

## THE USE OF BIG DATA FOR OUTAGE MANAGEMENT IN DISTRIBUTION SYSTEMS

Po-Chen CHEN  
Texas A&M University – USA  
pchen01@neo.tamu.edu

Tatjana DOKIC  
Texas A&M University – USA  
tatjana.djokic@neo.tamu.edu

Mladen KEZUNOVIC  
Texas A&M University – USA  
kezunov@ece.tamu.edu

### ABSTRACT

*The focus of this paper is an innovative use of the data correlation framework of big data analytics for improved outage management in distribution networks. Extended types of data sources and their correlation to the network model are described. Benefits of improved data analysis in view of the big data uses are discussed through few examples.*

### I. INTRODUCTION

The U.S. Department of Energy's (DoE) initiative "GRID 2030" has described a new vision of future electric grid that will benefit from advancements in measuring equipment and information technologies [1]. The correlation of the data collected from all Intelligent Electronic Devices (IEDs), as well as data coming from additional sources such as weather reporting system, lightning detection network, geographic information systems, etc, often called Big Data (or Large Data), has become of interest to the utility industry.

Numerous applications of big data for distribution system applications were foreseen in [2]. As an example, active machine learning technique may be used to analyze large amount of data [3]. Examining the syntax and semantic framework of big data analytics focused on outage management (OM) in distribution networks is foreseen to be the top priority for future grid performance optimization [4], [5]. In this paper, the data sources are specified and the methodology for

correlating big data to an expanded power system model is demonstrated.

### II. BACKGROUND: DATA SOURCES USED BY POWER INDUSTRY

In general, two groups of data used by the utility industry can be distinguished: one coming from the utility measurement infrastructure, and the other from other sources not necessarily being part of utility measurement infrastructure.

- Electric utility measurement data:
  - Synchrophasors also known as the Phasor Measurement Unit (PMU) data,
  - Smart meter data,
  - Data from other types of Intelligent electronic devices (IEDs),
  - Condition-based asset management data,
  - Supervisory control and data acquisition (SCADA) data.
- Data coming from additional sources:
  - Geographical information system (GIS) data,
  - Global positioning system (GPS) time-reference data,
  - Weather and lightning data,
  - Seismic data,
  - Animal migration data,
  - Electricity market data.

As part of utility infrastructure, the SCADA system scans consists of a set of measurements acquired at the substation, such as bus voltages, frequency, transformer tap position, and breaker status signals. Energy management system (EMS) collects the data every two to ten seconds. IEDs, such as Phasor Measurement Unit (PMU), Digital Protective Relay (DPR), Digital Fault Recorder (DFR), Sequence of Event Recorder (SER), and others, provide an additional extensive set of data that have much higher sampling rate than SCADA. In order to ensure interoperability of these devices, unified

---

This work was supported in part by NSF IUCRC "Power Systems Engineering Research Center" and in part by NSF Engineering Research Center (ERC) Center for Ultra-wide-area Resilient Electric Energy Transmission Networks (CURENT).

representation of data and model is done through the standard formats for data exchange such as IEC 61850-6 Substation Control Language or SCL used for correlating substation measurements and model [6], IEEE COMTRADE used for exchange of transient data [7], IEEE COMFEDE for exchange of event data [8], and Common Information Model or CIM for correlating power system model and SCADA data [9], [10].

### III. THE USE OF BIG DATA FOR DISTRIBUTION SYSTEM OM

Most of the outages are caused directly or indirectly by severe weather conditions. Understanding how these effects affect the network faults, as well as historical trends of when and where they are occurring for specific area can help in (1) developing more predictive estimates of the fault occurrence and (2) more detailed analysis of associated faults and operation of the protection equipment. The OM approach can be differentiated into outage mapping and precise FL technique. The outage mapping is a combination of techniques which are applied to narrow down the faulted area. Fault location (FL) techniques utilize field measurement data to precisely locate the fault to the affected feeder and its lateral. The different types of data should improve each of these two categories.

Fig. 1 shows the overall process of how the big data integration may be utilized to improve outage management solution (OMS). The use of the following data analytics as they are the key to the future implementation will be discussed next:

- FL techniques with various limitations,
- Customer calls and smart meter data
- GIS data,
- Weather data,
- DG/EV data,
- Condition-based asset management data.

#### A. The Use of Different FL Techniques

Fault Location (FL) is an essential application in OMS. It has a direct impact on determining the reliability of the feeder service and quality of electricity supply for customers. Improving the reliability of the network can be interpreted as improving the accuracy of FL methods. The accurate FL in distribution networks faces difficulties due to the following reasons.

- Complexity of distribution networks:
  - Large number of components,
  - Heterogeneous impedance of the feeders,
  - Existence of many feeder laterals and load taps,
- Significant changes in loading conditions over time,
- Unbalanced three phase system.

The implementation of precise FL using field data measurements collected from IEDs, which is referred as “Smart Fault Location” as discussed in [11], where a detailed literature survey can be found. FL algorithms require modeling of a system and/or signals, and are highly dependent on a quality of these models. Data coming from additional sources can be used to improve a simulation capability offering better understanding of what has happened, and a possibility to evaluate models and algorithms taking into account many different factors that were linked to an event.

#### B. The Use of Customer Calls and Smart Meter Data

Utilities usually apply SCADA systems to monitor substations in order to detect substation outages and/or large feeder outages. To localize outages, utilities mostly depend on customer calls [12]. Particularly inefficient is the situation when faults occur during late hours or during the nighttime when customers are not aware and hence do not place trouble report calls. During outages, 30% of customers call more than once because they do not know when their power will be restored [13]. During the process, utilities still have to

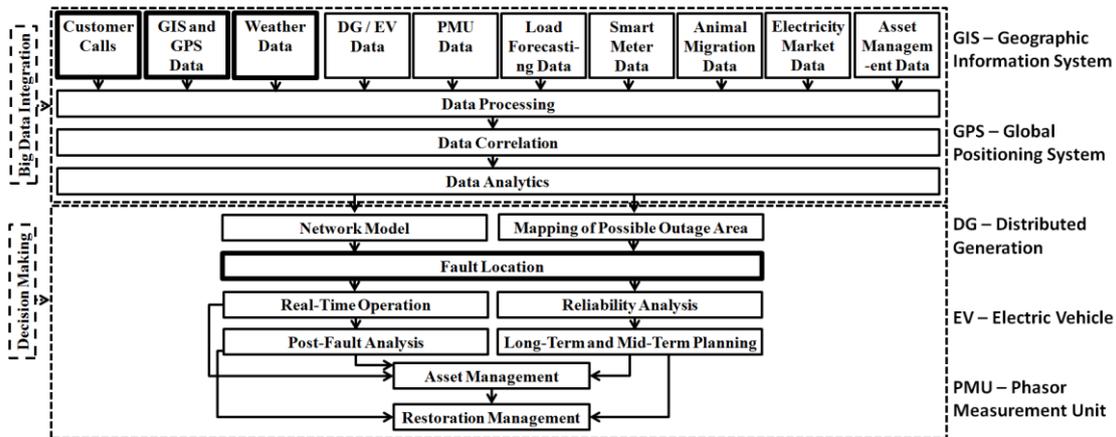


Fig. 1 Overall process of how the big data integration will improve OMS.

manage a significant number of customer inquiries (where some may be repetitive), which usually overwhelms the personnel. Therefore, interactive voice response (IVR) system is offered as a solution [14]. The virtual voice platform of IVR can support inbound and outbound interactions (i.e. phone, text, emails) and provide automation for organizing customers' report, which is more efficient and economical.

In addition, Advanced Metering Infrastructure (AMI) for processing smart meter data has become a viable option for outage reporting [15], [16]. The utilities can rely on AMI to know if there is an outage or whether the restoration is successful. With the use of AMI, the affected area with the fault can be detected promptly and based on a network model it may be possible to estimate the most likely location of a fault.

### C. The Use of GIS Data

The Geographic Information System (GIS) data is typically used to represent maps as data layers that can be used to correlate electrical networks to geographical locations [17]. The GIS data can be distinguished in two aspects: spatial and attribute data [18]. The spatial data provides the absolute and relative location of geographic features, such as coordinates of location where substation is situated. The attribute data provides

the characteristics of spatial features; the most common format is a table with information about specific attribute. In case of substation, attribute data can contain information about all the equipment located inside the substation.

The key to the data correlation is the integration of GIS and GPS with model of a power system. GPS provides time references that can be used to synchronize all events. Modern digital measuring devices such as PMUs, traveling wave fault locators, and lightning detectors and locators often have integrated GPS unit that is sending precise timestamp together with measured data. This information can be used to provide temporal correlation for different measurements that are linked to the fault. On the other hand, GIS model of a system can be correlated with an electrical model providing a better prospective on geographical characteristics of a system [19]. Conveniently, other sources of data such as weather, seismic, and animal migration data, also use GIS for their geographical presentation. Thus, spatial correlation of data coming from different sources can be done by including every set of data as a separate layer of a GIS. In this case, the utility's enterprise GIS database is designated for storing and maintaining the GIS data, where GIS

management is essential in keeping the network model up-to-date.

#### **D. The Use of Weather Data**

The following weather data including different ambient conditions is another important source for utility personnel to perform outage mapping [20], [21]:

- Lightning,
- Ice, snow, sleet, and blizzard,
- Freezing, fog, and frost,
- Wind and gale,
- Solar heat,
- Rain.

The other weather conditions such as earthquake, tornado, flood, and hurricane, they may not happen quite often but the damage is usually significant. For instance, in the end of Oct. 2012, Hurricane Sandy hit New York City in the USA and became the worst hurricane ever to hit the service territory of Consolidated Edison Company of New York (Con Ed). More than one million customers lost their electricity and the largest customer restoration efforts in Con Ed's history was undertaken [22], [23].

The historical weather data with post-fault analysis can provide an insight for various types of nature-caused events. For instance, thunderstorms (wind and rain) tend to produce different types of system interruptions than ice and snow [20]. Based on the weather data, the utilities can have a predicted preview of possible outage mapping. As another example, as reported in [20], the tree trimming may be the most common factor of storm-related outages. Therefore, the utilities should be prepared for the outages in the long-time unattended tree-trimming zone when a storm comes.

#### **E. The Use of DG/EV Data**

A large scale integration of distributed generation (DG) together with electric vehicle (EV) can be foreseen in

the future [24], [25]. The voltage quality and voltage stability are essential with DG and EV deployments [26]-[28]. The existence of DGs and EVs in a given service area may impact the fault location and outage management approaches [27]-[29]. It has also been shown that islanding of DG, as well as the EV connections in the grid to building (G2B) mode can provide continued electricity supply support to the customer during the outages [28], [30]. Still, how to manage interfacing of DGs and EVs for the proper intentional islanding and power restoration respectively is not fully understood, particularly under large DG and EV penetrations.

#### **F. The Use of Condition-based Asset Management Data**

The condition-based asset management data, specifically the data used for the planning and scheduling of system components maintenance, is also of great importance for an effective OM. This type of data helps in maintaining the availability of the critical components and system reliability consequently [31]. As an example, condition monitoring data of the circuit breakers (CB) may be used to predict the risk of fault clearing or feeder restoration being successful, which in turn may be incorporated in the overall outage management strategy. The CB reliability assessment example points to the ability to assess failure rate probabilities based on historical condition-based measurements, and as a consequence proper maintenance strategies may be developed in anticipation of the needs to have an efficient and cost effective OM, particularly system restoration [32].

### **IV. FUTURE SOLUTIONS**

Emerging from the discussions in Section III, the following Big Data analytics are proposed:

- Application of smart fault location techniques,
- Correlation of customer calls (i.e. IVR) with advanced fault location techniques,
- Outage reporting via AMI, where this information is correlated to the fault location techniques,
- Maintenance of utilities' enterprise GIS database to keep the distribution network model up-to-date,
- Estimation of outage mapping using weather, lightning, animal migration, seismic and other data.

The use of additional data sets offers new capabilities by providing: possibility to predict the sensitive area of a network and schedule proactive measures, allow better situational awareness during the outage and faster response with better precision of a fault location, and offer better understanding of specific factors that caused the outage during post-fault analysis.

## V. CONCLUSION

This paper makes several contributions:

- The correlation framework of big data analytics is examined and its importance is illustrated.
- Various data sources and their correlation to the network model are described.
- Benefits of improved data analytics for outage management using the big data are discussed
- Few examples of future applications are mentioned and the role of big data for improvement of these applications is discussed.

## REFERENCES

- [1] United States Department of Energy, "GRID 2030, A National Vision for Electricity's Second 100 Years."
- [2] M. Kezunovic, *et al.*, 2013, "The role of big data in improving power system operation and protection," *Bulk Power Syst. Dynamics and Control - IX Optimization, Security and Control of the Emerging Power Grid (IREP) Symp.*
- [3] V. Malbasa, *et al.*, 2013, "Power System Online Stability Assessment Using Active Learning and Synchrophasor Data," *IEEE Grenoble PowerTech (POWERTECH)*.
- [4] M. Smith, *et al.*, 2012, "2012 Annual Grid Analytics Report," *Utility Analytics Institute*.
- [5] Bridge Energy Group, 2014, "Utilities with Outage Management as a Top Priority Excelling in Customer Service Ratings." [Online]. Available: <http://www.bridgeenergygroup.com/news/press/utilities-outage-management-top-priority-excelling-customer-service-ratings/>
- [6] IEC Standard 61850-6, "Communication networks and systems in Substations – Configuration description language for communication in electrical substations related to IEDs".
- [7] IEEE Standard C37.111-1999, "IEEE Standard Common Format for Transient Data Exchange (COMTRADE) for Power Systems".
- [8] IEEE Standard C37.239, "IEEE Standard for Common Format for Event Data Exchange (COMFEDE) for Power Systems".
- [9] IEC Standard 61970, "IEC 61970 Energy management system application program interface (EMS-API) - Part 301: Common Information Model (CIM) Base".
- [10] IEC Standard 61968, "IEC 61968 Application integration at electric utilities - System interfaces for distribution management - Part 11: Common Information Model (CIM)."
- [11] M. Kezunovic, 2011, "Smart fault location for smart grids," *IEEE Trans. Smart Grid*, vol. 2, no. 1, 11-22.
- [12] United States Department of Energy, "The Smart Grid: an introduction."
- [13] 2012, "Business agility in outage management for energy and utilities," *IBM Software Energy Utilities*.
- [14] J. Kuru, *et al.*, 2009, "Innovative system integration for outage communication," *The 20th Int. Conf. and Exhibition on Electricity Distrib. - Part 2, CIRED*.
- [15] S.-H. Choi, *et al.*, 2011, "The design of outage management system utilizing meter information based on AMI (Advanced Metering Infrastructure) system," *IEEE 8th Int. Conf. on Power Electronics and ECCE Asia (ICPE & ECCE)*, 2955-2961.
- [16] H. Tram, 2008, "Technical and operation considerations in using Smart Metering for outage management," *IEEE/PES Trans. and Distrib. Conf. and Expo. (T&D)*.
- [17] Esri, 2013, "What is GIS?" [Online]. Available: <http://www.esri.com/what-is-gis>
- [18] D. J. Buckey, 1997, *The GIS Primer: An Introduction to Geographical Information System*, Pacific Meridian Resources, Inc., Colorado, USA.
- [19] P. H. Smith, 2005, "Electrical Distribution Modelling: An Integration of Engineering Analysis and Geographic Information Systems", *Master Thesis*, Virginia Tech.
- [20] R. J. Campbell, 2012, "Weather-Related Power Outages and Electric System Resiliency," *Congressional Research Service (CRS) Report for Congress*.
- [21] L. McColl, *et al.*, 2011, "A Climate Change Risk Assessment for the UK Electricity Networks," presented at *2011 Int. Conf. Energy & Meteorology (ICEM)*.
- [22] Official news of Consolidated Edison Company of New York. [Online]. Available: [http://www.coned.com/newsroom/news/pr20121108\\_4.asp](http://www.coned.com/newsroom/news/pr20121108_4.asp)
- [23] Official news of Consolidated Edison Company of New York. [Online]. Available: <http://www.coned.com/newsroom/news/pr20121112.asp>
- [24] P.-C. Chen, *et al.*, Oct. 2012, "Analysis of Voltage Profile Problems due to the Penetration of Distributed Generation in Low-Voltage Secondary Distribution Networks," *IEEE Trans. Power Del.*, vol. 27, no. 4, 2020-2028.
- [25] M. Kezunovic, 2013, "Electric Vehicles Could Offer More Gain than Drain," *Electric Energy T&D Magazine*, vol. 17, no. 6, 18-20.
- [26] P.-C. Chen, *et al.*, 2013, "Analysis of Voltage Stability Issues with Distributed Generation Penetration in Distribution Networks," *45th North American Power Symp. (NAPS)*.
- [27] C. Pang, *et al.*, 2012, "BEVs/PHEVs as Dispersed Energy Storage for V2B Uses in the Smart Grid," *IEEE Trans. Smart Grid*, vol. 3, no. 1, 473-482.
- [28] M. Kezunovic and S. Travis Waller, 2012, "Vehicle-Power Grid Interaction for Multi-fueling (EBV, PHEV)," Chapter on *Encyclopedia of Sustainability Science and Technology*, Springer.
- [29] S. M. Brahma, 2011, "Fault location in power distribution system with penetration of distributed generation," *IEEE Trans. Power Del.*, vol. 19, no. 1, 56-63.
- [30] P. Fuangfoo, *et al.*, 2007, "Impact Study on Intentional Islanding of Distributed Generation Connected to a Radial Subtransmission System in Thailand's Electric Power System," *IEEE trans. Ind. Appl.*, vol. 43, no. 6, 1491-1498.
- [31] M. Kezunovic, *et al.*, 2009, "Integration of Asset and Outage Management Tasks for Distribution Application," *Power Systems Engineering Research Center (PSerc) Final Project Report*.
- [32] P. Dehghanian, *et al.*, 2012, "Critical Component Identification in Reliability Centered Asset Management of Distribution Power Systems via Fuzzy AHP", *IEEE Syst. J.*, vol. 6, no. 4, 593-602.