Electric Power Systems Research 133 (2016) 51–62

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

Anomaly detection in electric network database of smart grid: Graph matching approach

Adnan Anwar*, Abdun Naser Mahmood

School of Engineering and Information Technology (SEIT), The University of New South Wales Australia, Canberra, ACT 2610, Australia

ARTICLE INFO

Article history:
Received 28 July 2015
Received in revised form
19 November 2015
Accepted 7 December 2015
Available online 5 January 2016

Keywords:
Smart grid
Database anomaly
Graph matching
Graduated assignment

ABSTRACT

Recent studies have shown that the operational modules of an Energy Management System (EMS) are vulnerable to the anomalies that exist in an electric topological and configuration database (DB). In this paper, we focus on the security of EMS modules by detecting anomalies in an electric network DB. Firstly, we explain how an EMS's Optimal Power Flow (OPF) module can be exploited by accidental or deliberate changes in a power system model. As a defense mechanism, for the first time, we propose a graph comparison-based approach for identifying anomalies in an electric network DB. In this study, we formulate the problem as a Quadratic Assignment Problem (QAP) and use the Graduated Assignment algorithm to perform graph matching. To evaluate the effectiveness of the proposed method, we consider different test scenarios considering the IEEE benchmark 24-bus, 30-bus and 118-bus test systems. The results obtained from this analysis show that the proposed method successfully captures DB anomalies at very high detection rates with a smaller time complexity than those obtained from studies published in relevant literature.

1. Introduction

In recent years, the importance of smart grid security has been highlighted due to several cyber incidents, particularly in the aftermath of the infamous *Stuxnet worm* [1] attack in 2010. The 'Industrial Control Systems Cyber Emergency Response Team (ICS-CERT)' has reported that, of the 257 cyber incidents that occurred in the sectors of the world's critical infrastructures in 2013, the maximum (approximately 145, that is, about 56%) were in the energy sector [2]. Realizing the importance and significance of smart grid security, researchers around the world are trying to identify smart grid vulnerabilities and potential strategies for protecting against them. An Energy Management System (EMS), which can be treated as the brain of a smart grid, also demonstrates vulnerability to corrupted measurement signals [3–9]. Specifically, researchers have proven that the operational modules of an EMS (e.g., state estimators, optimal power flow tools, etc.) can be compromised in such a way that a cyber attack remains undetectable using existing security measures [4].

While research on physical energy systems has a long history, that focusing on the cyber-security of an energy system is in its early stage. For the first time, a new class of cyber attack in an energy system, called 'False Data Injection (FDI) Attacks', is proposed in [4]. In this fundamental work, Liu et al. demonstrate the construction of a stealthy injection attack which remains hidden in an energy system's State Estimation (SE) module and affects the calculations of its state variables. A discussion of such kinds of FDI attacks can be found in [3]. Further enhancements of FDI attacks on the SE module are reported in [5–8]. While the above mentioned research works focus on how to develop these types of unidentifiable attacks, other studies relate to the development of defense models for protecting the SE module [10–15]. In an EMS, there are other modules which also play vital roles in operational decisions and any error in them will cause a serious operational failure and huge financial loss. One such operational module is the Optimal Power Flow (OPF) module which is used for economic dispatch and market analysis. Any wrong decision it makes will affect an energy grid's economics and produce misleading operating limits which will result in technical instability of the system. Although significant numbers of research works have been conducted based on the security of the SE module [5–8,10–15], considerably less attention has been paid to protecting the OPF module [16,17]. Teixeira et al. show that the OPF module of an EMS is vulnerable under maliciously biased estimates of measurement data [18]. The impact of an undetectable FDI attack on the OPF module is also investigated in [19,20]. It is important to mention that all these works consider that the OPF module is compromised due to the falsified measurement signals which bypass the SE module. In addition to the Bad Data Detection

[*] Corresponding author. Tel.: +61 451001357.
E-mail address: Adnan.Anwar@adfa.edu.au (A. Anwar).

http://dx.doi.org/10.1016/j.epsr.2015.12.006
0378-7796/© 2015 Elsevier B.V. All rights reserved.

(BDD) facility in the SE module, other defence strategies for protecting a grid against measurement data corruption are proposed in the literature [10–15]. The security issues of the OPF discussed in [18–20] can be protected by adopting a proper defense strategy, as discussed in [10–15]. However, complete security of the OPF module is not yet ensured as this module also needs inputs from other sources which may be corrupted. To elaborate, the OPF module also requires the power system's topological and configuration data for its operation which are typically stored in a DB. Therefore, any anomaly in this DB will produce misleading operational decisions which may critically affect the physical energy grid. It is worthwhile mentioning the following two issues which need attention in order to secure the operations of an energy system.

(1) Typically, the OPF module receives data once the operations of the SE and BDD modules have finished. After a successful SE/BDD operation, a deliberate change in the DB of the power system model (which is a cyber intrusion, as discussed thoroughly in [16]) will produce an erroneous OPF output and result in wrong operational decisions which can even lead to a severe failure [16,17]. A DB can also be changed due to other types of anomalies, e.g., erroneous inputs by a human operator.
(2) In addition to historical data, power system planning modules also need topological and configuration data of the *power system model* for decision-making purposes. Any error/anomaly that exists in the DB of the model will lead to a wrong planning decision which will not only affect the power system's economics but also degrade its reliability and stability.

Moreover, other operational modules of an EMS (e.g., Automatic Generator Controller (AGC), Contingency Analysis module, etc.) also need information of the power system model. Therefore, the focus of this work is to protect a power system's DB by detecting any anomalies that exist in it. In the most closest research, Valenzuela et al. propose an anomaly-based Intrusion Detection System (IDS) for ensuring data integrity which uses Principal Component Analysis (PCA) [16]. A similar problem is solved by Mousavian et al. [17] whereby the output power from a pre-trained Artificial Neural Network (ANN) and OPF module are compared to detect intrusions. Both these methods need real-time load data to calculate the most recent OPF output that is used for comparison against trained normal behavior during anomaly detection. In a practical network, load data may not be available for each node due to a communication failure or its corruption caused by injection attacks [21]. Besides, *demand uncertainty* is very common in power system operation. In those cases, those methods show their weakness. On the other hand, our proposed method does not need any OPF or power flow analysis; therefore, there is no need of any demand data. As a result, the demand uncertainty will not be able to affect the performance of the proposed anomaly detection technique. Although the methods of [16,17] secure the OPF operation, they cannot protect other operational modules of an EMS which also show vulnerabilities to corruption of its DB. In an EMS, the *Network Topology Processor* is responsible for converting the *power system data* into a bus/branch model (similar to a graph (node, edge) representation of a power system) which is widely used in different operational modules [22]. Therefore, we propose a graph matching-based anomaly detection technique which utilizes this bus/branch model. A reference graph is always updated based on system events (e.g., when system reconfiguration is performed) and stored in the anomaly detection modules. An anomaly is detected by comparing the reference graph against the query graph that is built every time prior to the operation of any EMS module. This proposed anomaly detection technique is executed in two stages: (i) performing a proper node correspondence between the two graphs, which is formulated as a Weighted Graph Matching Problem (WGMP) and solved as a QAP; and (ii) once pair-wise node correspondence is determined, using well-established similarity indices to detect any anomalies if exist. Unlike the methods in [16,17], which can protect only the OPF module, our proposed method are able to secure all EMS operational modules that are vulnerable to corruptions in a power system network DB, as discussed in detail in Section 3. Through extensive experiments considering different scenarios using multiple benchmark test systems, we show that the proposed method outperforms existing anomaly detection methods [16,17] which are used to solve similar problem. The proposed method also enhances the capability of post-anomaly processing. Although graph matching has been widely applied

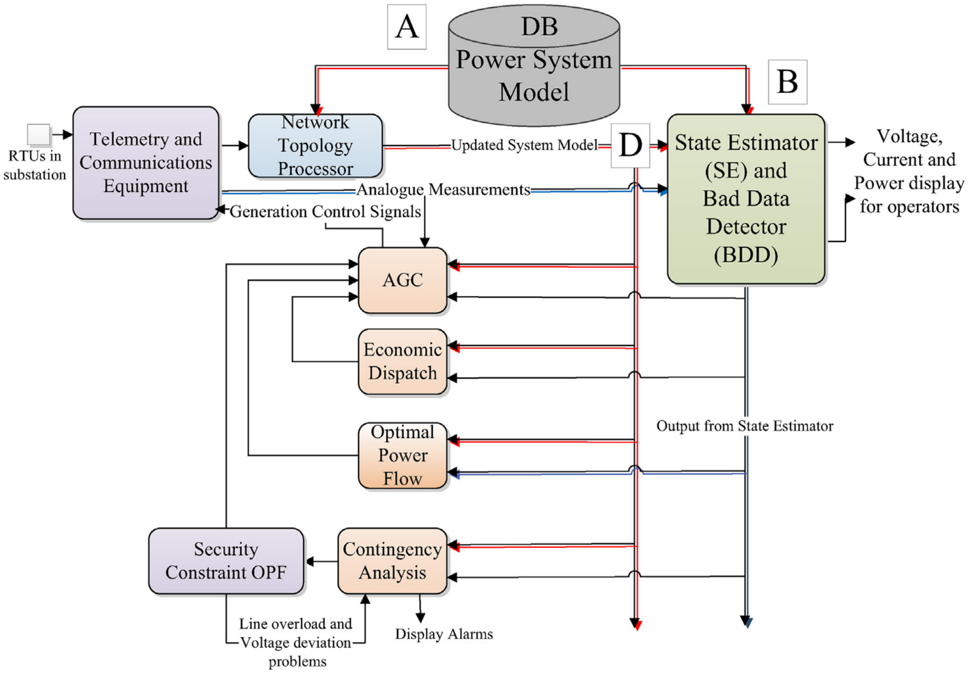

Fig. 1. Control centre architecture.