ELECON
ELECTRICITY CONSUMPTION ANALYSIS & ENERGY EFFICIENCY
www.elecon.ipp.pt

Towards Efficient European and Brazilian Electricity Markets

First ELECON Workshop

Institute of Engineering - Polytechnic of Porto, Porto, Portugal, September 24-25, 2013.

An automatic tool to Extract, Transform and Load data from real electricity markets

Ivo Pereira, Isabel Praça*, Tiago Pinto, Tiago M. Sousa, Ana Freitas, and Zita Vale

Polytechnic of Porto
Rua Dr. António Bernardino de Almeida, 431
4200-072 Porto

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 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, essential for understanding and forecast Electricity Markets behaviour. 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 the behaviour of the involved entities. In this paper we present an adaptable tool capable of downloading, parsing and storing data from market operators' websites, assuring actualization and reliability of 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

* Corresponding author.
E-mail address: icp@isep.ipp.pt

critical issue regarding these specific 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 tools to assist players in their decisions. Relevant research is being undertaken in this field, namely in what concerns player modelling and simulation, strategic bidding and decision-support.

The functioning of liberalized markets over the last years provides valuable information most of the times available to the community. Lessons can be learnt from these last years to improve knowledge about markets, to define adequate players' profiles and behaviours, but also to test and validate existing simulation tools, such as MASCEM (Multi-Agent System for Competitive Electricity Markets) [2, 3], making them suitable to represent reality and provide the means for a coherent and realistic analysis of its evolution (or possible alternative pathways for the future of the Electricity Markets sector).

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 this 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 we present an insight on the electricity markets data requirements, 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 we are developing 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. In section 4 we illustrate, by means of simple example, the processing of some of the data available at the EPEX Market operator homepage [4]. Finally, section 5 presents the most relevant conclusions and future implications of 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 [5, 6].

The functioning of liberalized markets over the last years provides valuable information most of the times available to the community trough market operators websites. Indeed, market operators such as the Iberian Market Operator [7], NordPool [8], EPEXSPOT (European Power Exchange) [4], MISO [9] and GME (*Gestore Mercati Energetici* – Italian Energy Market Operator) [10] 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 up in a very dynamic way, as it is put available in the various websites.

The Iberian Market Operator [7], which includes the electricity markets of Portugal and Spain, started on July 2006 with the futures market. One year later, in July 2007, both the day-ahead and the intraday markets started operating.

NordPool [8] is currently the largest energy market in Europe and includes the northern countries of this Continent, namely Norway, Denmark, Sweden, Finland, Estonia and Lithuania. There are two different markets in the NordPool: the ELSpot (day-ahead market) and the ELbas (intraday negotiations).

The EPEXSPOT [4], covers all the central area of Europe, including countries such as France, Germany, Belgium, Netherlands, Austria and Switzerland. The EPEX includes the day-ahead and intraday markets and an established deal with Czech Republic, Slovakia and Hungary to create a trilateral market between these countries.

The MISO [9] includes 15 U.S. states and the Canadian province of Manitoba and includes the day-ahead market.

GME [10], the Italian market operator, includes the day-ahead market (MGP), the intraday market (MI), ancillary services market (MSD) and The Forward Electricity Market.

With the information taken from the operation of different markets, lessons can be learnt from these last years to improve knowledge about markets, to define adequate players' profiles and behaviours, and realistic scenarios.

The need for understanding these market mechanisms and how the involved players' interaction affects the outcomes of the markets, contributed to the growth of usage of simulation tools, with the purpose of taking the best possible results out of each market context for each participating entity. Multi-agent based software is particularly well fitted to analyse dynamic and adaptive systems with complex interactions among its constituents, such as the electricity market. Several modelling tools directed to the study of restructured wholesale power markets have emerged. Some of the most relevant tools in this domain are:

- Electricity Market Complex Adaptive System (EMCAS) [11]: software agents with negotiation competence use strategies based on machine-learning and adaptation to simulate Electricity Markets;
- Agent-based Modelling of Electricity Systems (AMES) [12]: open-source computational laboratory for studying wholesale power markets, restructured in accordance with U.S. Federal Energy Regulatory Commission (FERC);
- Multi-Agent Simulator for Competitive Markets (MASCEM) [2, 3] : a platform based on multi-agent simulation firstly proposed in 2003 [2] that evolved into a complete tool acting in forward, day-ahead, and balancing markets, considering both simple and complex bids, and players strategic behaviour, based on ALBidS (Adaptive Learning strategic Bidding System) [13].

These simulators are very good examples of tools able to represent market mechanisms and players' interactions, but for them to be valuable decision support tools in foreseeing market behaviour, they need to be used in testing adequate and realistic scenarios. Real data analysis by means of a knowledge discovery process will be a crucial forward step to assure that MASCEM agents exhibit adequate profiles and strategies, namely by improving ALBidS strategies.

3. Automatic tool for real markets information extraction

After a careful analysis of the available data, since different operators make available different types of information, some of them providing even entities technical characteristics

and localization, a database model was defined. Figure 1 illustrates the Domain Model. As can be seen, the websites provide several files. Each file is related to a day, and may also include a session. The files have lines and those lines have: an agent, a proposal and the date. The proposals include the volume of traded energy, its price and whether it was a selling or buying proposal.

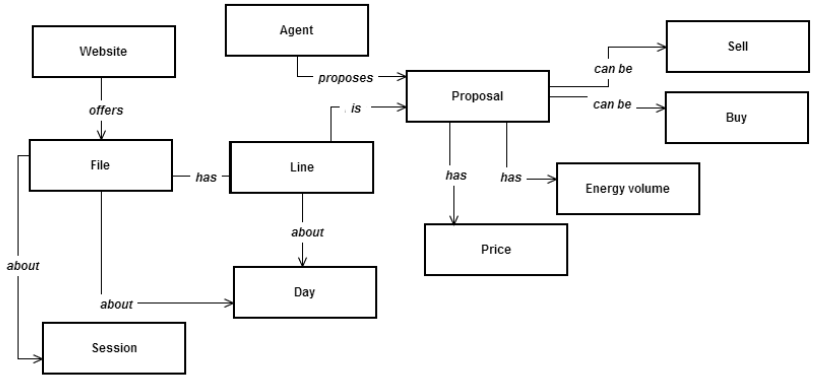

Fig. 1. Domain Model

Some requirements for the application are: the needs to assure the treatment of different file types, reliability in storing all the gathered data, as well as the needs to update the extracted data whenever it is available. Another relevant issue is efficiency regarding the treatment of great amounts of files, which, indeed in the initial use of the application may imply an enormous amount of files to assure gathering historic relevant data.

So that the files management can be performed in the best possible way, the adapter software standard is used for each file format. The adapter standard is used in circumstances in which a system needs to connect to an external service. In these cases the user should not need to allow such action, as it is intrinsic to the program. For this, it is necessary to create a class that provides the interface expected by the user and that uses the interface of the service provider. This means that it is necessary to create an interface with the signatures of the necessary methods for each adapter, and build each adapter in an independent fashion. The global class FileAdapter was built to provide the required abstraction that enables the system's ability to deal with new and different file formats.

The developed automatic tool includes four major steps:

- Download data - the download of several files containing the new data. The download depends on the website from which the data is being extracted, and it is performed accordingly to the data type of each file;
- Parse data - the extraction of the stored data from the downloaded files. The parsing of the data includes the analysis of the data fields of each file, from which the information and its associated value are taken;
- Store data - the storage of collected data in the database. The storage of the parsed data takes into account the necessary connections between different sets of data. This enables the data to be stored appropriately, respecting the interconnectivity and dependencies between all data;

- Mechanism for automatic data updates – machine learning techniques to automatically define downloads periodicity. The availability timings of type of file are analyzed so that the developed tool can process all available data the sooner as possible.

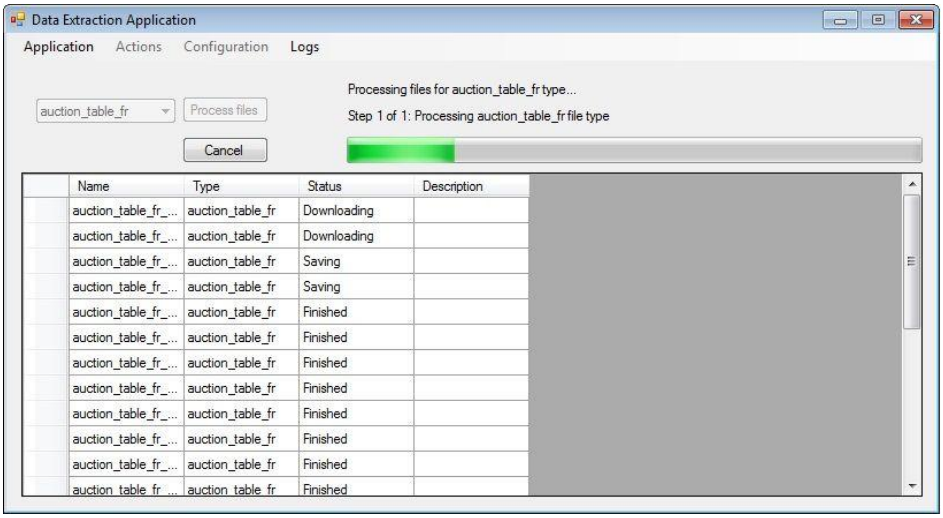

Fig. 2. Data extraction from the EPEX website

From figure 2 it is possible to see the interface of the developed tool, showing the treatment of different files in parallel, while displaying the status of each of the processed files.

4. EPEXSPOT Test-Case

The EPEX SPOT webpage [4] provides several information regarding the France, Germany, Austria and Switzerland market transactions. The existing information starts on 2006, and it is put available daily. This is a very complete website, easy to access and navigate, however, the information is not that easy to manipulate. The EPEX includes the day-ahead and intraday markets in separated webpages and both files considered by the data extraction tool. The "Auction" page provides information about the period, price and volume for each day. The "Intraday" page considers for each period the lowest, highest, last, weighted average and index prices (€/MWh) and also the buy and sell volumes (MW).

The example that is shown in this section refers to the file "Auction". This information can be accessed via browser on the EPEX website [4].

In this case, the presented information is shown on the webpage, and the page link has the following format: *http://www.epexspot.com/en/market-data/auction/auction-table/2013-09-10/FR*, where "*auction/auction-table*" corresponds to the market type, in this case the day-ahead market, "*2013-09-10*" corresponds to the date (in the format: yyyy-mm-dd) and indicating the country is "*FR*", which, in this case, corresponds to France.

In the first phase of the whole process - *Download Data*, the file to be analysed needs to be downloaded. Since the information shown on the webpage concerns several days, only the information of the day we are analysing is copied to a text file and then stored. The files are placed in a temporary folder, so that at the end of the process are removed since they are no longer required.

In the second phase - *Parse Data*, it is necessary to parse each extracted file. As the files to be read come in plain-text form it is very easy to open them. To execute the parsing of the files, the file structure was studied, bringing to attention that the format is similar to an HTML table, in which each line is a period of the day, and each of these lines are subdivided on other two lines, one for the price (€/MWh) and the other for the volume transacted (MWh). The data type of each field is shown in table 1.

Table 1. Field Types for the considered file

Field	Data Type
Period	Integer value indicating the period
Price (€/MWh)	Numeric Value
Volume (MWh)	Numeric Value

Finally, each line of the file will result in a database record being further included in a database corresponding to the *StorageData* phase of the process. This way, the information of all files is centred on the same database, in order to be used by other systems such as market simulators.

5. Conclusions

Electricity markets worldwide suffered profound transformations. The privatization of previously nationally owned systems; the deregulation of privately owned systems that were regulated; and the internationalization of national systems, are some examples of such transformations. With the increase of the competitiveness and consequent decrease of electricity price in sight, the restructuring of electricity markets brought a significant enlargement of the complexity in this sector.

Data regarding electricity market players of very distinct nature, with enormous differences in their characteristics; data regarding the market mechanisms of different types from country to country; data regarding the interactions and negotiations between players in different market environments; data regarding the decision support and strategic behaviour of such players; are only a few examples of what is required for electricity market simulators to adequately model the electricity market environment, so that realistic scenarios can be built and an advantageous decision support can be provided.

Even though most of the referred data is available, it is of very difficult access, for diverse reasons; therefore it is not used in the way and extent it should. In order to overcome the problem of the data access and treatment, making it available in an useful way for electricity market studies, in this paper we present our work regarding the development of a tool that provides a database with available information from real electricity markets.

The presented tool has the capacity of collecting, analysing, processing and storing real electricity markets data available on-line. Additionally, this tool is built in a way that it can deal with different file formats and types, some of them inserted by the user, resulting from information obtained not on-line but based on the possible collaboration with market entities.

This tool includes the capability of managing files using parallel processing, allowing the system to deal with multiple data sources at the same time. The different data files are accessed through a machine learning approach for automatic downloads of new information available on-line. All procedures are secured by a reliability mechanism that prevents from the storage of incomplete or unviable information.

The final result from the continuous execution of the presented tool is the definition and

implementation of a database that gathers information from different market sources, even including different market types.

This is a crucial tool to go a step forward in electricity markets simulation, since the integration of this database with a scenarios generation tool, based on knowledge discovery techniques, will provide a framework to study real market scenarios allowing simulators improvement and validation.

The possibility of using electricity market simulators capable of providing scenarios based on real data is an enormous asset for the study of electricity markets. Market operators and regulators are able to experiment and test new market rules and mechanisms, and obtain valuable insights regarding the consequences of such changes, both in what affects the market itself, and also in what way it influences the market players. In what concerns the advantages for market players, scenarios based on real data provide the means for testing different strategic behaviours, and analysing their results. Real market players are also able to thoroughly studying competitor players' actions, coming to understand how they behave, act and react in different circumstances and contexts, meaning an invaluable tool for adapting their own behaviours to the expected actions from behalf of competitors.

Acknowledgements

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/ under project ELECON - Electricity Consumption Analysis to Promote Energy Efficiency Considering Demand Response and Non-technical Losses, REA grant agreement No 318912.

This work is also supported by FEDER Funds through COMPETE program and by National Funds through FCT under the projects FCOMP-01-0124-FEDER: PEstOE/EEI/UI0760/2011, PTDC/EEA-EEL/099832/2008, PTDC/SEN-ENR/099844/2008, and PTDC/SEN-ENR/122174/2010

References

[1] G. Rothwell and T. Goméz, "Electricity Economics: Regulation and Deregulation", IEEE Series on Power Engineering, 2003, pp.15-42

[2] I. Praça, C. Ramos, Z. Vale, M. Cordeiro, "MASCEM: A Multi-Agent System that Simulates Competitive Electricity Markets", IEEE Intelligent Systems, vol. 18, No. 6, pp. 54-60, Special Issue on Agents and Markets, 2003

[3] Z. Vale, T. Pinto, I. Praça, H. Morais, "MASCEM - Electricity markets simulation with strategically acting players", IEEE Intelligent Systems, vol. 26, n. 2, Special Issue on AI in Power Systems and Energy Markets, 2011

[4] EPEX – European Power Exchange – homepage, http://www.epexspot.com, accessed on April 2013

[5] L. Meeus, et al., "Development of the Internal Electricity Market in Europe", The Electricity Journal, vol. 18, no.6, pp. 25-35, 2005

[6] M. Shahidehpour, et al., "Market Operations in Electric Power Systems: Forecasting, Scheduling, and Risk Management", Wiley-IEEE Press, pp. 233-274, 2002

[7] OMIE - Operador del Mercado Ibérico de Energia – Iberian Electricity Market Operator, homepage, http://www.omie.es, accessed on April 2013

[8] Nord Pool – homepage, http:// www.nordpoolspot.com, accessed on April 2013

[9] MISO energy – homepage, http:// www.misoenergy.org, accessed on April 2013

[10] GME - Gestore Mercati Energetici – Italian Energy Market Operator, homepage, http://www.mercatoelettrico.org, accessed on April 2013

[11] V. Koritarov, "Real-World Market Representation with Agents: Modeling the Electricity Market as a Complex Adaptive System with an Agent-Based Approach", IEEE Power & Energy magazine, pp. 39-46, 2004

[12] J. Sun and L. Tesfatsion, "Dynamic Testing of Wholesale Power Market Designs: An Open-Source Agent-Based Framework", Computational Economics, vol. 30, nr. 3, 2007, pp.291-327

[13] T. Pinto, et al., "Strategic Bidding Methodology for Electricity Markets using Adaptive Learning", Modern Approaches in Applied Intelligence, vol. 6704, pp. 490-500, 2011