TIME-FREQUENCY CHARACTERISATION FOR ELECTRIC LOAD MONITORING

Mabrouka El Guedri [1, 2], *Guy D'Urso* [3], *Christian Lajaunie* [2], *Gilles Fleury* [1]

[(1)] Department of Signal Processing & Electronics Systems – Supélec, 91192 Gif sur Yvette, France
[(2)] Centre de Géostatistique – Ecole Nationale des Mines de Paris, 75006 Paris, France
[(3)] Electricité de France – R&D, 4 Qui Watier, 78401 Chatou, France
mabrouka.elguedri@supelec.fr, gilles.fleury@supelec.fr, christian.lajaunie@ensmp.fr

ABSTRACT

Electric utilities and consumers are increasingly interested in energy monitoring for economic and environmental reasons. A non-intrusive solution may rely on information extracted from the electric consumption measured at a centralized part of a distribution network. The problem at hands consists in the separation of the electric load into its major components. This problem of source separation from one sensor is quite tractable under certain conditions.

In this work, the focus is made on the most consuming household appliance in France: the space-heating. It is a sum of an unknown number of pseudo-periodic signals embedded in the global active power. An unsupervised algorithm to determine the space-heating schedule from the global consumption based on the interpretation of the space-heating signature in the time-frequency domain is proposed. The proposed method conjoins a time-frequency detector and a frequent itemsets extraction. First results on real data are quite satisfying.

1. INTRODUCTION AND BACKGROUND

The electric power industry and consumers recently face many challenges such as energy saving and greenhouse gas emissions reducing. Accurate and reliable information about the nature and the state of the electric systems will undoubtedly be helpful to meet these challenges. Actually, a good knowledge of the electric load and the targeted appliances help consumers understanding their bills and better control their consumption

A non-intrusive and economical solution may rely on information extracted from the electric consumption measured at a centralized easily accessible part of a distribution network. Non-intrusive electric load monitoring has been subject to several approaches over the last twenty years. General overviews can be found in [1, 2 , 3]. The available solutions require measurements of the active and the reactive power, which carry out the finger-prints of the electric appliances. They are mostly made up of three steps. Event detection determines the appliances operating schedule. Load identification uses steady state powers and transient patterns, if available, to recognize the elementary components.

In this paper, the focus is made on the most consuming end-use in France: the space-heating. The observed signal (the active power) is a sum of pseudo-periodic square waves characterized by slowly time-varying duty cycles. The problem statement is given in section 2. In section 3, the detection of locally-stationary signals embedded in a non-stationary observation in the time-frequency (T-F) domain is explained. A novel procedure of interpretation of the time-frequency content is investigated. Experimental tests have been carried out on real data provided by Electricité de France.

2. PROBLEM STATEMENT

In this paper, we are interested in the breakdown of the whole electric consumption given the active power $y(t)$ available at the electricity meter (sampled at a sampling rate T_e=2s) in France. In fact, this measurement is available and does not need any installation at home customers. Moreover, the recorded measurement could be sent to a centralized platform through radio transmission. Notice that only one measurement is required compared to existent solutions based on three measurements of current and voltages at higher sampling frequency. The electric load decomposition consists in the detection and the estimation of its major components. The observed signal is the sum of an unknown number of loads y_k of the electric appliances as detailed in equation 1,

$$\forall t, y(t) = \sum_{k=1}^{K} y_k(t) + b(t) \quad (1)$$

where b is an additive Gaussian noise.

Classical methods of sources separation [4] are not suited for this problem. Recent techniques [5] developed for audio signals are based on the temporal and spectral parsimony of the elementary signals. In this work, we propose a dedicated method to electric load characterization. An analysis and a characterization of the individual signals are crucial to deal with this problem.

In this paper, the focus is made on the space-heating by convectors for two reasons. Firstly, it is the most consuming household appliance in France (~70% of the national consumption in the residential sector). Secondly, a convector load obviously differs from the other signals. When it is on, it generates periodic square signal with constant period (40s or 80s) and a slowly time-varying duty cycle. **Figure 1** illustrates three signals of three (electronic) convectors operating simultaneously at a customer house. During this study, we analyzed the elementary signals recorded for eleven different convectors during three weeks. This data basis is certainly not sufficient, but necessary for a feasibility study. In fact, gathering such data is a costly and difficult task. Some common properties of the convectors loads have been established thanks to this analysis. Actually, we noted that a convector signal is locally stationary, where the local stationarity means that the covariance function of the studied signal can be considered as the time modulation of a stationary covariance function [6].

Let us underline that more data were gathered during the winter 2008-2009 from two different sites. The active power of each convector (almost thirteen convectors) has been measured. The signals analysis have the same properties of those used during this study.

The detection of space-heating from the active power is formulated as the detection of a mixture of an unknown number of periodic locally stationary signals, where the set of possible values for periodicities is known.

Since the observed signal y is non-stationary, time-frequency representations [6] are suitable tools to characterize its components. In this space, the signal components are described by structures called spectral patterns. Given that targeted signals are piecewise stationary, the appropriate TFR is the *Short Time Fourier Transform* (STFT) [6]. A sample of an electric load, where three convectors are simultaneously in use, and the magnitude of the corresponding STFT, namely the spectrogram, are illustrated by **Figure 2**.

Equidistant horizontal lines correspond to the space-heating signature in the time-frequency domain, because of the periodic property of the convectors signature. The interpretation of the space-heating signature obviously becomes straightforward in this domain. Note that other time-frequency features may also be distinguished.

The interpretation of the time-frequency representation content mostly consists of two processes: the TFR segmentation and the classification of the extracted patterns. Several parametric or non parametric, supervised or unsupervised methods have already been proposed in literature [7, 8].

This work investigates a new unsupervised and semi-parametric procedure, dedicated to the characterization of an unknown mixture of periodic locally-stationary signals embedded in a non-stationary observation. The segmentation of the TFR is tuned with a time-frequency detector based on a statistical model of the TFR. The classification process aims at separating the extracted spectral patterns into two families: the space-heating features and the other devices' ones. This operation is performed by a data-mining method, namely the frequent itemsets extraction. The proposed method is based on priors extracted from convectors measurements with electronic regulation. As for other convection systems, such as mechanics convectors, we will show that our method might be still appropriate since the space-heating load is locally stationary.

The proposed approach to characterize the space-heating load is described in **Figure 3**.

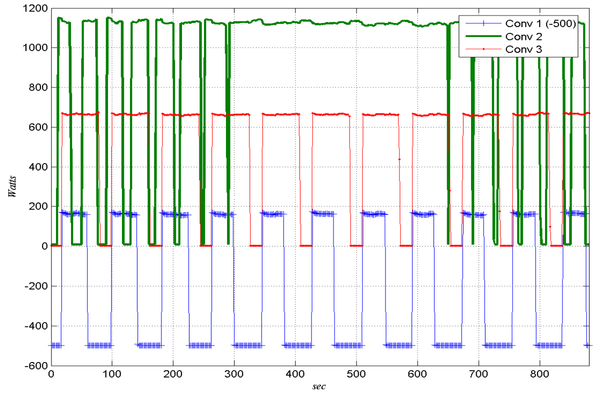

Figure 1 – Active power of three convectors

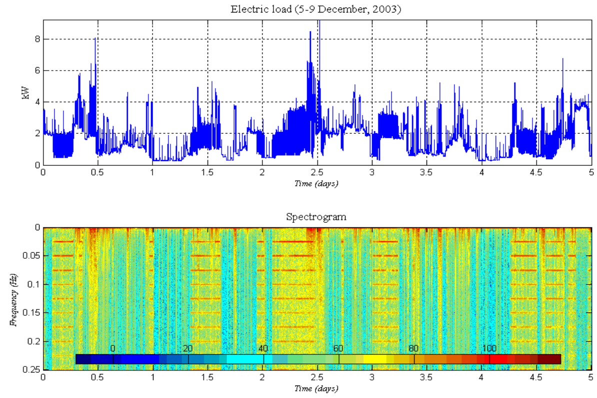

Figure 2 - A weekly electrical load and its spectrogram (total number of convectors: 3)

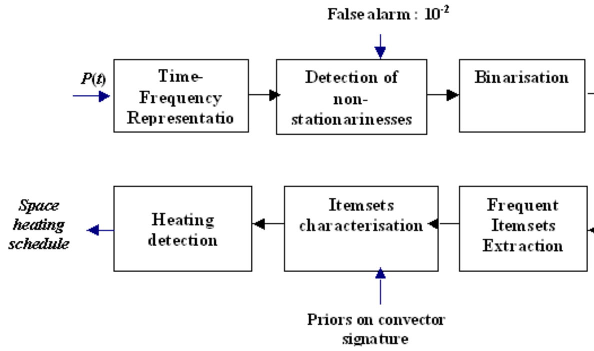

Figure 3 - Space-Heating detection in the time-frequency domain (where P(t) is the measured active power)

3. DETECTION AND EXTRACTION OF PIECEWISE STATIONARY SIGNALS FROM THE SPECTROGRAM

3.1 Spectrogram segmentation based on a series of hypothesis tests

A TFR might be considered in two different ways. In [7, 9] it is considered as an image. The segmentation is performed by active contours and mathematical morphology. These tools do not provide an automatic detection and characterization of the time-frequency features. In [10, 11], the TFR is modelled with a mixture of probability densities.

Its segmentation is based on the statistical properties of a specific time-frequency representation: the spectrogram, which is the magnitude of the *STFT* [6].

As the observation y is a sequence of N samples, each spectrogram coefficient is given as following (h is the analysis window of length $N_f < N$):

$$\rho_y[k,\nu] = \frac{1}{N}\left|\sum_{m=1}^{N} h[m-k]y|m]e^{-2j\pi\frac{m\nu}{Nf}}\right| \qquad (2)$$

The segmentation procedure consists in labelling each T-F coefficient using a statistical model of the chosen TFR. In this work, our purpose is to provide two regions: a region which contains the spectral patterns and a region defined as the background of the TFR. The segmentation task might be performed through a hypothesis test. The method described below has been developed in the GIPSA-lab as part of studies conducted in this lab. The segmentation procedure consists in labelling each T-F coefficient using a statistical model of the chosen TFR. In this work, our purpose is to provide two regions: a region which contains the spectral patterns and a region defined as the background of the TFR. The segmentation task might be performed through a hypothesis test. The method described below has been developed in the GIPSA-lab as part of studies conducted in this lab.

Statistical hypothesis test

The detection of abrupt temporal changes is formulated as following:

H_0: The observed signal r is stationary, possibly added with a deterministic signal. It follows a known probability density p_0, whose parameters are unknown.

H_1: Non-stationary signal of unknown probability density p_1

Lack of information about ruptures makes the definition of an optimal detector not feasible. In [12], the authors propose a recursive detector based on the knowledge of the p_0's family. They proposed a new detector in a particular TFR: the spectrogram, and reformulated the previous hypothesis test in the new space. As the prior density of observations and the cost function are unknown, the decision rule's threshold might be determined for a given false alarm Pfa

Consequently, if $\rho_y[k,\nu] > \eta_{Pfa[\nu]}$, then the alternate hypothesis H_1 is accepted. Otherwise, one accepts the null hypothesis.

Let us now give some definitions and notations. For a given signal *y*, the observation space is $L = \{(k,\nu), k \in K, \nu \in F\}$ where $K \times F$ is the time-frequency domain. For each normalised frequency, we define:

$L_{H0}(\nu) = \{(k,\nu) \in L(\nu) \ / \ p_{\rho_y/H0} = p_0 \text{ and } k \in I_s)\}$

where : $I_s = \arg\max\{card(I), I \subset K, \mathrm{E}(\rho_y[k,\nu]) < \mu_s\}$

μ_s is a threshold chosen a priori.

$L_{H1}(\nu)$ is the complementary subspace of $L_{H0}(\nu)$.

Notice that the decision rule requires the knowledge of the statistical distributions of the T-F coefficients.

Statistical model of the spectrogram coefficients

For better explanation, the statistical test is presented in the particular case of a stationary white Gaussian signal.

H_0^*: Random signal b, stationary, with probability density $N(0,\sigma^2)$.

H_1^*: The alternate hypothesis.

The statistical properties of the spectrogram coefficients under the null hypothesis have to be determined in order to evaluate the detection threshold. For each frequency, $(\rho_b[k,\nu])$ is a mixture of zero mean independent Gaussian random variables. It follows asymptotically a central χ^2 law [13] with 2 degrees of freedom.

Therefore, the spectrogram might be modelled as a mixture of central and non-central chi-square laws [13]. Non-central parameters are proportional to the signal to noise ratio evaluated for the spectrogram coefficients. The estimation of the distributions' parameters helps to make a decision as for the category of each spectrogram coefficient.

Decision Rule

The threshold value is determined for a fixed false alarm Pfa. The decision rule is:

If $\rho_y[k,\nu] > \eta_{Pfa[\nu]}$, the alternate hypothesis H_1 is true, and $\rho_y[k,\nu] \in L_{H1}(\nu)$. Otherwise, the null hypothesis is satisfied and we have: $\rho_y[k,\nu] \in L_{H0}(\nu)$

L_{H1} is the set of all the spectral patterns.

Algorithm

The key idea of this detector is to construct the subspace of interest iteratively. More precisely, let $L^i_{H0}(\nu)$ be the current estimation of $L_{H0}(\nu)$ at iteration *i*, $P^i = card\{L^i_{H0}(\nu)\}$ and $\mu^i[\nu]$ be the empirical mean of $L^i_{H0}(\nu)$. For more details, we refer to these works [12, 13].

The stopping criterion is the stability of L_{H0}. As for initialisation, one affects to each $L_{H0}(\nu)$ the *p*% lowest values of the observed data. The parameter *p* is chosen *a priori*. This procedure is used to detect all the spectral and the temporal changes in the T-F domain. It provides a mapping of the spectral patterns including those associated to the piecewise-stationary periodic signals. The next step aims at automatically extracting these patterns so that one can select the convectors patterns.

3.2 A novel method to extract "some" time-frequency patterns: Frequent Itemsets extraction

The subspace L_{H1} is a time-frequency map of two families of features: the first one is associated to periodic signals, the second one to other signals. Our objective is to extract the first category of patterns. The extraction algorithms of compact subsets [13] are not suited to our problem because we aim at extracting only some itemsets having some specific properties. Actually, the features associated to the convectors have three properties. Each pattern is nearly rectangular. Besides, for a given temporal support, the convectors structures have almost the same size. Finally, whenever the space-heating is "on", the patterns of each convector are equidistant. These priors are used to model the targeted structures and to develop an automatic and unsupervised procedure of extraction. A specific method namely the *Frequent Itemsets Extraction* selects these itemsets. Early works on this method [14] led to the APRIORI algorithm, where the processed data are binary matrix. The rows are called *objects* and the columns are the *attributes*.

Definition: A frequent itemset is an itemset for which the number of objects including this itemset, called support, is greater than a threshold named cut frequency (f_c) [15].

Frequent Itemsets Extraction looks for the frequent itemsets in a large data basis. This procedure is carried on a condensed representation: a *Galois lattice*. The lattice construction is illustrated in **Figure 5** for a non realistic binary matrix (**Figure 4**). The frequent itemsets for a cut frequency $f_c = 2$ are presented in **Figure** 6. The APRIORI algorithm proceeds repeatedly as following:

1srt iteration: evaluate the itemsets' frequencies

k^{th} iteration:

- Frequent itemsets of iteration (k-1) are used to determine candidates for k-frequent itemsets, with a greater frequency than the cut frequency.
- Update data: pruning the non-frequent itemsets.

In our case, time instants are the objects (*transactions*). Frequencies are the *attributes*. A spectral pattern could be considered as an association between an itemset of attributes and the objects (instants) containing these attributes (frequencies). Note that the matrix data must be binarised and transposed for computing reasons. Moreover, a morphological filtering is used to eliminate all the patterns with duration inferior to the minimum duration of a convector. The patterns extraction is then performed by specific software: *MvMiner* [16]. As the space-heating spectrum covers the entire frequential axis, the extraction of the spectral signature of the convectors might be performed by selecting the longer itemset of attributes. As for the cut frequency, it could be defined regarding the minimal duration of a convector's operating t_{ch0} given consumer habits an operating constraints as following: $f_c = \dfrac{t_{ch0}}{T_e\, N_f}$.

The output of this module is a map including mostly the binarised signature of the convectors in the spectrogram. We can then deduce the schedule of this end-use from a daily electrical load.

4. EXPERIMENTAL RESULTS

Real data have been gathered only at a customer house, because collecting real data is a difficult and costly task. The results given herein are not complete yet. We only demonstrate the efficiency of the proposed method on real data in the particular case of electronics convectors. The evaluation of the detector on large data basis for different customers will be performed in future works, using data recorded at tow houses.

Time / *Frequency*	0	512	2x 512	3x 512	4x 512	5x 512	6x 512	7x 512	8x 512
$f_1 = 0$	1	1	1	0	0	0	1	1	1
$f_2 = 0.01$	1	1	1	0	1	1	1	1	1
$f_3 = 0.02$	0	0	0	1	1	1	0	0	1
$f_4 = 0.03$	1	1	1	1	0	0	0	0	0

Figure 4 - An illustration of the processed data

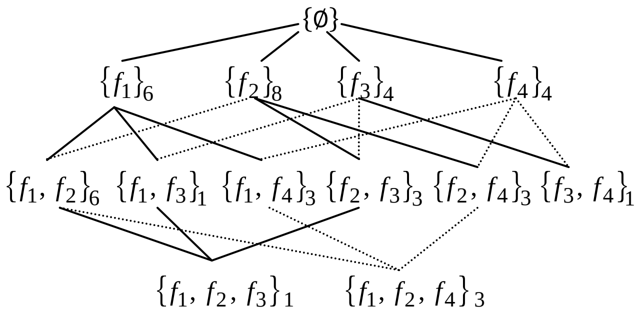

Figure 5 - A condensed representation of the matrix: a lattice

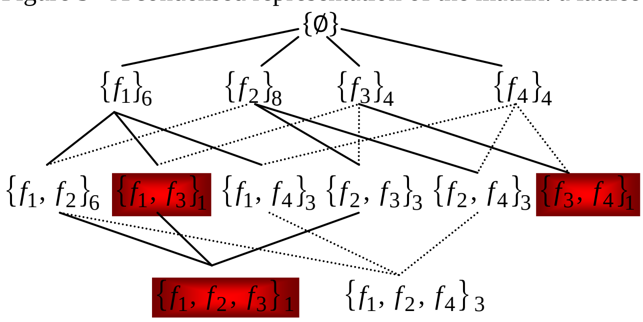

Figure 6 - Frequent itemsets for $f_c = 2$

The frequent itemsets map and the longer frequent itemsets obtained for a weekly electric load (Figure 2) are presented in Figure 7.

Notice that some itemsets are split up resulting in non-detected intervals Z_1, Z_2 and Z_3.

The space-heating detector in the temporal domain (Figure 8) is thereafter obtained given the patterns matrix and a model of a convector. Besides the non-detection due to split up of some itemsets, Figure 8 (black ellipses) illustrates some non-detection due to the cut frequency chosen and the length of extracted itemsets.

A post-processing might be added to the global chain, so that adjacent regions are interpolated. Thereafter, the space-heating detector from the selected itemsets is deduced. The temporal support of non-detection intervals is around 10% of the space-heating support, without any post-processing. We also stress that the chosen method of extraction does not lead to any false alarms. The detector described above is based on the spectral properties of a convector load and should be applicable to other convection system since the corresponding patterns are frequent compared to the other devices ones.

5. CONCLUSIONS AND FUTURE WORKS

In this work, a new method of interpretation of a particular time-frequency representation, namely the spectrogram in the case study of electric load monitoring, is investigated. The method conjoins a time-frequency detector and a data mining method to extract some pertinent patterns.

The T-F detection is based on a statistical model of the spectrogram. An existing detector of non-stationarities is implemented and tested on real data. A model of the convectors given the priors on the temporal and the spectral properties of the electric signature is used. The itemsets extraction method operates on the binarised time-frequency map and provides the frequent itemsets.

Finally, the detection of the periodic signals from the unique mixture is performed by selecting the largest patterns of attributes. The method presented herein is dedicated to the detection and the extraction of periodic locally stationary signals embedded in a non-stationary observation. An application of this algorithm is the estimation of the space-heating operating schedule given the unique measurement of the active power, without any intrusion in the house and any knowledge on the electric installation.

The designed detector has no false alarm, which is an important criterion in the point of view of the customer. The detector would be evaluated on a larger data basis, where each device active power is recorded. If the performances of the proposed detector are satisfying, it would be evaluated in real life through tests on an industrial platform. Other space-heating system would be thereafter studied.

REFERENCES

[1] G. Hart, “Residential energy monitoring and computerized surveillance via utility power flows,” IEEE Technology and Society, pp. 12-16, June 1989.

[2] C. Laughman, K. Lee, R. Cox, S. Shaw, L. Norford and P. Armstrong, “Power Signature Analysis,” IEEE Power and Energy Magazine, pp. 55-63. March 2003.

[3] S. Leeb, A conjoint pattern recognition approach to non-intrusive load monitoring, PhD. Dissertation, Massachusetts Institute of Technology, February 1993.

[4] A. Hyvärinen and P. Pajunen, “Independent component analysis: Algorithms and Applications », Neural Networks, vol. 13, pp. 411-430, 2000.

[5] L. Benaroya, F. Bimbot, “Wiener based source separation with HMM/GMM using a single sensor”, Proceedings of 4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003), Nara, Japan, 2003, p. 957–961.

[6] P. Flandrin, *Temps-Fréquence*, Hermès Paris, 1993.

[7] V. Pierson, N. Martin, “Watershed segmentation of time-frequency images”, Proceedings of the IEEE Workshop on Nonlinear Signal and Image Processing, 1995.

[8] I. Vincent, C. Doncarli, E. Le Carpentier, “Nonstationary signals classification using time-frequency distributions”, Proceedings of the International Synopsium on TFTS, p. 233-236, 1994.

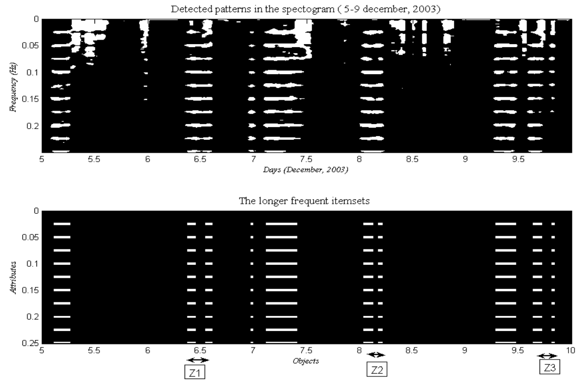

Figure 7 - Extraction of the convectors’ patterns

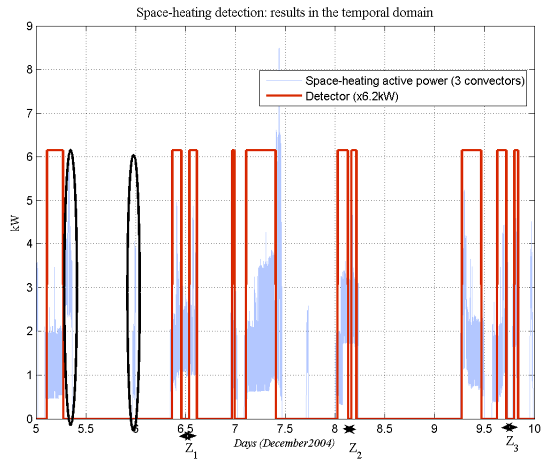

Figure 8 – Space-heating detection: results in the temporal domain

[9] B. Leprettre and N. Martin, “Extraction of pertinent subsets from time-frequency representations for detection and recognition purposes”, Signal Processing, vol. 82, no. 2, pp. 229-238, February 2002.

[10] M.J. Coates, *Time-frequency modeling*, PhD dissertation, University of Cambridge, August 1998.

[11] M.J. Coates and W.J. Fitzgerald ,”Regionally optimised time-frequency distributions using finite mixture models”, Signal Processing, no 77, pp. 247-260, 1999.

[12] C. Hory, Mixtures of Chi2 distributions for the interpretation of a time frequency representation, PhD dissertation, December 2002.

[13] N. Martin, C. Doncarli, *Décision temps-fréquence*, Hermès, Paris 2004.

[14] R. Agrawal. “Fast Discovery of association rules”, Advances in Knowledge Discovery and Data Mining. 1996

[15] B.Crémilleux et Rioult.F. « Optimisation d’extraction de motifs : une nouvelle méthode fondée sur la transposition de données ». Cap 2003

[16] F. Rioult, “Mining strong emerging patterns in wide sage data”, ECML/PKDD’04 Discovery Challenge, pp. 127-138, Pisa, Italy