Dae-Hyun Choi
cdh8954@neo.tamu.edu

A Framework for the Impact Analysis of Data Quality/Integrity/Privacy in Cyber-Physical Electric Energy System

1. Background/Motivation

"***Even small changes in the data could affect the stability of the grid and even jeopardize human safety***" [1].

Managing and understanding the cyber-physical energy grid data is an increasing concern for Independent System Operators (ISOs), utilities, Load Serving Entities (LSEs) and market participants. As more and more cyber sensing devices such as smart meters (distribution level) and synchrophasors (transmission level) are deployed to the power system, it is prudent for future electric grid operators to understand the fundamental impact of *multi-scale* spatial **data quality** on resilient physical and market operations. In addition, **data integrity** is closely related to cybersecurity that may occur in the power industry. For instance, data manipulated by an intelligent adversary may result in misleading system operators and smart grid control systems, thus resulting in the severe malfunction of physical and financial grid operations. Furthermore, maintaining smart meter **data privacy** is essential to design a successful demand response program and a secure advanced metering infrastructure (AMI). Built upon our preliminary success of the development of analytical platform for data analysis in supervisory control and data acquisition (SCADA) systems, we propose frameworks and algorithms to analyze and design *robust cyber-physical electric energy systems* with the presence of bad/malicious data.

2. Proposed Research/Work

We plan to design new operational tools, architecture, and algorithms for cyber-physical electric energy systems. The approach is to (1) develop a unified system-wide monitoring tool for multi-scale spatial grid data quality analysis, (2) create a resilient multi-area state estimation architecture and sensing/communication system to mitigate the risk of data integrity attack, and (3) develop a novel data privacy-preserving algorithm and infrastructure to prevent malicious monitoring. By applying advances in *large data acquisition and processing*, *distributed estimation, cyber security, optimization, and power engineering/economics,* proposed research is conducted along the following three directions related to **data quality**, **data integrity, and data privacy**:

- ***Data Quality-Aware Multi-Scale Decision Making in Cyber-Physical Power Systems***

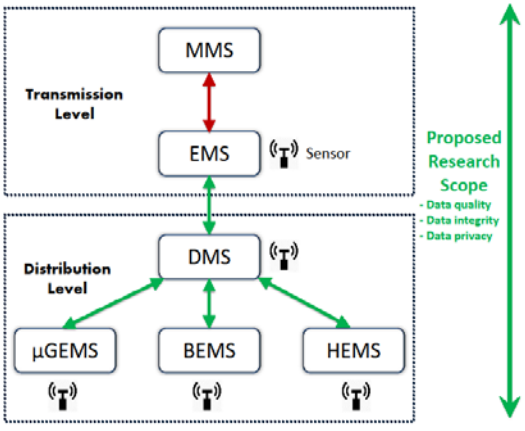

Fig. 1. Hierarchical Architecture for Future Grid Operations

In Fig. 1, I plan to develop a unified decision-making framework to investigate the impact of multi-scale spatial data quality at the transmission and distribution level on energy management system (EMS) and market management system (MMS) operations. Data in the transmission level refer to sensing for physical grid condition (e.g., power flow, voltage phasor, network topology), electricity market operation (e.g., market participant's bidding price), and control command (e.g., open/close of circuit breakers, AGC signal). Data in the distribution level refer to smart metering (e.g., individual

2014 NSF Young Professional Workshop on Exploring New Frontiers in Cyber-Physical Systems

[1] Data Management and Analytics for Utilities [Online]. Available: http://www.smartgridupdate.com/dataforutilities/pdf/DataManagementWhitePaper.pdf

energy consumption), distributed renewable solar and wind generation, and energy storage/electric vehicle (e.g., location information). A possible solution will be to adopt sensitivity based-KKT condition perturbation approach for rigorous understanding of the fundamental coupling among multi-level EMS, MMS, DMS and lower level management systems. The proposed sensitivity index is the combination of the coupled partial derivatives quantifying the sensitivity of parameters in each management system.

- ***Data Integrity-Resilient State Estimation***
  Multi-area state estimation is becoming increasingly popular as the interconnected power system consists in multiple subsystems. An interconnected power system is operated by multiple control centers based on their administrative boundaries (e.g., New England, New York). However, malicious measurement data in one administrative area may contaminate the performance indices of other administrative areas due to the physical coupling among the subsystems. Therefore, it is important to assess which region may be affected by any given malicious data. The concept of error residual spread area (ERSA) is proposed to decompose the whole system into several non-overlapping regions, in which the corrupted data only contaminate measurements within each region. ERSA decomposition relies on network topology and measurement configuration. The proposed ERSA decomposition will localize and prevent data integrity attacks from spreading to a remote location. I am interested in developing a new architecture and algorithm for *attack-resilient multi-area state estimation* based on ERSA decomposition. Research in this area will have great potential for interdisciplinary collaborations with researchers in power system engineering and statistical signal processing.
- ***Data Privacy-Preserving Design of Advanced Metering Infrastructure***
  The massive deployment of smart meters in the distribution network raises a series of concerns on the loss of privacy. Since metering data of individual homes/buildings is accumulated every 15 min, it is possible to infer the pattern of electricity consumption of individual users. I am interested in developing an algorithm and protocol for privacy-preserving smart metering to protect the privacy of individual users while preserving the ability of distribution utility to compute the current electricity consumption. Gaussian Mixture Model (GMM) based on the distributions of actual meter data and corrupted data for privacy protection will be considered. Statistical inference algorithms such as expectation maximization (EM) technique can help distribution utility to estimate the distribution of aggregated actual meter data. This research has a potential to generate new paradigms of secure and privacy-preserving distribution network design from interdisciplinary collaborations with researchers in power system engineering and cyber security.

3. Potential Impact on CPS

The proposed research introduces a novel and resilient operating paradigm against bad/malicious data for cyber-physical electric energy system operations. The proposed analytical frameworks will provide input for the development of visualization tools on the level of multi-scale spatial data quality as well as system operators with a unified view on the impact of data quality on physical and economical operations of the future grid. The proposed architecture and algorithms for maintaining data integrity and data privacy will provide system operators with analytical tools to detect potential cyber data attacks and protect consumer privacy from malicious monitoring. Overall, the proposed research will lay a foundation for the development of a next-generation management system infrastructure for the future cyber-physical electric energy system that is reliable, economical and secure.