Early Classification of Individual Electricity Consumptions

Asma Dachraoui[1], Alexis Bondu[1], and Antoine Cornuejols[2]

[1] EDF R&D,
1 avenue du General De Gaulle, 92140 Clamart, France
{asma.dachraoui, alexis.bondu}@edf.fr
[2] AgroParisTech, UMR-MIA-518 AgroParisTech - INRA,
16 rue Claude Bernard, F-75231 Paris Cedex 05, France
antoine.cornuejols@agroparistech.fr

Abstract. EDF[3] hires special contracts with costumers to flatten the consumption peaks. Smart meters are able to record consumptions and will be set up over 35 millions households. In this paper, we highlight the interest of early classification for detecting the households which probably contribute to the evening peak. The proposed approach is based on a collection of classifiers. In our experiments, we "early classified" the consumption of 1000 irish households and we obtained promising results.

Keywords: Early Classification, Smart Grid, time series, MODL

1 Introduction

The french electrical grid is being modernized by exploiting information and communications technology. One objective of the emerging *"smart grid"* is to coordinate the electrical demand in order to flatten the consumption peaks and to limit their impact on the environment. The *"smart meters"* will be set up within a few years and are able to record the individual power consumptions in real time. EDF[3] increasingly needs online analysis of time series, for instance to improve the maintenance of its industrial equipments [1]. This article deals with the ability of EDF to flatten the consumption peaks by hiring special contracts with the customers. The price of electricity may vary over time which aims at modulating the demand. The early detection of the households which contribute to the evening peaks is an important issue for EDF. Our objective is to wisely target the customers for whose the electricity price may be modified during the day. We propose to address this problem by applying an early classification approach on individual electricity consumptions *(see Section 2)*. In Section 3 our approach is applied on 1000 irish households characterized by their electricity consumptions. At last, future works are discussed in Section 4.

2 Related work and proposed approach

Related work : "Early classification of time series" consists in training a classifier which is able to make a prediction on uncompleted time series. In others words,

[3] EDF *(Electricité de France) is the main french electricity provider.*

the objective is to make the predictions as soon as possible on incoming time series, while controlling the error. Among the recent related papers, A. Bregon [2] applied a KNN classifier on uncompleted time series, and Z. Xing [3] proposed a *"Shapelets"* based approach which compares the uncompleted time series with *"typical subseries"*. As described in this section, our approach exploites several classifiers dedicated to each instant of the day.

Modeling choices : We assume the smart meters generate daily consumptions as time series composed by 48 measuring points *(each 30 minutes)*. A specific time period corresponding to the daily peaks is exploited to label the training dataset. The class values *"high consumption"* and *"low consumption"* describe the contribution of a given household to the peak of a specific day. Each time series is automatically labeled by comparing the cumulative consumption between *6:00 pm* and *8:00 pm* to a threshold. In our application context, there is no rule for determining this threshold. Thus, it seems interesting to define this threshold by varying the labeling percentage of *"high consumption"* examples. Our objective is to asses the influence of this threshold on the quality of the classifier.

General schema : Our approach is based on several classifiers trained in parallel. Each classifier corresponds to one instant of the day before *6:00 pm*. Consequently, 36 classifiers are trained for each household. These classifiers do not exploit the same explicative variables which correspond to the measuring points of individual consumptions. In the context of the smart grid the data is continuously generated, so online learning approaches are particularity relevant. Our approach uses historical data to train the classifiers off-line and the early predictions are conducted online. The progressive arrival of the time series is simulated by hiding the forthcoming measuring points : the input of the current classifier is only composed by the previous measuring points up to the current instant. In practice our approach would be difficult to be deployed, because the training of 36 classifiers for 35 millions households should be implemented in a distributed way. Furthermore, in our experiments the measuring points of the previous days are not exploited by the classifiers. We consider our approach as a baseline to be compared with alternative early classification algorithms. Our objective is to bring out the interest of early classification in the context of the smart grid.

The MODL supervised discretization approach : In this paragraph we present the classifier which is exploited for our experiments. The MODL[4] approach was chosen because : i) it is devoid of prior knowledge; ii) it is robust to the outliers; iii) it avoids the over-fitting; iv) it is parameter free; v) it asymptotically estimates any conditional density. The supervised discretization of a continuous variable consists in estimating the conditional distribution of classes owing to a piecewise constant estimator. The MODL approach [4] turns the discretization into a model selection problem. First, a family of discretization models is defined. The parameters of a discretization model are the following : the number of intervals, the bounds of intervals and the number of examples belonging to each class

[4] MODL *(Minimum Optimized Description Lenght)*

into each interval. A prior distribution $P(M)$ is defined over the family of models. This prior exploits the hierarchy of the model parameters : the number of intervals is first defined, then the bounds location and last, the conditional distribution are described in each interval. All possible values of model parameters are considered as equiprobable at each level of the hierarchy. A Bayesian approach is applied to select the best model, that is defined by maximizing the probability $P(M|D)$ of the model M given the data D. Exploiting the Bayes rule, and since the probability $P(D)$ is constant under varying the model, this amounts to maximizing $P(M)P(D|M)$. This approach leads to an analytical evaluation criterion. The optimization of this criterion defines the most probable model given the data. This approach is intrinsically regularized, a compromise is naturally reached between the complexity of the models and their generalization ability. For our experiments we use a selective naive Bayes classifier provided with the MODL supervised discretization. This approach selects a subset of informative variables in order to maximize the quality of the classifier. Numerous comparative experiments indicate that this classification approach provides good results in practice and avoids over-fitting [4].

3 Experiments

Real data : We exploite a real dataset[5] provided by the Irish CER which contains the individual consumptions of 6000 households during 500 days, sampled each 30 minutes. We chose to process the data separately for each household, and we formated this dataset into daily consumption containing 48 measuring points.

Experimental protocol : The training set *(70% of the data)* and the test set *(30% of the data)* are constituted by randomly chosen daily consumptions in order to limit the impact of potential non-stationarities. For instance, these non-stationarities can be due to the relocation of a customer, a new electrical device, a change of behavior ... etc. For each household, the data is repeatedly labeled by varying the proportion of *"high consumption"* between 8% and 50%. Then, several naive Bayes classifiers *(described in Section 2)* are trained for each instant of the day and for each value of the classes proportion.

Evaluation criterion : For a single household, the quality of the early classifier over time is represented by the AUC[6] plotted on each instant of the day. Then, the ALC[6] [5] which integrates the AUC over time is exploited to evaluate the earliness and the quality of the classifier.

Results : The households with a large value of ALC can be early detected as participating *(or not)* to the evening peak. The ALC criterion evaluates the earliness of the detection of the households which contribute to the evening peak. The most predictable households are particularly interesting for EDF to hire special contracts. The left chart of Figure 1 gives an example of a single household and plots the AUC of the classifiers depending on the instant of the day. The

[5] CER *(Comity of Energy Regulation)* Smart Metering Trial Data Publication. 2012
[6] AUC *(Area Under ROC Curve), ALC (Area under Learning Curves)*

AUC reached at *5:00 am* is close to the final AUC obtained at *8:00 pm*. This household is particularly "early predictable". The right chart of Figure 1 plots the average ALC obtained over 1000 households depending on the proportion of the *"high consumption"* class *(which varies the labeling threshold)*. This plot shows an optimum when 15% of examples are labeled by the class value *"high consumption"*. EDF is mostly interested by high levels of consumption which corresponds to low proportion of *"high consumption"* labeled examples.

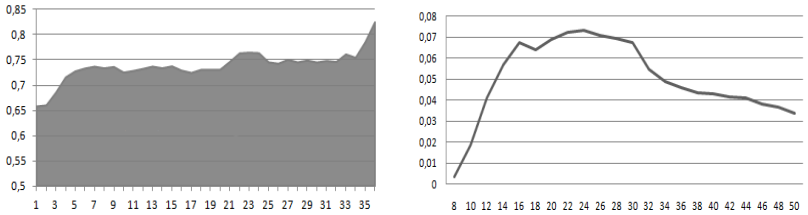

Fig. 1. AUC over time for a single household, and average ALC according to the labeling threshold

4 Conclusion

In this paper, we early classified the individual consumptions of 1000 households, for each a collection of 36 classifiers is exploited corresponding to the instants of the day. We showed the interest of early classification of time series : it allows EDF to target the predictable customers and to early detect the high levels of consumption during the day. In practice, our approach could be easily distributed but its time complexity should be reduced to be reasonably applied on 35 millions households. Future works will be done on alternative early classification methods. The representation and the compression of time series will be studied.

References

1. A. Bondu, B. Grossin, and ML. Picard. Density estimation on data streams: an application to Change Detection. In *EGC (Extraction et Gestion de l'Information)*, 2010.
2. A. Bregón, M.A. Simón, J.J. Rodríguez, C. Alonso, B. Pulido, and I. Moro. Early fault classification in dynamic systems using case-based reasoning. In *Proceedings of the 11th Spanish association conference on Current Topics in Artificial Intelligence*, CAEPIA'05, pages 211–220, Berlin, Heidelberg, 2006. Springer-Verlag.
3. Z. Xing, J. Pei, P.S. Yu, and K. Wang. Extracting Interpretable Features for Early Classification on Time Series. In *Proceedings of the Eleventh SIAM International Conference on Data Mining, SDM 2011, April 28-30, 2011, Mesa, Arizona, USA*, pages 247–258. SIAM / Omnipress, 2011.
4. M. Boullé. MODL: A bayes optimal discretization method for continuous attributes. *Machine Learning*, 65(1):131–165, 2006.
5. C. Salperwyck and V. Lemaire. Learning with few examples: An empirical study on leading classifiers. In *IJCNN (International Joint Conference on Neural Networks)*, pages 1010–1019. IEEE, 2011.