	.9	0	0	0
Lamp CFL	0	0	0	0	0	0	0	0	0	0	.17	0	.77	0	.07
Monitor	0	0	0	0	0	0	0	.27	0	.07	0	0	0	.67	0
Shaver	.13	0	.2	0	0	0	.03	0	0	0	0	0	.07	0	.57

of accuracy rate. The greatest misclassification errors are between Mobile phone - Shaver and Kettle - Shaver. Kettle and Computer station obtain an accuracy rate of 100%.

For the *unseen appliance* protocol the best overall accuracy rate is 53.56% when using k-NN. The confusion matrix is reported in Table II. Kettle is the only category performing well (97%). On the other hand Shaver, Laptop, Printer and Television provide poor performances (less than 33%). The greatest misclassification errors are between Computer - Television, Laptop - Monitor and Mobile phone - Shaver.

For GMM we start from $I = 1$ and we increment progressively the number of Gaussians. In Figure 2 (black lines) the trend of the accuracy rate when varying the number of Gaussians is depicted. The line with filled points represents the intersession protocol, while the line with asterisks represents the unseen instances protocol. As expected, the first protocol performs better than the second. The difference between the accuracy rates of the two protocols increases when increasing the number of Gaussians. In both cases the performances tend to saturate after approximately 15 mixtures.

For the *intersession* protocol the best accuracy rate is 89.78% when using GMM with $I = 29$. The confusion matrix is reported in Table III. With the exception of Monitor, Coffee machine and Printer, all the categories are over 87% in terms of accuracy rate. The greatest misclassification errors are between Laptop - Monitor and Coffee machine - Oven. Hi-Fi, Computer stations, Laptop, Oven and Kettle obtained an accuracy rate of 100%.

For the *unseen appliance* protocol the overall accuracy rate is 73.33% when using GMM with $I = 33$. The confusion matrix is reported in Table IV. Kettle and Fan achieve more than 90%. Shaver, Hi-Fi and Printer provide poor performances (57%). The greatest misclassification errors are between Printer - Coffee machine, Laptop - Monitor and Mobile phone - Shaver.

VI. Conclusion

In this paper we present the ACS-F2 database for the appliance identification. The database contains electrical consumption signatures acquired at low frequency 10^{-1} Hz. The database is an extension of the previous ACS-F1. The database contains 225 appliances uniformly spread among 15 categories and every appliance is recorder two times for one hour. Every category contains 15 appliances of different brands and/or models. The database is freely available to the scientific community for the experiment reproducibility and algorithm comparison.

We propose two protocols: the intersession and the unseen appliance protocol. The protocols principally indicate how separate train and test set. Specifically the first protocol tries to identify instances coming by the appliances already seen in the training phase. The second protocol tries to identify instances coming by the appliances never seen in the training phase.

We also report on recognition results following these protocols and using machine learning algorithms like k-NN and GMM. As best case we obtained an accuracy rate of 83.11% and 89.78% using k-NN and GMM respectively on the intersession protocol. For the unseen appliance protocol we obtained 53.56% and 73.33% using k-NN and GMM respectively. We reported the confusion matrix for every best case. As expected, the unseen appliance protocol is harder compared to the intersession protocol and GMM algorithm performs better than k-NN in terms of accuracy rate for both protocols.

Through the analysis of confusion matrices we observe that Kettle is always well discriminated (90% in the worst case) while Printer provide poor performances (73% in the best case). Some misclassifications appear recurrent, as between Shaver - Mobile phone, Printer - Monitor and Laptop - Monitor.

VII. Acknowledgments

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] Energy Optimizers Limited, PLOGG Wireless Energy Management. http://www.plogg.co.uk.

[2] S. Barker, A. Mishra, D. Irwin, E. Cecchet, P. Shenoy, and J. Albrecht. Smart *: An Open Data Set and Tools for Enabling Research in Sustainable Homes. In *Proceedings of the 2nd ACM Workshop on Data Mining Applications in Sustainability (SustKDD)*, 2012.

[3] A. Dempster, N. Laird, and D. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of Royal Statistical Society*, 39(1):1–38, 1977.

[4] C. Gisler, A. Ridi, D. Zufferey, O. Abou Khaled, and J. Hennebert. Appliance consumption signature database and recognition test protocols. In *Proceedings of the 8th International Workshop on Systems, Signal Processing and their Applications (WOSSPA)*, pages 336–341, 2013.

[5] H. Gonçalves, A. Ocneanu, and M. Bergés. Unsupervised disaggregation of appliances using aggregated consumption data. In *SustKDD Workshop on Data Mining Applications in Sustainability*, 2011.

[6] L. Harrington, J. Brown, S. Holt, A. Meier, B. Nordman, and M. Ellis. Standby Energy : building a coherent international policy framework – moving to the next level. *Summer Study Proceedings, European Council for an Energy Efficient Economy*, 3:1285–1296, 2007.

[7] G. Hart. Nonintrusive Appliance Load Monitoring. *Proceedings of the IEEE*, 80(12):1870–1891, 1992.

[8] J. Kolter and M. Johnson. REDD : A Public Data Set for Energy Disaggregation Research. In *Proceedings of the ACM Workshop on Data Mining Applications in Sustainability (SustKDD)*, 2011.

[9] A. Kyle, A. Filip, D. Benítez, D. Carlson, A. Rowe, and M. Bergés. BLUED : A Fully Labeled Public Dataset for Event-Based Non-Intrusive Load Monitoring Research. In *Proceedings of the 2nd ACM Workshop on Data Mining Applications in Sustainability (SustKDD)*, 2012.

[10] M. Maasoumy, B. Sanandaji, K. Poolla, and A. S. Vincentelli. Berds - berkeley energy disaggregation data set. In *Proceedings of the Workshop on Big Learning at the Conference on Neural Information Processing Systems (NIPS 2013)*, 2014.

[11] A. Marchiori, D. Hakkarinen, Q. Han, and L. Earle. Circuit-Level Load Monitoring for Household Energy Management. *IEEE Pervasive Computing*, 10(1):40–48, 2011.

[12] 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.

[13] A. Ridi, C. Gisler, and J. Hennebert. Automatic identification of electrical appliances using smart plugs. In *Proceedings of the 8th Internation Workshop on Systems, Signal Processing and their Applications (WOSSPA)*, pages 301–305, 2013.

[14] A. Ridi, C. Gisler, and J. Hennebert. A survey on intrusive load monitoring for appliance recognition. In *Proceedings of the 22nd International Conference on Pattern Recognition (ICPR), to appear*, 2014.

[15] J. Ross and L. Berkeley. Whole-House Measurements of Standby Power Consumption. In *Proceedings of the 2nd International Conference on Energy Efficiency in Household Appliances (EEDAL)*, 2000.

[16] F. Sultanem. Using Appliance Signatures for Monitoring Residential Loads ot Meter Panel Level. *IEEE Transactions on Power Delivery*, 6(4):1380–1385, 1991.