Big Data Research 2 (2015) 94–101

Contents lists available at ScienceDirect

Big Data Research

www.elsevier.com/locate/bdr

# Big Data Analytics for Dynamic Energy Management in Smart Grids ☆

Panagiotis D. Diamantoulakis [b,*], Vasileios M. Kapinas [b], George K. Karagiannidis [a,b]

[a] Department of Electrical and Computer Engineering, Khalifa University, PO Box 127788, Abu Dhabi, United Arab Emirates
[b] Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, GR-54124 Thessaloniki, Greece

ARTICLE INFO

*Article history:*
Received 21 December 2014
Received in revised form 27 February 2015
Accepted 13 March 2015
Available online 28 April 2015

*Keywords:*
Big data
Smart grids
Dynamic energy management
Predictive analytics
Artificial intelligence
High performance computing

ABSTRACT

The smart electricity grid enables a two-way flow of power and data between suppliers and consumers in order to facilitate the power flow optimization in terms of economic efficiency, reliability and sustainability. This infrastructure permits the consumers and the micro-energy producers to take a more active role in the electricity market and the dynamic energy management (DEM). The most important challenge in a smart grid (SG) is how to take advantage of the users' participation in order to reduce the cost of power. However, effective DEM depends critically on load and renewable production forecasting. This calls for intelligent methods and solutions for the real-time exploitation of large volumes of data generated by the vast amount of smart meters. Hence, robust data analytics, high performance computing, efficient data network management, and cloud computing techniques are critical towards the optimized operation of SGs. This research aims to highlight the big data issues and challenges faced by the DEM employed in SG networks. It also provides a brief description of the most commonly used data processing methods in the literature, and proposes a promising direction for future research in the field.


## 1. Introduction

A smart grid (SG) is the next-generation power system able to manage electricity demand in a sustainable, reliable and economic manner, by employing advanced digital information and communication technologies. This new platform aims to achieve steady availability of power, energy sustainability, environmental protection, prevention of large-scale failures, as well as optimized operational expenses (OPEX) of power production and distribution, and reduced future capital expenses (CAPEX) for thermal generators and transmission networks [1]. The upcoming technology in the framework of SG facilitates the development and efficient interactive utilization of millions of alternative distributed energy resources (DER) and electric vehicles [1–3]. To this end, each consumer location has to be equipped with a smart meter for monitoring and measuring the bi-directional flow of power and data, while supervisory control and data acquisition (SCADA) systems are needed to control the grid operation.

While dynamic energy management (DEM) in conventional electricity grids is a well-investigated topic, this is not the case for SGs. This is due to its much more complicated nature, since complex decision-making processes are required by the control centers [4,5]. Energy management systems (EMSs) in SGs include i) real-time wide-area situational awareness (WASA) of grid status through advanced metering and monitoring systems, ii) consumers' participation through home EMSs (HEMS), demand response (DR) algorithms, and vehicle-to-grid (V2G) technology, and iii) supervisory control through computer-based systems [6]. A typical overview of the SG and the included systems and technologies is given in Fig. 1. The quality and reliability of the data collected is a key factor for the optimized operation of the SG, thus rendering data mining and predictive analytics tools essential for the effective management and utilization of the available sensor data [7]. This is because effective DEM relies dramatically on short-term power supply and consumption forecasting, which handles prediction horizons from one hour up to one week [8]. Additionally, the sensor data contains important correlations, trends, and patterns that need to be exploited for the optimization of the energy consumption and the DR, among others [4]. Most of the research related to data mining in SGs deal with predictive analytics and load classification (LC), which are necessary for the load forecasting, bad data correction, determination of the optimal energy resources scheduling, and setting of the power prices [9,10]. The efficient processing of the produced vast amount of data requires increased data storage and computing resources, which imply the need for high performance computing (HPC) techniques.

☆ This article belongs to BDA-HPC.
* Corresponding author.
*E-mail addresses:* padiaman@auth.gr (P.D. Diamantoulakis), kapinas@auth.gr (V.M. Kapinas), geokarag@auth.gr (G.K. Karagiannidis).

http://dx.doi.org/10.1016/j.bdr.2015.03.003
2214-5796/

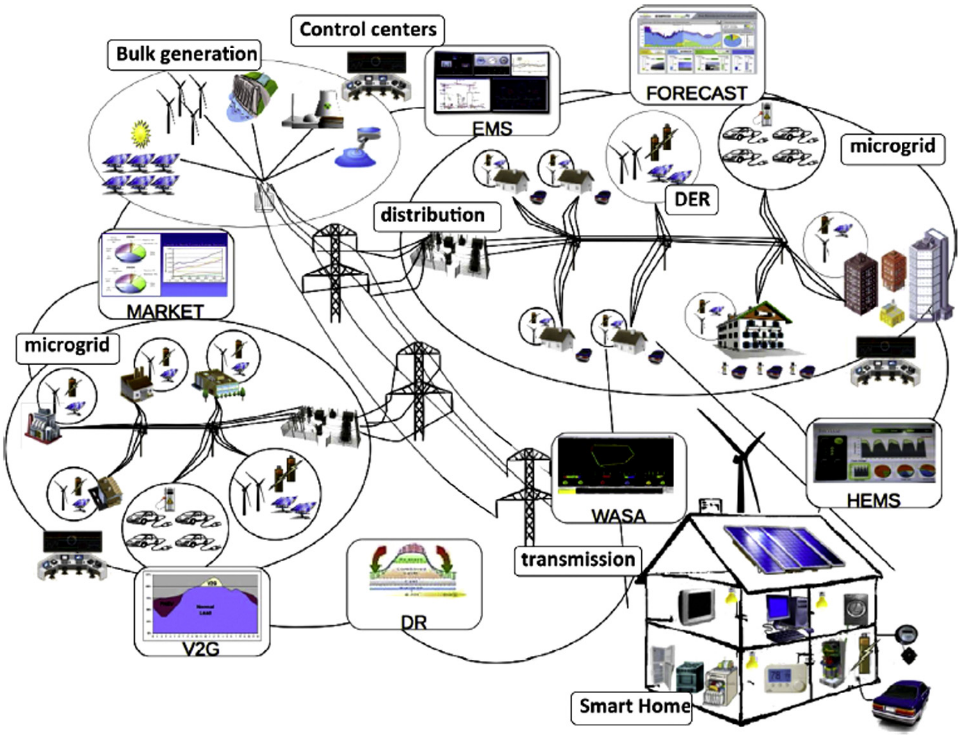


**Fig. 1.** Smart Grid overview [6].

This work differs from other related surveys in the literature, such as [11–16], in being the first meta-analytic review on efficient SG data processing with focus on DEM. Besides, it gives useful insights into technologies and methods from the area of big data analytics (BDA) that have to be further explored into the framework of DEM, demand forecasting, and dynamic pricing. Our investigations reveal that there is free space for research in the following topics:

- Design and development of algorithms that can accurately extract the load patterns from large-scale datasets.
- Design of machine learning (ML)-based algorithms with improved forecasting performance, low memory requirements, and scalable architecture.
- Development of novel data-aware resource management systems that can provide powerful data processing in distributed computing systems and clusters for real-time processing.

It is also highlighted that scalability and flexibility, achieved through the construction of robust algorithms and fast provisioning of HPC resources, can enable the efficient processing of the large data volumes involved in DEM and short-term power demand/supply forecasting.

The rest of the paper is organized as follows. Section 2 gives insight on the reasons why conventional data processing techniques are not appropriate for DEM in SGs. Section 3 focuses on smart meter data stream mining and presents the most commonly used methods. Section 4 is dedicated to the appropriate HPC techniques. Section 5 provides promising future research directions, while Section 6 concludes the paper.

## 2. Dynamic energy management in SGs: a big data issue

DEM requires power flow optimization, system monitoring, real-time operation, and production planning [17]. In more detail, DEM in an SG is a complicated, multi-variable procedure, since the latter enables an interconnected power distribution network by allowing a two-way flow of both power and data. This is in contrast to the traditional power grid, in which the electricity is generated at a central source and then distributed to consumers. Thanks to the bi-directional flow of information and power between suppliers and consumers, the grids become more adaptive to the increased penetration of DER, encouraging also users' participation in energy savings and cooperation through the DR mechanism [10,18,19].

DR can be applied to both residential (e.g., cooling, heating, electric vehicles (EVs) charging, etc.) and industrial loads and includes three different concepts: i) energy consumption reduction, ii) energy consumption (or production) shifting to periods of low (or high) demand, and iii) efficient utilization of storage systems [20]. It should be noticed here that plug-in EVs can be considered as storage devices, while the careful scheduling of their charging and discharging can benefit both their owners and the utilities. Obviously, this further increases the parameters that the DEM algorithms have to take into account, such as the EVs charging profiles. Consequently, the associated complexity is also increased, creating at the same time storage capacity prediction problems [21]. Thus, a crucial issue in SGs is how to manage DR in order to reduce peak electricity load, utilizing at the same time renewable energies and storage systems more efficiently. Finally, effectiveness of DR algorithms depends critically on demand, price, load, and renewable energy forecasting, which highlights the need for sophisticated signal processing techniques [22].

The electricity demand and renewable production in the SG environment is affected by several factors, including weather conditions, micro-climatic variations, time of day, random disturbances, electricity prices, DR, renewable energy sources, storage cells, micro-grids, and the development of EVs [23–26]. High forecasting accuracy accommodates the generation and transmission planning, i.e., deciding which power plants to operate and how much power should be generated by them at a specific time-period, with the aim to reduce the operating cost and increase the reliability [27]. It also enables the utilities to successively estimate the electricity cost and correctly set the electricity prices, capturing the in-