

ADVANCED DATA ANALYSIS OF SMART GRID CONSUMPTION DATA

Iuliana ȘIMONCA (BOTH)

Bucharest University of Economic Studies, Romania

iuliana.botha@ie.ase.ro

Alexandra CORBEA (FLOREA)

Bucharest University of Economic Studies, Romania

alexandra.florea@ie.ase.ro

Abstract. *The article highlights a part of the research conducted within the SMARTRADE project for developing an informatics system prototype for forecasting, analysis and decision models for electricity suppliers in order to estimate electricity consumption and production efficiently. The paper focuses on data presentation layer, exemplifying some interesting aspects regarding NoSQL databases integration with Business Intelligence solutions.*

Keywords: Data Analysis, BI Technologies, NoSQL Databases, Smart Grid

JEL classification: C88, L88

DOI: 10.24818/ie2020.03.06

1. Introduction

This paper presents some research conducted within the SMARTRADE (Intelligent System for Trading on the Wholesale Electricity Market) research project, which has as objective the design and development of an informatics system prototype for forecasting, analysis and decision models for electricity suppliers in order to estimate consumption and production in an appropriate way for efficient transactions in the wholesale energy market.

A series of our own studies, like [1] or [2], as well as some established ones, such as [3], [4] or [5], offer detailed definitions of the smart grid concept. Thus, if the term *grid* refers to the electrical network, the term *smart* refers to digital technologies which ensure the interaction of customers with the electrical devices.

The European Commission defines the smart grid as “a necessary response to the environmental, social and political demands placed on energy supply” [6].

Over the past decades, the energy was largely centralized, therefore produced in a power plant and delivered from there constantly to the consumer. In the current era, energy is generated decentralized, even by the consumers themselves, the surplus being stored in the network. Therefore, there are important challenges for all actors involved, such as energy suppliers, consumers, distributors, etc. in order to deliver sustainable energy supplies. These will be achieved by using new technologies, products and services, in order to create an integrated environment for the actions of all users connected to it.

While [1] and [2] focus on data integration, exemplifying the process of integrating streaming data with static data that is received from smart meters, this paper focuses on data presentation, exemplifying some interesting aspects regarding NoSQL integration with Business Intelligence solutions.

2. Business Intelligence solutions for Smart Grid

Most software manufacturers offer complete BI solutions that include data integration and organization in databases, data warehouses or data lakes, advanced and multidimensional OLAP analysis, trend forecasting and analysis models, planning and monitoring tools, dynamic data presentation and analysis tools.

In the energy sector, the need for precise information is a critical desiderate in order to better make decisions and the specialized literature expresses this aspect very well.

In this regard, [7] includes a review of the main methods for data analysis in the field of smart grids, as well as a brief presentation of big data architecture and platforms. The paper [8] indicates some data mining algorithms applied for fault detection, electric device state monitoring, load profiling or energy forecasting. Some features of BI and data analytics are highlighted in paper [5], the authors developing a framework of BI&DA for smart grids, with a great emphasis on the data presentation sub-module. Also, the study [9] offers insights into the development of a BI solution for the Serbian energy market.

Going on the idea drawn in the analyzed studies, we have also reached the conclusion that the analysis of current and historical data, as well as forecasting of future ones are critical activities within a BI solution.

In order to extract valuable information from the source data, we chose Microsoft PowerBI. That is because the latest analysis of the market structure of BI and Analytics platforms (February, 2020) keeps the last two years distribution unchanged, with Microsoft Power BI dominating the entire market [10]. Thus, according to this Gartner classification, Microsoft Power BI is among the leaders of the products in its range, being the most popular data presentation tool in various formats, suitable for strategic management, but also a complete Business Intelligence product.

3. NoSQL Databases for Smart Grid

For a better understanding of the context, in the paper [11] we have presented the architecture of the SMARTRADE prototype. Then, through the studies in paper [1] we highlighted the process of streaming data integration in an Oracle database using Spark. However, given the large volume of data that is generated and need to be stored and processed, we have also considered a solution based on non-relational databases. This can be a better solution for an efficient data storage, since the data received from smart meters can have an approximate structure (easy to transpose into a CSV or JSON file) and, more important, is generated at an incredible speed, producing enormous quantities that need to be stored (the readings are received from smart meters every 1 minute).

Non-relational databases (or, briefly, NoSQL – “Not-only SQL” databases) are designed to solve scalability and performance issues for distributed, cloud-integrated IT solutions. They appeared in the context of large volumes of data, to solve the problems caused by the need for their storage and fast retrieval.

The paper [12] carries out a relevant study on the storage of data from the smart grid using different databases (relational - MySQL, Microsoft SQL Server, PostgreSQL, or NoSQL - MongoDB, Cassandra), depending on the type of data which need to be stored. As a result, NoSQL databases have shown an advantage in read / write performance against the other relational databases tested. Going into more detail, the paper [13] performs a comparative analysis on four open source NoSQL databases (MongoDB, Cassandra, Hbase and Elasticsearch). MongoDB had a good overall performance regarding the satisfaction of all requirements for smart grid data analysis.

During the research project, in order to offer access for BRPs to consumers data, we have considered the possibility of using MongoDB NoSQL database.

MongoDB (from *humongous*) is a document-oriented database, easy to understand and manipulate. It is ideal for both small, test projects and projects involving a large volume of data.

In the context of the Business Intelligence module that we need to develop, MongoDB can be considered as a very important data source, which can be integrated with Microsoft Power BI. In this way, the NoSQL data (structured, semi-structured or unstructured) can be converted into some interactive reports or readable diagrams.

In order to connect any Business Intelligence tool to MongoDB data sources, it is necessary to install and configure MongoDB Connector for BI [14]. Thus, the components needed to integrate MongoDB with Microsoft PowerBI are highlighted in the following figure:

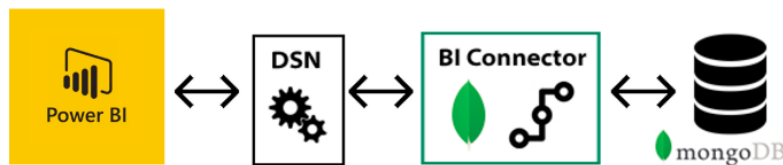


Figure 1. MongoDB – Microsoft PowerBI integration
Source: [14]

Once the connector is installed and a DSN is configured, we can access the non-relational data sources as any other data, which become available for building various BI interfaces.

The stored collections present documents with different structures, depending on how they were acquired from the data sources. However, since JSON documents are self-describing, an approximate structure can be deduced from them. Such an example is indicated in the following figure, being described a JSON document regarding the recorded consumption data:

```

{
  "_id" : ObjectId("5e97a6fe8f607aee5c881c71"),
  "ID_LOCATION" : 2,
  "TEMPERATURE" : 65.98,
  "WIND_DIRECTION_1" : "",
  "WIND_DIRECTION_2" : "",
  "WIND_BEARING" : 301,
  "WIND_SPEED_55" : "",
  "WIND_SPEED_100" : "",
  "PRESSURE" : 1012.63,
  "HUMIDITY" : 0.79,
  "SOLAR_IRRADIATION" : "",
  "INSOLATION" : "",
  "VISIBILITY" : 9.79,
  "APPARENT_TEMPERATURE" : 65.98,
  "PRECIPROBABILITY" : 0,
  "ICON" : "clear-night",
  "WIND_SPEED" : 5.05,
  "WIND_DIRECTION_3" : "",
  "CLOUD_COVER" : 0.08,
  "NEBULOSITY" : "",
  "PRECIPINTENSITY" : 0,
  "DEWPOINT" : 59.18,
  "GRADES" : "F",
  "DAY_CODE" : 3,
  "RECORDING_DATE" : "12-08-2015",
  "ID_PROFILE" : 34,
  "SEASON" : 3,
  "TOTAL_CONSUMPTION" : 37.886910001
}
  
```

Figure 2. Example of a MongoDB document

4. Advanced Data Analysis Module

In recent years, we are facing the situation where many organizations, flooded with a lot of data, must be able to decide which of them are relevant to the business and which are useless. The main goal of collecting large volumes of data is to process and analyze them in order to obtain meaningful information and, therefore, value.

For this reason, a vital condition for the proper functioning of the prototype created within the SMARTRADE research project, is the inclusion of a module totally dedicated to the advanced analysis of the data acquired from the data sources, as well as of the forecasted ones.

During the current phase of the research project we developed this advanced data analysis subsystem (further referred to as ABI module). This module represents the Business Intelligence solution that will be implemented within the informatic prototype and contains components dedicated to each of the actors indicated in the following figure:

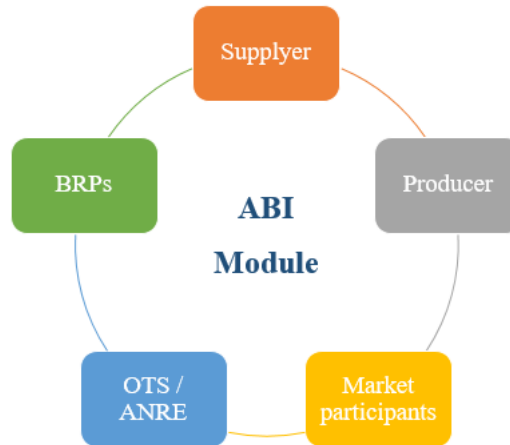


Figure 3. ABI components

For now, we have developed two of the most important sub-modules dedicated for advanced analysis of the activities within the Balance Responsible Parties (BRPs):

- 1) The analysis module for BRP affiliated electricity suppliers which contains a series of interactive reports and graphs for visualizing the electricity consumption for different consumers and profiles, over time periods and locations. Using Microsoft PowerBI, we developed attractive interfaces for:
 - Analyzing consumption profiles and distribution of consumers by clusters;
 - Visualizing energy consumption history over different time periods, by locations and types of consumers;
 - Analyzing deviations between forecasted values and actual energy consumption over different time periods, by location and by consumption profiles;
 - Visualizing energy consumption through interactive maps;
 - Analyzing the consumption according to the time period and the types of electrical devices installed at consumer locations;
 - Visualizing the distribution of consumption by types of consumers: residential, commercial, industrial.
- 2) The analysis module for BRP affiliated electricity producers which contains reports and graphs for visualizing the production of electricity on different plants, groups (like turbines or photovoltaic panels) over time periods and locations. The production reports are correlated with the weather data and some of the reports are developed in order to visualize the generated power prediction based on the forecast algorithm implemented within the prototype.

Thus, we have built various visualizations needed for the dashboards to be integrated in the final prototype. They are dedicated for electricity producers who access the ABI module. We have developed in Microsoft PowerBI the following sets of interfaces that contain interactive reports and graphs for:

- Visualizing the production history on electrical groups for different periods of time;
- Analyzing deviations between the predicted values and the actual ones on electrical groups, for different periods of time, based on the power prediction algorithm generated by types of groups and power stations (photovoltaic / wind);

- Visualizing notifications sent to BRP / OPCOM on each control panel;
- Visualizing the production on electrical groups through interactive maps.

For the purpose of further integration with the final prototype, it was necessary for all the developed interfaces to be published in the web environment, so that they could be opened and analyzed by simply accessing a URL. Thus, we published the reports containing the views created in Microsoft PowerBI Desktop in Power BI Dashboard, hosted in Azure Cloud. Subsequently, for each of the two main reports (the one containing the interfaces dedicated to the Suppliers and the one intended for the Producers), we requested the access link, as well as the HTML code that can be incorporated into the prototype web pages.

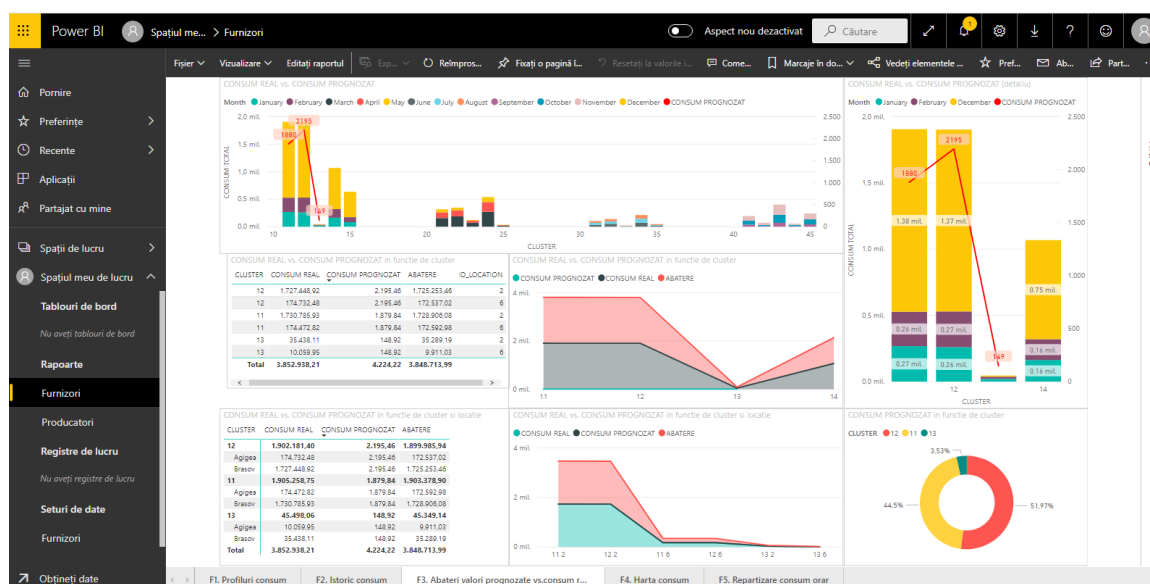


Figure 4. Microsoft PowerBI interactive report dedicated for energy suppliers

Furthermore, algorithms for data processing and predictive analytics can be implemented through scripts in Python or R, directly in PowerBI Desktop, in order to obtain graphical visualizations of the data. Thus, parts of the component for data analytics included in the SMARTRADE prototype, can be easily implemented through a single tool, by merging analytics and analysis.

6. Conclusions

The paper presents some aspects regarding the data presentation layer from the architecture of the solution proposed within the SMARTRADE project. We studied and exemplified the use of Microsoft PowerBI to develop interactive visualizations based on various data sources, including NoSQL databases.

Acknowledgment

This paper presents the scientific results of the project “Intelligent system for trading on wholesale electricity market” (SMARTRADE), co-financed by the European Regional Development Fund (ERDF), through the Competitiveness Operational Programme (COP) 2014-2020, priority axis 1 – Research, technological development and innovation (RD&I) to support economic competitiveness and business development, Action 1.1.4 - Attracting high-level personnel from abroad in order to enhance the RD capacity, contract ID P_37_418, no. 62/05.09.2016, beneficiary The Bucharest University of Economic Studies.

References

- [1] V.Diaconita, I.Botha, A.Florea. (2019). “Big Data integration of smart grid consumption data”. Proceedings of the 18th International Conference on Informatics in Economy: Education, Research & Business Technologies, 2019, ISSN: 2247 – 1480
- [2] I.Botha, A.Belciu, I.Lungu. (2018). “Integrating Big Data technologies with the data warehouse architecture in the context of smart grid”. Proceedings of the 17th International Conference on Informatics in Economy: Education, Research & Business Technologies, 2018, ISSN: 2247 – 1480
- [3] Initiatives that catalyze the industry to modernize the grid. Internet: <https://www.smartgrid.gov/>
- [4] M.Williams, M.Hardy.(2018). “Smart grids, big data and business intelligence transforming the energy sector”. Power Engineering International Magazine. Available: <https://www.powerengineeringint.com/editors-picks/smart-grids-big-data-and-business-intelligence-transforming-the-energy-sector/>
- [5] G.Escobedo, N.Jacome, G.Arroyo-Figueroa. (2016). “Business Intelligence and Data Analytics (BI&DA) to Support the Operation of Smart Grid”. Conference: Special Session on Recent Advancement in IoT, Big Data and Security, pp.489-496, 2016
- [6] European Commission (2006). “European Smart Grids Technology Platform. Vision and Strategy for Europe’s Electricity Networks of the Future”. [Online]. Available: https://ec.europa.eu/research/energy/pdf/smartgrids_en.pdf
- [7] B.P.Bhattarai, S.Paudyal, Y.Luo et.al. (2019). “Big Data Analytics in Smart Grids: State-of-the-Art, Challenges, Opportunities, and Future Directions”. IET Research Journals - Smart Grid, pp.1-16, Feb.2019, ISSN 1751-8644.
- [8] Y.Zhang, T.Huang, E.F.Bompard. (2018). “Big data analytics in smart grids: a review”. Energy Informatics, 1, 8, 2018. Available: <https://energyinformatics.springeropen.com/articles/10.1186/s42162-018-0007-5>
- [9] M.Radenković, J.Lukić, M.Despotović-Zrakić, A.Labus, Z.Bogdanović. (2018). “Harnessing business intelligence in smart grids: A case of the electricity market”. Computers in Industry, Vol. 96, 2018, pp 40-53
- [10] Gartner (2020). 2020 Gartner Magic Quadrant for Analytics and Business Intelligence Platforms, [Online]. Available: <https://info.microsoft.com/ww-landing-2020-gartner-magic-quadrant-for-analytics-and-business-intelligence.html>
- [11] A.Bara, S.V.Oprea, I.Botha, O.B.Tor (2016). “Conceptual design and architecture of an informatics solution for smart trading on wholesale energy market in Romania”. Database Systems Journal vol. VII, no. 4/2016, pp. 3-12 [Online]. Available: http://dbjournal.ro/archive/26/26_1.pdf
- [12] E.Yılmaz, H.Polat, S.Oyucu, A.Aksöz, A.Saygin. (2018). “Data storage in smart grid systems”. Conference: 2018 6th International Istanbul Smart Grids and Cities Congress and Fair (ICSG), 110-113.
- [13] M.H.Ansari, V.Tabatab, B.Bahrak. (2019). “Evaluation of big data frameworks for analysis of smart grids”. Journal of Big Data, 6, 109, 2019. Available: <https://link.springer.com/article/10.1186/s40537-019-0270-8>
- [14] A.Ashraf (2019). “sData visualization using Power BI and MongoDB” [Online]. Available: <https://medium.com/@alameerashraf/data-visualization-using-power-bi-and-mongoddb-e694328220b5>