

Application of Big Data in Smart Grid

+Chun Sing Lai, *Student MIEEE*; *Loi Lei Lai, *FIEEE*

+Engineering Science, University of Oxford, UK; chun.lai@eng.ox.ac.uk

*State Grid Energy Research Institute, Beijing China; l.l.lai@ieec.org

Abstract—In this paper, the state-of-the-art of big data is reviewed. Challenges, opportunities and tools will be discussed. Some emerging technologies will be looked to promote big data applications. The applications of big data in smart grid in some countries will be summarized too.

Index Terms—Big data, smart grid

I. INTRODUCTION

The terms “big data” and “data analytics” can mean very different things, depending on the problems to be solved [1-4]. Many researchers and practitioners manage “big data” challenges by focusing on its high volumes of information rather than the information management. In general, big data is a term used to recognize the exponential growth, availability and use of information in the data-rich environment of today and beyond. The term “big data” puts a focus on the issue of information volume. Many researchers and practitioners manage “big data” challenges by focusing on its high volumes of information rather than the information management and this can lead to short-sighted decisions that will restrict the information architecture in which decision makers and managers try to expand and upgrade to meet changing needs for the society, environment and business. Too narrow a focus will force massive reinvestment to address other factors due to big data to reduce risks and increase costs.

Data volume is growing annually at a minimum rate of 59 percent worldwide annually. Although volume is a significant challenge in managing big data, data information variety and velocity must be focused on as well.

Volume: The increase in data volumes comes from traditional data types and new types of data. A very large size is a storage issue and too much data is also a massive analysis problem.

Variety: It is always a logical sequence to obtain data, derive information from it and then make decisions. There are now many more varieties of information to analyze. The varieties include tabular data, documents, metering data, video, images, audio, financial transactions and others.

Velocity: This involves streams of data, structured record creation, and availability for access and delivery. In some cases, real-time decision making is needed. Therefore it is important to have the required velocity to produce the data and process it to meet the need. It is essential to make sense of big

data and get patterns from it to help organizations to make better decisions.

Big Data technologies as a new generation of technologies and architectures designed to extract value economically from very large volumes of a wide variety of data by enabling high-velocity capture, discovery, and/or analysis. The total amount of data has grown exponentially, it has been estimated that more data was produced between 2010 and 2012 than in all of preceding human history [5].

The big data in Smart Grid is generated from various sources, such as (i) power utilization habits of users, (ii) phasor measurement data for situational awareness, (iii) energy consumption data measured by the widespread smart meters, (iv) energy market pricing and bidding data collected by automated revenue metering (ARM) systems, (v) management, control and maintenance data for devices and equipment in the power generation, transmission and distribution networks acquired by intelligent electronic devices, (vi) operational data for running utilities, such as financial data and (vii) very large data sets, not directly obtained through the grid measurement but widely used in decision making, such as weather data, data from the National Lightning Detection Network (NLDN), and Geographic Information System (GIS) data.

The application of big data in power utilities is just started. By 2020, the number of installed smart meters in Europe will reach 240 million while North America will have 150 million smart meters in use, China is forecasted to install about 400 million smart meter by that date. Japan would deploy about 60 million smart electricity meters and South Korea would plan to deploy between 500,000 and 1.5 million smart meters per year in homes before 2020. With so many smart meters being deployed, utilities’ data center will increase the amount of data by several TB per day. Research has shown that many of the current utilities have not fully explored the value of “big data”. However, serious attention has been given to this field and some well-known institutions have updated their teaching and research programme to education students who have can take up the challenges of big data research and applications in the near future. Data analytics competitors are competing to bring a set of IT tools and capabilities that are largely new to the utility industry.

II. RELATED TECHNIQUES AND TOOLS

Over the past few years, nearly all major companies, including IT giants like Oracle&IBM, grid giants such as General Electric, Siemens/eMeter, ABB/Ventyx&Schneider Electric/Telvent, and startups like AutoGrid, Opower, & C3. All have started their big data projects and are competing to bring a set of IT tools that are largely new to the utility industry. Scientists have developed a wide variety of techniques and technologies to capture, curate, analyze and visualize Big Data. Big Data needs powerful techniques to efficiently process very large volume of data within limited time. The techniques involve a number of disciplines, including statistics, data mining, machine learning, signal processing, pattern recognition, optimization methods and visualization method.

A. Techniques

Optimization methods have been applied to solve quantitative problems in a lot of fields, such as engineering, economics, physics, chemistry and biology. In smart grid area, to tackle communication problems [6-9] and data processing problems [10], a variety of new scalable and distributed architectures or frameworks have been proposed.

Data mining is a set of techniques to extract valuable information (patterns) from data, including clustering, classification and regression. For meter data analytics, data mining methods have been adopted. The applied methods can be grouped in two classes, namely statistical methods to make estimation based on historical data; and AI methods to model risk and uncertainty [11]. Recently developed data mining methods include fuzzy wavelet neural network [12] and nonparametric estimation [13]. Customers are grouped according to their electricity consumption patterns and/or other characteristics, such as activities. Clustering methods are applied in generating residential load profiles [14] and to differentiate pattern variations due to seasonal and temporal impacts [15]. Fuzzy decision tree is used to classify power quality disturbances [16]. Data mining approaches have been used for price classification [17].

Machine learning is an important area in artificial intelligence. It is used to discover knowledge and make intelligent decisions automatically.

Visualization method is used to create tables, images and diagrams to understand data.

B. Tools

One of the most famous and powerful batch process-based Big Data tools is Apache Hadoop. It provides infrastructures and platforms for specific Big Data applications in business and commerce.

For stream data applications, for example electric power system operation, would require real-time response for data processing platforms such as Storm which is designed especially for real-time stream data analytics.

For interactive analysis processing, the data is presented in an interactive environment. Users are directly connected to the computer and can interact with it in real time. Apache Drill is a distributed system for interactive analysis of Big Data. It has

the capability to process petabytes of data and trillions of records in seconds.

III. TECHNOLOGIES

The on-going or emerging technologies that are closely related to big data include cloud computing, Internet of things (IoT), granular computing, data center and quantum computing [18].

A. Cloud computing

Cloud computing delivers applications and services over the Internet. Cloud computing is closely related to big data. Big data puts stress on the storage capacity of a cloud system. The main objective of cloud computing is to use huge computing and storage resources to provide big data applications with fine-grained computing capacity [19]. The emergence of big data also accelerates the development of cloud computing. The distributed storage technology and the parallel computing capacity can effectively manage big data and improve the efficiency of acquisition and analysis of big data. CloudView is a framework for storage, processing and analysis of massive machine maintenance data in a cloud computing environment, which is formulated using the Map/Reduce model and reaches real-time response. Cloud computing promotes transferring and sharing data.

B. Internet of Things (IoT)

The Internet of Things (IoT) is the network of physical objects or "things" embedded with electronics, software, sensors and connectivity to enable it to achieve greater value and service by exchanging data with the manufacturer, operator and/or other connected devices.

The internet of things represents an enormous amount of networking sensors embedded into various devices and machines in real world. Such sensors deployed in different fields to collect various kinds of data, such as environmental data, geographical data, operation data, customer data in power utilities. Many devices including mobile equipments, home appliances could be used for data acquisition. IoT enables advanced applications for smart grid. The big data generated by IoT has special characteristics because of the different types of data collected, which have the characters of heterogeneity, variety, unstructured feature, noise, and high redundancy. It is expected that by 2030, the IoT data will be the most important part of big data, big data in IoT has three features that conform to the big data paradigm: (i) Huge amount of data is generated from terminals; (ii) Data generated by IoT is usually semi-structured or unstructured;(iii) The data processing capacity of IoT has fallen behind so it is extremely urgent to promote the development of IoT based on the introduction of big data technologies.

C. Granular computing

In granular computing (GrC), granules such as classes, clusters, subsets, groups and intervals are used to build computational models for complex applications with huge amounts of data, information and knowledge. Granular computing can reduce the data size into different level of granularity. However, not all the Big Data applications can

use the GrC techniques. It depends on the confidence and accuracy of results required. For example, power network sensor data needs to be processed and responded in time and with high accuracy for decision making.

D. Data center

The data center is in charge of acquiring, managing & organizing data, and leveraging the data values & functions. The emergence of big data brings development opportunities and great challenges to data centers. Organizations are experiencing rapid IT growth but their data centers are aging. International Data Corporate (IDC) puts the average age of a data center at nine years old. Gartner, another research company says data centers older than seven years are obsolete [20].

As the electric grid gets smarter, vast quantities of data arrive at utility companies, which promote the explosive growth of the infrastructure and related software of data center. The big data has more requirements on storage capacity and processing capacity, as well as network transmission capacity. Enterprises must take the development of data centers into consideration to improve the capacity of rapidly and effectively processing of big data under limited price/performance ratio.

E. Quantum computing

Quantum computing studies quantum computers that make direct use of quantum-mechanical phenomena, such as superposition, to perform operations on data [21]. Quantum computers are different from digital computers which require data to be encoded into binary digits. Quantum computation uses quantum bits.

The development of actual quantum computers is still in its infancy, but experiments have been carried out in which quantum computational operations were executed on a very small number of qubits. It is foreseen that large-scale quantum computers will be able to solve certain problems much more quickly than any classical computers that use even the best currently known algorithms such as integer factorization using Shor's algorithm. Small-scale quantum computers are existed, for example, D-Wave Systems Company developed their quantum computer, called "D-Wave one" with 128 qubits processor and "D-Wave two" with 512 qubits processor in 2011 and 2013 respectively.

IV. CHALLENGES AND OPPORTUNITIES

When handle Big Data problems, difficulties lie in data capture, transmission, processing, storage, searching, sharing, analysis and visualization. Data is increasing at exponential rate, but the improvement of information processing methods is relatively slow. In many important Big Data applications, the state-of-the-art techniques and technologies cannot solve the real-life problems practically, especially for real-time analysis. Some of them are listed below:

A. Uncertainty

Utilities are uncertain about the costs and requirements for building a big data analytics infrastructure. There is also uncertainty about how big data analytics will fit into the current systems. There is a lot of data and people do not know

how to manage, store it and at what point does it become useless. Data curation needs to be considered. It the data management through its lifecycle of interest and usefulness is to power utilities. Curation activities enable data discovery and retrieval, maintain quality, add value and provide for re-use over time. The existing database management tools are unable to process Big Data that grow so large and complex. The size of Big Data keeps increasing exponentially, but current capability to work with is only in the relatively lower levels of petabytes, exabytes and zettabytes of data. New framework for modeling uncertainty and predicting the change of the uncertainty is required.

B. Security

Most Big Data are stored in a distributed way, cloud data storage is often used. However, the network bandwidth capacity is the bottleneck in cloud and distributed systems, especially when the volume of communication is large. Smart meter installations have generated concerns about data privacy. Tremendous amounts of data about individuals, e.g. internet activity, energy usage, social interaction, are being collected and analyzed, which have the risk to cause damage to data provider.

Significant security problems include data security protection, intellectual property protection, personal privacy protection, commercial secrets, network security and financial information protection. Most developed and developing countries have already made related data protection laws to enhance the security.

C. Data quality

As the size of data set is very large, sometimes in the region of several gigabytes or more and also the data origin is from many sources, real-world databases include inconsistent, incomplete and noisy data. Therefore, a number of data preprocessing techniques, including data cleaning, integration, transformation and reduction need to be applied to minimize noise, inconsistencies and incompleteness in data. In many cases, the current techniques will be too slow to achieve a workable solution for real-life problems.

D. Data analysis and visualization

For real-time Big Data applications such as power system operation and protection, it is necessary to guarantee the time response requirement when the data volume is very large. Presently, it is a big challenge for stream processing. Big Data has encouraged the development of the hardware and software architectures such as the advancement in cloud computing which distributes multiple workloads into a large cluster of processors. In this direction, distributed computing is being developed at a very high speed. It is required to understand how Big Data works with real-time systems. however, it is not easy to link system state, measurements and network topologies together with respect to time. It is particularly difficult to conduct data visualization because of the high dimension and size of Big Data since the current Big Data visualization tools mostly have poor performance in functionalities, scalability and response time.

E. Data explosion

Data from the UK Department of Energy and Climate Change [22] are used to describe appliance ownership patterns. This data enables energy efficiency ratings to be accounted for. However, smart metering will create an explosion in data availability. For example, in the UK, the 27 million domestic electricity consumers currently just have over 100 million data points per year collected quarterly or half-yearly for energy suppliers to record, store, and use in billing and other business operation. When smart metering fully deployed and operated at 30-minute sampling rate, energy suppliers will need to ingest, store and process at least around 4500 to 9000 times more of the current data size, reaching 50 terabytes. To manage data sets in such a large volume, the main problem of using relational database management systems is its low scalability [23] and this require a scalable solution that can grow for practical use.

F. Lack of standards

To take advantage of these large new data sets, it is essential to gain access to data; develop the data management and programming capabilities to work with large-scale data sets. New approaches to summarize, describe, and analyze the information contained in big data must be developed [24]. Integration of many different forms of data from systems like meter data management, outage management, customer management, billing platforms and asset management is required.

Standards for data description and communication are essential. These facilitate data reuse by making it easier to import, export, compare, combine and understand data. Standards also eliminate the need for each data originators to develop unique descriptive practices [25].

The lack of worldwide industry standards around data from smart grids and meters could lead to concerns about sharing data with competitors, it also brings worry around data ownership, and concerns about data accuracy. A further worry is on data privacy and data protection.

G. Lack of talents

The shortage of talent will be a significant constraint to obtain values from Big Data. Big Data is expected to rapidly become a key element of competition internationally. This kind of specialist is difficult to educate as it takes many years to train Big Data analysts that must have strong mathematical background and related professional knowledge. Specialized resource is a critical success factor for better data management. Utilities are concerned about the shortage of available data specialists and will need to undertake efforts to locate or develop such talents. To ensure a continued supply of skills in the future, in-house training is a likely means but a talent war is likely to get bigger.

H. Current workers may be reluctant to use new technologies

V. BIG DATA APPLICATION EXAMPLES IN SMART GRID WORLDWIDE

Table 1 summarizes some potential application areas of big data in smart grid.

Table 1. Potential application areas of big data in smart grid

1	Prediction and analysis of economic situation and social impact
2	Development of scientific reasoning for decision making
3	Performance analysis for generation and storage systems
4	Load management with demand response and energy utilization & efficiency analytics
5	Consumer behavior analysis
6	Using AMI and intelligent electronic devices for state estimation
7	Pricing analytics and incentive implementation analysis
8	Grid infrastructures optimization
9	Demand and generation forecast under high uncertainties
10	Asset management
11	Service quality analytics

The key goals of U.S. utilities' big data efforts, mainly focused on converting the tens of billions of data points coming from the millions of smart meters deployed around the country and turning them into actionable information for the grid operations. For U.S. utilities most of the data analytics focus on improving grid reliability, outage response and lowering the cost of distribution operations.

European utilities can either be vertically integrated service providers for entire nations, or contenders in deregulated and competitive energy markets. For example, Italy and the Scandinavian countries have largely completed their smart meter rollouts while the UK, France, Germany are just getting started. For European utilities, they concentrate on integrating large-scale renewables & distributed generations and managing new loads such as plug-in electric vehicles.

A. EDF [26]

EDF has applied big data technology for weather forecasting for risk assessment and modeling the impacts of energy operations on the environment. As EDF begins its plan to roll out 35 million smart meters across France, it will incorporate this data into doing business. Smart meter data will be used to better estimate the condition of the grid, promote demand response, predict the lifetime of power lines, transformers and other equipment. Demand response is an important part of managing grid assets in a world where new customer loads such as plug-in electric vehicles and the increasing share of power being generated by customers via distributed generations are altering supply-demand balances.

B. E.ON [27]

In Sept 2013, IBM announced that it has been selected by E.ON Metering to operate its Smart Metering IT infrastructure in a private cloud. The new platform will improve the deployment and management of smart meters, simplify the integration of renewables, and other innovative services, while also allowing E.ON to deliver personalized services that will put customers in better control of energy usage. Customers will have the ability to view their usage profiles for information about time-of-use-rates and changes in use patterns that can be compared with historical data. In addition, the platform's scalability and low start-up and operation costs will provide the flexibility for future growth. Real-time data evaluation in a smart grid is growing increasingly important for energy utilities due to the increase use of intermittent renewable energy.

C. USA EXELON [28]

In 2014, Baltimore Gas & Electric (BGE) started a project in the hope to save hundreds of millions of dollars by using C3's cloud-based data analytics to manage operations of the 2 million smart meters deployed, and by tapping their data to discover and prevent energy theft and revenue losses. Trillions of data points are regressing and correlating against each other to extract values. BGE believes it can better detect, isolate and reduce meter tampering, and unbilled energy delivery.

BGE parent company Exelon also plans to apply the same C3 platform for projects at its Chicago-based Commonwealth Edison and Philadelphia-based PECO utilities, which are deploying a collective 8 million smart meters between them.

Integration of many different forms of data from systems like meter data management, outage management, customer management, billing platforms and asset management is required.

On the smart grid side, C3 is working with utilities including Southern California Edison, Northeast Utilities and Entergy.

It will be very difficult if not impossible to encode every rule for theft patterns, because they change over time. Machine learning is used where you have to learn from the data without being able to code an algorithm. C3 is a grid data analytics company to cite machine learning as part of its suite of tools, this is a cutting-edge claim to apply techniques as yet untested fully in the utilities.

D. Pacific Gas and Electric Company [29]

Smart meters are providing utilities with unprecedented amounts of energy data. Utilities across the US, including Pacific Gas & Electric Company (PG&E) in California, are revolutionizing the use of such data to empower their customers.

PG&E's groundbreaking work with its interval data analytics (IDA) program was developed to maximize the value of data captured by 9.4 million smart meters. PG&E is the largest US utility to install smart meters across its entire service territory.

PG&E captures hourly or sub-hourly reads from each home. With more robust data, PG&E is also able to provide detailed advice to help customers better manage their energy use.

PG&E's My Energy portal, which is available online to all customers, demonstrates the energy-saving benefits of personalized data and insights. Customers can examine their time-specific energy use, see how they compare within their neighborhood, understand how and why their consumption varies over time, set their usage goals, and, of course, discover more ways to save energy.

E. KEPCO [30]

In 2014, State-run electric utility Korea Electric Power Corp (KEPCO) launched two pilot projects on ways to use big data to improve demand management and risk forecasting. The first pilot set up an energy consulting business based on advanced metering infrastructure (AMI) data while the second established a risk forecasting system analyzing social networking service data.

The objective of the first project is to help customers save electricity by providing comparable data on energy usage for similar business types based on AMI big data, while allowing KEPCO to manage demand and reduce brownouts. The second project aims to analyze a variety of business risks including blackouts, customer complaints and climate change by merging information from social network service data, internet data and complaints. KEPCO has stated it aims to train 300 big data professionals by 2016.

To summarize, power utilities have diversity in the application of big data due to different, business structure, technology development etc. Having more completed operational data, most of the companies put the demand side response and user service in a higher priority. It can be seen that in general, power utilities are cooperating with IT companies to develop big data applications.

VI. CONCLUSIONS

Big data is still unclear for many power utilities. Smart grid operation and future energy management will be hugely data-intensive. There are many obstacles which affect the success of big data applications in smart grid. Presently experience in integrating big data with smart grid is limited. The utilities need to focus on turning the data that they collect into business intelligence to improve processes and customer experience. For example, the detected patterns of spatial and temporal electricity consumption can be used to help optimizing demand response management. This will lead to solutions that improve efficiency and enable alternative approaches to handle various aspects of the utility to drive company performance and minimize risks from new regulations and political interference under a low carbon economy.

ACKNOWLEDGMENTS

Financial support from State Grid Corporation of China is very much appreciated.

REFERENCES

- [1] L. Wiggins, "Bringing big data up to the big leagues", IBM Data magazine, 2013.
- [2] James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh and Angela Hung Byers, "Big data: The next frontier for innovation, competition, and productivity", McKinsey Global Institute, 2011, http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation (Visited on 1 April 2015)
- [3] Laney, Douglas, "3D Data Management: Controlling Data Volume, Velocity and Variety", Application Delivery Strategies, Meta Group, 2001. <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf> (Visited on 1 April 2015)
- [4] "Gartner Says Solving 'Big Data' Challenge Involves More Than Just Managing Volumes of Data", Gartner, Inc., June 2011. <http://www.gartner.com/newsroom/id/1731916> (Visited on 1 April 2015)
- [5] Mike Lynch, "Data wars: Unlocking the information goldmine", April 2012. <http://www.bbc.com/news/business-17682304> (Visited on 1 April 2015)
- [6] Y.-J. Kim, M. Thottan, V. Kolesnikov, and W. Lee, "A secure decentralized data-centric information infrastructure for smart grid", IEEE Communications Magazine, Vol. 48, No. 11, November 2010, pp. 58–65.
- [7] J. Zhou, R. Hu, and Y. Qian, "Scalable distributed communication architectures to support advanced metering infrastructure in smart grid", IEEE Transactions on Parallel and Distributed Systems, Vol. 23, No. 9, Sept 2012, pp. 1632–1642.
- [8] X. Fang, S. Misra, G. Xue, and D. Yang, "Managing smart grid information in the cloud: opportunities, model, and applications", IEEE Network, Vol. 26, No. 4, July 2012, pp. 32–38.
- [9] Z. Fan, P. Kulkarni, S. Gormus, C. Efthymiou, G. Kalogridis, M. Sooriyabandara, Z. Zhu, S. Lambotharan, and W. H. Chin, "Smart grid communications: Overview of research challenges, solutions, and standardization activities", IEEE Communications Surveys Tutorials, Vol. 15, No. 1, 2013, pp. 21–38.
- [10] M. Arenas-Martinez, S. Herrero-Lopez, A. Sanchez, J. Williams, P. Roth, P. Hofmann, and A. Zeier, "A comparative study of data storage and processing architectures for the smart grid", IEEE International Conference on Smart Grid Communications (SmartGridComm), Oct 2010, pp. 285–290.
- [11] C. Borges, Y. Penya, and I. Fernandez, "Evaluating combined load forecasting in large power systems and smart grids", IEEE Transactions on Industrial Informatics, Vol. 9, No. 3, Aug 2013, pp. 1570–1577.
- [12] M. Amina, V. Kodogiannis, I. Petrounias, and D. Tomtsis, "A hybrid intelligent approach for the prediction of electricity consumption", International Journal of Electrical Power and Energy Systems, Vol. 43, No. 1, pp. 99 – 108, 2012.
- [13] N. Ding, Y. Besanger, F. Wurtz, and G. Antoine, "Individual nonparametric load estimation model for power distribution network planning", IEEE Transactions on Industrial Informatics, Vol. 9, No. 3, Aug 2013, pp. 1578–1587.
- [14] W. Labeeuw and G. Deconinck, "Residential electrical load model based on mixture model clustering and markov models", IEEE Transactions on Industrial Informatics, Vol. 9, No. 3, 2013, pp. 1561–1569.
- [15] [15] A. M. Ferreira, C. A. Cavalcante, C. H. Fontes, and J. E. Marambio, "A new method for pattern recognition in load profiles to support decision making in the management of the electric sector", International Journal of Electrical Power and Energy Systems, Vol. 53, 2013, pp. 824 – 831.
- [16] M. Biswal and P. Dash, "Measurement and classification of simultaneous power signal patterns with an s-transform variant and fuzzy decision tree", IEEE Transactions on Industrial Informatics, Vol. 9, No. 4, 2013, pp. 1819–1827.
- [17] D. Huang, H. Zareipour, W. Rosehart, and N. Amjady, "Data mining for electricity price classification and the application to demand-side management", IEEE Transactions on Smart Grid, Vol. 3, No. 2, 2012, pp. 808–817.
- [18] C. P. Chen and C.-Y. Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on big data", Information Sciences, Vol. 275, pp. 314 – 347, 2014.
- [19] Intel, "Big data in the cloud: Converging technologies", Intel IT Center, September 2014.
- [20] <http://www.forbes.com/2010/03/12/cloud-computing-ibm-technology-cio-network-data-centers.html> (Visited on 10 April 2015)
- [21] http://en.wikipedia.org/wiki/Quantum_computing (Visited on 6 April 2015)
- [22] Department of Energy and Climate Change (DECC), "Energy Consumption in the UK: Domestic Data Tables", UK, Tech. Rep. URN 12D/270, 2012.
- [23] M. Ferguson, "Architecting a big data platform for analytics", IBM Data magazine, 2012.
- [24] Liran Einav and Jonathan Levin, "Economics in the age of big data", Science, Vol. 346 Issue 6210, 7 November 2014.
- [25] Clifford Lynch., "How do your data grow?", Nature, Vol. 455, September 2008, pp. 28-29.
- [26] EDF's Big Data Vision for France <http://www.greentechmedia.com/articles/read/edfs-big-data-vision-for-france> (Visited on 11 April 2015)
- [27] <https://www-03.ibm.com/press/us/en/pressrelease/41921.wss> (Visited on 4 April 2015)
- [28] <http://www.greentechmedia.com/articles/read/c3-energy-unveils-first-big-test-of-smart-grid-data-analytics> (Visited on 8 April 2015)
- [29] Aaron Tinjum, "PG&E is revolutionizing how utilities use data to empower their customers — one smart meter read at a time", May 9, 2014 <http://blog.opower.com/2014/05/pg-e-is-revolutionizing-how-utilities-use-data-to-empower-their-customers-one-smart-meter-read-at-a-time/> (Visited on 10 April 2015)
- [30] Metering International, "KEPCO pilots big data projects for AMI and customer service systems", July 14, 2014. <http://www.metering.com/kepco-pilots-big-data-projects-for-ami-and-customer-service-systems/> (Visited on 11 April 2015)