

Data Extraction Tool to Analyse, Transform and Store Real Data from Electricity Markets

Ivo F. Pereira¹, Tiago M. Sousa¹, Isabel Praça¹, Ana Freitas¹,
Tiago Pinto¹, Zita Vale¹, and Hugo Morais²

¹ GECAD – Knowledge Engineering and Decision-Support Research Center
Institute of Engineering – Polytechnic of Porto (ISEP/IPP), Porto, Portugal
{ifdsp, tmsbs, icp, 1080598, tmcfp, zav}@isep.ipp.pt

² Automation and Control Group, Technical University of Denmark (DTU)
morais@elektro.dtu.dk

Abstract. The study of electricity markets operation has been gaining an increasing importance in the last years, as result of the new challenges that the restructuring process produced. Currently, lots of information concerning electricity markets is available, as market operators provide, after a period of confidentiality, data regarding market proposals and transactions. These data can be used as source of knowledge to define realistic scenarios, which are essential for understanding and forecast electricity markets behavior. The development of tools able to extract, transform, store and dynamically update data, is of great importance to go a step further into the comprehension of electricity markets and of the behaviour of the involved entities. In this paper an adaptable tool capable of downloading, parsing and storing data from market operators' websites is presented, assuring constant updating and reliability of the stored data.

Keywords: Databases, Electricity Markets, Machine Learning, Multi-Agent Simulators, Real Electricity Markets Data.

1 Introduction

Electricity markets are complex environments with very particular characteristics. A critical issue regarding these characteristics concerns the constant changes they are subject to. This is a result of the electricity markets' restructuring, which was performed so that the competitiveness could be increased, but it also had exponential implications in the increase of the complexity and unpredictability in those markets scope [1]. Electricity markets, as competitive environments, require good decision-support mechanisms to assist players in their decisions. Significant development concerning electricity markets players modelling and simulation including decision-support capabilities can be widely found in the literature [2-5].

Liberalized markets provide valuable information generally available to the community. Knowledge about markets grows from these last years and assists the definition of adequate players' profiles and behaviours. This information is also essential to test and validate existing simulation tools, such as MASCEM (Multi-Agent System

for Competitive Electricity Markets) [2, 3], making them capable of simulating realistic scenarios.

The realistic modelling of electricity markets, which provides the means for a suitable knowledge extraction from the study of advantageous simulations, requires an extensive search and organization of as much information as possible concerning these markets characteristics, particularities and constraints. Automatic tools, able to gather, store, update and organize data from distinct real electricity markets will be a key issue to improve markets simulators and the modelling of the participating entities, enabling researchers and professionals to extract knowledge and really learn from the last years' experience.

This paper presents a tool that was developed with the purpose of automatically searching for new electricity market data, extracting it from various websites, parsing the information, and storing it in the appropriate database, so that it can be used by the electricity market simulators to model realistic scenarios. This tool is adaptive to the data availability timings; it is capable of dealing with different data formats, and it includes parallel processing capabilities, in order to deal with multiple data sources processing.

This paper is organized in 5 sections. In section 2 an insight on the electricity markets data requirements is presented, both in what concerns the distinct nature of different countries' electricity markets, and the requirements from the currently most important electricity market simulators, namely MASCEM, which is being developed by the authors' research team since 2003 [2]. Section 3 presents the system capable of downloading, analysing and saving information from real electricity markets to provide a database with real historical data. Section 4 illustrates, by means of simple example, the processing of some of the data available at the GME (*Gestore Mercati Energetici* – Italian Energy Market Operator) homepage [6]. Finally, section 5 presents the most relevant conclusions and future implications concerning the presented work.

2 Electricity Markets Data

The liberalization of the electricity sector provides new market rules, the emergence of new market players and new forms of interactions among them [1, 7].

The functioning of liberalized markets over the last years provides valuable information most of the times available to the community through market operators websites. In fact, market operators such as MIBEL (Iberian Market Operator) [8], Nord Pool [9], EPEXSPOT (European Power Exchange) [10], MISO (central U.S. energy market) [11] and GME [6] provide on their web sites information regarding market proposals and transactions, usually after a period of confidentiality. The available information depends on each different market operator, however, essential information such as market proposals, with quantity and price; accepted proposals and established market prices is usually always available. This information grows in a very dynamic way, as it is put available in the various websites.