ISSN: 0971-1260

Vol-22-Issue-17-September-2019

A Preview of Big Data Analytics in Power Systems

Dr. P. K. Choudhary
Asst. Professor, SEEE
Lovely Professional
University. Phagwara, Punjab,
India
prashant.24788@lpu.ac.in

M Choudhary
Asst. Professor, CSE
RCET, Bhilai, Chhattisgarh,
India
minukum@gmail.com

Dr. Bhagwan Shree Ram
Professor, SEEE
Lovely Professional
University, Phagwara, Punjab,
India
bhagwan.24828@lpu.co.in

Abstract — Over the past few years, power systems have faced various challenges and technological innovations and are rapidly evolving into digital systems by deploying the concept of smart grids. Smart grid is a technology that can meet modern energy requirements by incorporating advanced information and communications technology (ICT), particularly smart metering, will generate large energy data in terms of volume, speed and variety. Big data generated can bring enormous benefits to better energy planning, efficient generation of energy, and distribution. Power systems can benefit from the application of big data analytics to generate large amounts of data, which can help leverage the optimization processes currently taking place in power grids. Big data methods can be used to benefit the entire quality of the electric power chain. This paper presents a brief overview of potential applications and challenges to be considered for effective use.

Keywords— *Big Data; Analytics; Power System;, Smart Grid; Cyber Security*

INTRODUCTION

The world has undergone a digital revolution over the last decade, and the increased digitalization of different sectors, together with the rapid development of the idea of the Internet of Things (IoT) and an increase in the number of connected devices, have resulted in the collection of large amounts of data. This is still growing and the data generated and exchanged is expected to reach 44 ZB by 2020 [1]. Digitalization is a reality in the energy sector, with the introduction of the idea of smart grids and the growing penetration of ICT. This digitalization can be seen from generation to distribution / retail in all stages of the energy value chain [2]. At the same time, power systems face numerous operational challenges such as the need to minimize operating costs and increase performance, increase the share of renewable energy, and address environmental issues, among others. Together with the advances presented by big data, these challenges represent new opportunities, particularly since the next generation of power systems, the smart grids, will be extremely data-intensive [3–6]. Ultimately, applying big data techniques to power systems can lead to real-time optimization of power generation and

ISSN: 0971-1260

Vol-22-Issue-17-September-2019

transmission, precise load demand prediction, analysis of consumption patterns leading to new services, and dynamic pricing strategies [7].

However, current status of both power systems and big data technologies still need to overcome different challenges to meet these goals. Most important challenges include: effective data acquisition and storage; data curation methodologies; how to exactly use the stored data to extract business value and how to mitigate privacy while using these data. These challenges become even more important in a sector which is traditionally adverse to change which is the case of the energy sector. Nonetheless, in order to meet these targets, the current status of both power systems and big data technology also needs to overcome various challenges. The most critical issues include: effective data acquisition and storage; methodologies for data curation; how to use the stored data correctly to derive business value and how to protect confidentiality while using such data. In a market that is historically resistant to change, which is the case of the energy sector, these problems become even more significant.

Reliability and long return-ofinvestment (ROI) scenarios are two of the main explanations for this activity. The energy sector needs to ensure high reliability and once a system is proven safe from an operational point of view, many utilities are reluctant to switch to another technology. Furthermore, this is an asset-intensive sector where most assets have a high cost and a large ROI indicating that the same assets must ensure long-term safe operation.
In view of this, the adoption of new technologies and the shift to a data- driven approach must be accomplished in a timely manner, taing advantage of opportunities arising from the challenges facing the sector.

This paper provides a short summary of possible applications and challenges of big data in power systems in order to better understand how big data can be applied to power systems and how they can benefit from techniques that are already familiar to the ICT world but still considered unusual to the world of power systems. The power system community has a different concept of security from the cyber security community, while traditional cryptography in the cyber security community does not find real-time variables and strong cyberphysical couplings in advanced power systems such as smart grids. These parallel and unconverged research efforts in security would leave significant loopholes. The paper is structured as follows: section II briefly describes the main characteristics of big data and how they can be adapted to meet the power system needs; section III describes safety challages for big data analytics; finally, section IV contains some comments and conclusions about the paper subject.

II. BIG DATA IN POWER SYSTEMS

A. *Characteristics of big data*

ISSN: 0971-1260

Vol-22-Issue-17-September-2019

Big data is more than just a large data set. A common definition of big data involves classifying its features using the 4V's [8]. Such features refer to various sectors that can benefit from the use of big data. The 4V's can be briefly explained as follows, taking into account the particularities of the energy sector.

Volume. A large volume of data will be generated by the digitization of the power system, including the large-scale deployment of smart metering devices (multimillion-unit scale) and changes at the substation level. It will be beneficial to generate data at the minute level and even at the second level [9-11]. According to New York State's 2012 housing units, 127.1 TB is required to store the information on electricity consumption every day [9]. It presents new possibilities, but also problems when it comes to storing and processing such data [12], [13]. ***Velocity.*** This function refers to the processing speed needed in power systems to capture, process and use big data. The data collection time interval can be as short as a few milliseconds and considering that the processing of these data can be collected in a real-time or near-real-time basis, a challenge can arise.

Variety. It applies to the various data that need to be taken into account. Multiple sources and types of structured and unstructured data are applicable to energy generation, energy planning, and distribution, etc. Structured data (e.g. sensor and meter data), semi-structured data (e.g. weather forecast information) and unstructured data (e.g. costumer activity data) are included in energy systems. In addition to the data generated by the Advanced Metering Infrastructure (AMI) and smart meters on household appliance-level energy consumption, other data such as various renewable energy, weather, and market etc. are collected for optimal system operation. Processing these various datasets is a challenge, particularly when unstructured data sources are increasing [14], [27].

Veracity. This feature tackles the need for information ambiguity to be taken into account. Confidence in the veracity of used data is crucial in a key sector of society, such as the power one, and big data technologies have to ensure this veracity as one of the most important aspects. Such four features allow the use of big data techniques to extract business value. Value is often marked as the fifth V [15] for this reason. Taking into account the particular application in the field of power systems, these 5V characteristics can be complemented by the 3E's [16], which stand for: Energy that can be saved from the implementation of big data techniques; Exchange of information between energy and other sectors with the goal of maximizing the quality of big data; *Empathy* meaning that by providing better energy services needs of users can be achieved and their satisfaction increased.

B. ***Big data and smart energy management***

Different measures need to be taken in order to achieve a data-driven smart energy management system based on big data analysis. These include the different stages of big data analytics operations which are summarized in the value chain of Fig. 1.

ISSN: 0971-1260

Vol-22-Issue-17-September-2019

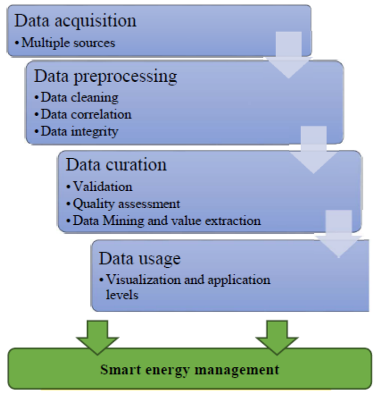

Fig. 1: Value chain of big data analytics in power systems

In the application of big data analytics in power systems, the four main activities of the identified value chain are of great importance. The processes of data acquisition and preprocessing deal with all the requisite topics to collect and prepare data from multiple sources by establishing a common understanding of all used datasets. Data curation is the core process and deals with all the extraction of business value from data collected in the preceding stages.

Data visualization and representation allows players in the energy sector to develop their business with the information-driven approach and achieve various goals such as energy efficiency, customer interaction, real-time tracking, demand response, smart management and (among others) dynamic pricing strategies. Big data strategies can be implemented across the energy sector's entire value chain, including: power generation, transportation, supply and demand side management (DSM). The application depends on the company's core business using this approach. Next are given some examples of possible applications, arranged according to the core stage of the management of power systems [11].

Power generation: capacity enhancement, asset management and scheduling of generation;
Transmission: asset management, network planning, grid failure identification;
Distribution: real-time sensing, voltage regulation, asset management forecast;
Demand side management: demand response, load prediction, adaptive pricing strategies, real-time consumer tracking, fraud detection.

Such systems depend on numerous different data sources at the various stages of the value chain of the power sector. There is a wide range of data sources that can be separated into various categories [17]: information from smart meters by deploying advanced metering infrastructure (AMI) [18]; network resource data (primary equipment sensors); third party data (off-grid data sets) and asset management data (smart devices in the grid). Due to the very different nature of these data sources, data integration / correlation becomes a key step in the value chain for data-driven power systems. Big data analytics need to take these different data sources into account and take into account the risks and privacy issues that may arise in these different steps from data utilization. Two different aspects need to be taken into account: consumer data privacy, which is important for consumer engagement in the process, and data privacy and data exchange control between different players. The need for data privacy is also a challenge to be addressed by

ISSN: 0971-1260

Vol-22-Issue-17-September-2019

creating / adapting sound and safe solutions to ensure that utilities always have their data protected. Another difficulty is to properly define ownership of data. Different players in the energy sector use some datasets, but the interfaces between them and the data owners need to be clearly defined otherwise this may result in disputes that hinder the implementation of a data driven approach.

C. *Current applications and opportunities*

Resource intensive is the major steps of the value chain of power systems (generation, transmission and distribution). These resources produce large amounts of data, mainly related to monitoring and control of the situation. Through providing a new vision related to operation and maintenance activities, the use of big data technologies will increase the life span of resources. A data-driven approach can improve the operating conditions of the various existing assets in various aspects such as real-time monitoring, fault diagnosis, maintenance scheduling, thus enhancing the sustainability and reliability of the grid. The main challenge associated with this request is the need for data privacy, as discussed earlier. Power systems consist mainly of natural monopolies (transmission and distribution grids) and massive (generation) operators. Nevertheless, the product side, through the delivery of DSM systems, is the one that provides more possibilities for big data applications. Big data application areas are within a very wide spectrum. Smart meter deployment, real-time monitoring of load profiles, and load classification techniques will enable many different services to be created, such as consumer segmentation and dynamic pricing. Such tactics will ultimately change customer behavior and make it a much more effective player in the energy sector than it is now. Big data can also generate new consumer-related business opportunities, providing new energy-related services and products. Again, data privacy needs to be ensured, and to avoid market imbalances, the use of consumer data from utilities needs to be clearly regulated.

One area where the value of big data is already very important and continues to grow is linked to renewable energy management. Energy from renewable sources, particularly in European countries, is already a major source of electricity. For example, 57% of electricity was generated from renewable sources in Portugal in 2016 [19]. A precise forecast of renewable energy sources such as wind and solar is essential to ensure grid stability and power quality [20-21]. There are a wide variety of prediction models, which can benefit from more data sources and big data analytics techniques being implemented. Offshore renewables are now being looked at by various countries as the next sustainable source of energy. Those include not only offshore wind farms that are already a mature technology, but also ways to extract energy from the ocean, i.e. wave and tidal. Because these areas are the ones that still need to overcome various technical challenges and still need a sound research effort to reach maturity.

III. BIG DATA ANALYTICS SECURITY CHALLANGES

Smart grid infrastructure integrates state-of- the-art ICT technology. Networking, smart communications technology and information processing functions are embedded in every aspect of the energy system, ranging from power generation, power transmission, power distribution, and consumer appliances [22-24]. A large-scale smart grid will consist of thousands of

ISSN: 0971-1260

Vol-22-Issue-17-September-2019

microgrids operating in both isolated and interconnected modes [22]. Smart metering is an integral component of a smart grid with large-scale installation of smart meters in homes and other parts of the system. In terms of volume, velocity, and variety, the widespread implementation of advanced ICT, especially smart metering, will generate large energy data.

There are many potential benefits to be gained from smart grid data [24]: automatic and real-time tracking of the energy consumption of consumers, automated billing storage, identification of energy losses (possible fault and/or fraud), early outage warning, rapid detection of energy supply fluctuations, intelligent and real-time energy planning and pricing. Smart energy data includes the private information of individual users that must be covered under different legal regulations [25], [26]. The data may also include an organization's sensitive information. More importantly, these data can be used to make decisions that affect the critical infrastructure's safe operation. Security and privacy are therefore going to be an important issue. However, due to the big data nature of the smart energy data, tight cyberphysical couplings, distributed and open infrastructure environment, this is also very challenging. In data analytics and security, the big data complexity of smart energy presents new challenges that existing technologies are unable to deal with. In addition to the typical big data analytics and safety issues, energy big data analytics should add another dimension of complexity in coping with the specific variable of close cyber-physical couplings. The power system community has a different concept of security from the cyber security community, while conventional cryptography in the cyber security community does not consider real-time factors and tight cyberphysical couplings in advanced power systems such as smart grids. These parallel and unconverged research efforts in security would leave significant gaps.

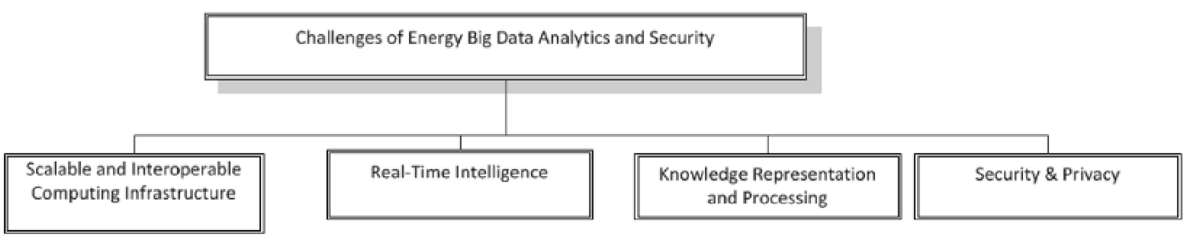

Fig. 2: Taxonomy of Energy Big Data Challanges

CONCLUSIONS

Recent developments in ICT technologies and the increase in the number of available data have revolutionized various sectors, including power systems, which can benefit from the use of big data. Given this, this paper identified the main characteristics of big data (when applied to power systems) and provided a succinct analysis of the major advantages and challenges associated with this synergy. One of the main areas of power systems that can benefit from the application of techniques related to big data is the development of renewable energy. We have also provided a comprehensive coverage of analytics of energy big data and security / privacy. In addition to the usual 3Vs energy big data challenges, we also covered the real-time and tight cyber-physical coupling features that are prominent in a smart grid. We also proposed an energy-big dataoriented taxonomy to better understand the complicated and intriguing relationships between different components, security issues and related solutions.

REFERENCES

[1] IDC, "The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things," 2014. [Online]. Available: https://www.emc.com/leadership/digital-universe/2014iview/executivesummary. htm.

[2] A. Ipakchi and F. Albuyeh, "Grid of the future," IEEE Power Energy Mag., vol. 7, no. 2, pp. 52–62, Mar. 2009.

[3] T. Garrity, "Getting Smart," IEEE Power Energy Mag., vol. 6, no. 2, pp. 38–45, Mar. 2008.

[4] P. Zhang, F. Li, and N. Bhatt, "Next-Generation Monitoring, Analysis, and Control for the Future Smart Control Center," IEEE Trans. Smart Grid, vol. 1, no. 2, pp. 186–192, Sep. 2010.

[5] (International Energy Agency) IEA, "Smart Grids – Technology Roadmap," 2011.

[6] S. Massoud Amin and B. F. Wollenberg, "Toward a smart grid: power delivery for the 21st century," IEEE Power Energy Mag., vol. 3, no. 5, pp. 34–41, Sep. 2005.

[7] K. Zhou, C. Fu, and S. Yang, "Big data driven smart energy management: From big data to big insights," Renew. Sustain. Energy Rev., vol. 56, no. January, pp. 215–225, 2016.

[8] IBM Corporation, "The Four V's of Big Data," 2016. [Online]. Available: http://www.ibmbigdatahub.com/infographic/four-vs-bigdata.

[9] Z. Huang, H. Luo, D. Skoda, T. Zhu, and Y. Gu, "E-Sketch: Gathering large-scale energy consumption data based on consumption patterns," in Proc. IEEE Int. Conf. Big Data (Big Data), Washington, DC, USA, 2014, pp. 656–665.

[10] J. Yin, P. Sharma, I. Gorton, and B. Akyoli, "Large-scale data challenges in future power grids," in Proc. IEEE 7th Int. Symp. Service Orient. Syst. Eng. (SOSE), Redwood City, CA, USA, 2013, pp. 324–328.

[11] M. Aiello and G. A. Pagani, "The smart grid's data generating potentials," in Proc. Federated Conf. Comput. Sci. Inf. Syst. (FedCSIS), Warsaw, Poland, 2014, pp. 9–16.

[12] S. Yin and O. Kaynak, "Big Data for Modern Industry: Challenges and Trends," Proc. IEEE, vol. 103, no. 2, pp. 143–146, 2015.

[13] D. Cai et al., "Electric Power Big Data and Its Applications," in International Conference on Enery, Power and Electrical Engineering (EPEE), 2016, pp. 181–184.

[14] IBM Corporation, "Managing big data for smart grids and smart meters," White Pap., p. 8, 2012.

[15] IBM Corporation, "Extracting business value from the 4 V's of big data," 2016. [Online]. Available: http://www.ibmbigdatahub.com/infographic/extracting-business-value-4-vs-big-data.

[16] E. P. I. S. C. CSEE, "White paper of China's electric power big data," 2013.

[17] S. Witt, "DATA MANAGEMENT & ANALYTICS FOR UTILITIES," SmartGrid Updat., 2014.

[18] D. G. Hart, "Using AMI to realize the Smart Grid," in 2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008, pp. 1–2.

[19] REN, "2016 Technical Data," 2017.

[20] N. P. da Silva, L. Rosa, W. Zheng, and R. Pestana, "Wind Power Forecast Uncertainty Using Dynamic Combination of Predictions," Period. Polytech. Electr. Eng. Comput. Sci., vol. 59, no. 3, pp. 78–83, 2015.

[21] S. Pelland, J. Remund, J. Kleissl, T. Oozeki, and K. De Brabandere, "Photovoltaic and Solar Forecasting: State of the Art," Int. EnergyAgency Photovolt. Power Syst. Program. Rep. IEA PVPS T14, pp. 1–40, 2013.

[22] J. Hu, H. R. Pota, and S. Guo, "Taxonomy of attacks for agentbased smart grids," IEEE Trans. Parallel Distrib. Syst., vol. 25, no. 7, pp. 1886–1895, Jul. 2014.

[23] N. Bui, A. P. Castellani, P. Casari, and M. Zorzi, "The Internet of energy: A Web-enabled smart grid system," IEEE Netw., vol. 26, no. 4, pp. 39–45, Jul./Aug. 2012.

ISSN: 0971-1260

Vol-22-Issue-17-September-2019

[24] D. Alahakoon and X. Yu, "Smart electricity meter data intelligence for future energy systems: A survey," IEEE Trans. Ind. Informat., vol. 12, no. 1, pp. 425–436, Feb. 2016, doi: 10.1109/TII.2015.2414355.

[25] D. A. Powner, "Electricity grid modernization: Progress being made on cybersecurity guidelines, but key challenges remain to be addressed, GAO report," United States Gov. Account. Office, Washington, DC, USA, Tech. Rep. GAO-11-117, Jan. 2011. [Online]. Available: http://www.gao.gov/new.items/d11117.pdf

[26] S. Simitis, "From the market to the polis: The EU directive on the protection of personal data," Iowa Law. Rev., vol. 80, no. 3, p. 445, 1994.

[27] J. Hu and A. V. Vasilakos, "Energy Big Data Analytics and Security: Challenges and Opportunities," IEEE Trans. Smart Grid, vol. 7, no. 5, pp. 2423-2436, Sept. 2016.