Data Stream and Load Forecasting : some ideas for a research project at Electricité De France

Alain Dessertaine

EDF R&D OSIRIS
1 avenue du Général De Gaulle
92141 Clamart Cedex
alain.dessertaine@edf.fr

Abstract. In this article, we are going to highlight the interest in using the data resulting from the future 32 millions of communicating meters which will equip all the French customers from now to 2013 in order to build up Court-terms forecasts and Means-terms concerning the EDF electric consumption (total, or by wallets) in an environment of data stream. First of all we will develop our reflections and afterwards we will evoke some tracks of research. These tracks should enable us to approach some modelling and forecasts by aggregation/disintegration of curves, as well as modelling and forecasts on Hilbertian data. Finally we will place side by side these ideas with the Stream-Mining type approaches.

1 Context

1.1 Some words about the potential use of the data Streams at EDF

The volume of data treated and analyzed by EDF is getting increasingly important. The installation of systems of measurement, becoming more and more efficient, will increase consequently this volume. Our aim will be to have a lighting on these data and information delivered in a current way for a better reactivity about some decision-makings. Then, for example, a rise in competence on the use and modelling of structured data stream should allow the calculus and the analysis of monitoring indicators and performances of the power stations of production, in an environment of data stream.

Moreover, the installation planed from now to 2013 of more than 32 millions of communicating meters of all of EDF' consumers should allow a better analysis of the EDF customers' spending patterns. We could also analyze "on line" consumption on more or less incorporated levels and predict it in order to adapt the production and the potential purchases on the electricity markets. Indeed, these meters will be used like sensors in order to measure all the load curves of each customer to very fine temporal granularities (going until the