Classification and Detection of PMU Data Manipulation Attacks Using Transmission Line Parameters

Seemita Pal, Biplab Sikdar and Joe Chow

***Abstract*—Modern power grids are increasingly relying on real-time data, such as those from Phasor Measurement Units (PMUs), for their control and management operations. Due to its dependence on the Internet for data transfer, the grid is susceptible to a wide range of cyber attacks. Among these, data manipulation attacks are of particular interest in the context of PMU data, due to their potential for causing widespread damage. In such attacks, the adversary changes the measurements in order to bias the estimate of system states. In this paper we propose an effective and simple-to-implement mechanism for detecting such attacks. The proposed methodology is based on evaluating the equivalent impedances of transmission lines. Being independent of the conventional bad data detection scheme, it is also able to detect the so called "false data injection attacks". Extensive simulation results using real PMU data have been provided in order to verify the accuracy of the proposed detector.**

I. Introduction

PMUs are one of the most important grid-monitoring tools which provide dynamic visibility of the state of the power system. PMUs measure the voltage and current phasors as well as the frequencies at the buses where they are installed. The data generated by the PMUs are used along with SCADA other sources of information for various power system applications like state estimation, optimal power flow, real-time congestion control etc. This data, also known as synchrophasor data, is extremely critical since the control centers may base their control decisions either directly on these measurements or the output of the various applications.

The importance of synchrophasor data for the operation and management of smart grids makes them an attractive target for malicious attacks. An attacker may access and modify the synchrophasor data in three ways: by attacking the PMUs, by tampering with the communication network, or by breaking into the synchrophasor system through the control center office [1]. If a data manipulation attack is suspected yet not confirmed, it can create distrust of the measurements leading to confusion regarding the true states of the system and thereby hamper observability. Thus reliable techniques for detecting possible manipulation of measurements are essential. An undetected attack may obscure the control center from impending problems or mislead it into taking erroneous actions. It may cause uneconomic dispatch choices, congestion, failure of generators, failures of transmission lines, as well as cascading failures leading to blackouts.

This work was supported primarily by the ERC Program of the National Science Foundation and the Department of Energy under NSF Award Number EEC-1041877 and the CURENT Industry Partnership Program.

The authors are with Rensselaer Polytechnic Institute, Troy, NY, USA and the National University of Singapore, Singapore.

While power systems have traditionally accounted for the presence of "bad" data in measurements (e.g. due to malfunctioning instruments), malicious modifications of data are a relatively newer concern. While bad data detection techniques are effective at handling measurement, communication and structural errors, it has been shown that it is possible to construct data manipulation attacks that bypass bad data detection techniques. Existing work on detecting data manipulation attacks assume that at least some of the PMUs are secure, and focus on obtaining the minimum number of such secure PMUs (and their placement) that are required in order to detect an attack. However, the assumption that a PMU provides accurate data at all times without any possibility of being compromized by attackers is not realistic.

In this paper, we propose a mechanism for detection of PMU data manipulation attacks. The proposed mechanism is based on continuously monitoring the equivalent impedances of transmission lines and classifying observed anomalies for detecting the presence and location of attacks. Using this method, data integrity can be tested in a distributed and non-iterative manner, thereby requiring less memory and processing, and making early detection using legacy systems possible. Also, this is completely independent of traditional bad data detection schemes and there are no requirements of having a set of PMUs or state variables that are immune to attacks and available for verification purposes.

The rest of paper is organized as follows. The related work is presented in Section II. In Section III, we define the threat model. Section IV introduces the scheme for detection of PMU data manipulation attacks. Section V describes the detection mechanism in detail and provides the mathematical validation of the methodology. In Section VI, simulation results are provided for verifying the effectiveness of the detector. Finally Section VII concludes the paper.

II. Related Work

State estimation, which is a part of the Energy Management System (EMS), is used to obtain the best estimates of the system states (bus voltage magnitudes and phase angles) using network topology and available measurements [2]. Conventional state estimation techniques inherently have bad data detection mechanisms that use the redundancy in the available measurement data to detect any gross errors caused due to sensor problems or telemetry failures. In conventional bad data detection techniques, the 2-norm of the difference between the observed measurement vector and the estimated states is compared against a threshold to detect the presence

of bad measurements [3], [4]. Let $x = (x_1, x_2, \cdots, x_n)^T$ and $z = (z_1, z_2, \cdots, z_m)^T$ denote state variables and meter measurements respectively, where n is the number of state variables and m is the number of meter measurements and $m \geq n$. Therefore, presence of bad measurements is inferred if, $||z - Hx|| > \tau$, where H is an $m \times n$ full rank system matrix to allow estimation of x from z and τ is the threshold. Although this technique is quite effective against gross errors, it fails to detect highly structured bad data that conforms with the system topology and some applicable physical laws. This concept of false data injection attacks in smart grids was first introduced in [5]. Let $z_a = z + a$ represent the vector of observed measurements that contains the attack vector $a = a_1, a_2, \cdots, a_m$. If a is a linear combination of the column vectors of H, that is $a = Hc$, then it has been shown that the manipulated measurements can pass the bad data detection test [5]. So given a set of compromised or targeted PMUs, the attacker may be able to construct an attack vector that is a linear combination of the column vectors of H, which thereby makes the attack undetectable.

Based on this result, the authors of [1] introduced indices that quantify the least effort needed by attackers to achieve attack goals while avoiding bad data detection. The proposed method of protection against such attacks is placement of strategically placed meter measurements in grid [6], [7]. The authors in [7] have investigated the false data injection attacks from an operators point of view in order to determine how to defend against such attacks. It was shown that it is a necessary but not a sufficient condition to protect at least a certain number of measurements in order to be able to ensure observability of the system, and detect false data injection attacks. A Bayesian framework that leverages the knowledge of prior distribution on the states to detect false data injection attacks was proposed in [8]. The problem of determining the smallest number of meters that need to be tampered by the attacker has been modeled as an optimization problem in [9], [10]. In [9] a greedy algorithm is proposed for selecting a subset of the measurements in the system where secure PMUs can be placed. Giani et al. defined irreducible attacks, their conditions, and proposed an algorithm based on graph theory for finding all irreducible attacks [11]. In [12] the sensitivity of real and reactive power measurement residuals to false data injection attacks in a nonlinear state estimator is presented and a detector is proposed based on topology perturbation of monitored parameters [13]. In [14], a mechanism based on evaluation of the equivalent line impedances for detecting data manipulation attacks in synchrophasor data has been proposed. However, it has comparatively higher detection delays and lower detection accuracy. The authors in [15] propose an mixed integer programming based optimal response scheme that is to be followed in order to mitigate risk when certain PMUs are detected to be compromised and the threat is suspected to have propagated in the PMU network. In [16], GPS spoofing attacks on PMUs are presented and a cross-layer detection mechanism is proposed which is based both on the physical layer information as well as the power grid measurements. In this paper, however, we focus on manipulation of electrical data generated by PMUs and we have not considered GPS spoofing attacks.

Existing work on detection of false data injection attacks is focused on two aspects. Firstly, looking at the problem from the attackers point of view, they suggest techniques for determining the minimum number of PMUs that the attacker needs to access and corrupt in order to influence the state variables without raising an alarm. Conditions for making the attack unobservable are also suggested in this case. Secondly, looking at the problem from the defenders point of view, schemes for determining the minimum number of trustworthy or secure PMUs required for being able to detect the attacks has been proposed. In some cases the possible placement of these PMUs have been also suggested. It is quite clear that most of the cases assume that the attacker has limited resources or the defenders have some number of absolutely secure PMUs. But in reality, that may not be the case. No PMU can be expected to provide absolutely accurate data and zero possibility of corruption by attackers at all times. Therefore, detection techniques are required that can detect false data injection attacks without requiring such assumptions.

III. Threat Model

The threat model assumed in this paper is that the adversary has compromised one or more of the PMUs, PDCs, network routers or/and communications links. At each of the compromised nodes, the adversary has the ability to manipulate PMU data in order to bias the power system state estimates (the data may be unencrypted or the encryption broken). By doing so, the control center may be influenced into taking suboptimal dispatch decisions or wrong control actions, leading to the adversary's monetary gains, outages and/or damages.

Under the adversary model described above, the following cyber attack is considered. We consider a scenario where PMU data is carried in packets from the PMU to the PDC, then on to the Super PDC, and finally to the control center, via a number of intermediate routers and communication links. It is assumed that an adversary compromises one or more of these mentioned nodes or links in the network and manipulates the PMU data in the packets. To maximize the damage, the objective of the adversary is to manipulate data to the maximum extent possible without detection [17]-[21]. The data manipulated by the attacker changes the estimated system states from their true values and larger biasing is more likely to lead to erroneous actions of greater consequence. However, even relatively small changes can cause uneconomic dispatch choices or billing manipulation. Our objective is to develop a mechanism that will effectively detect such PMU data manipulation attacks.

The adversary may manipulate the following measurements: (1) *current magnitude,* (2) *current angle,* and (3) *voltage angle*. Since the voltage magnitude is always expected to be around 1 p.u. (per unit), manipulation of this data will readily raise the 'attack' alarm. Hence, it is expected to be avoided by the attacker. We consider two kinds of attacks: ramp and step. In ramp attacks, the attacker slowly and monotonously changes the data from its original value to make detection difficult. In step attacks, the attacker abruptly changes the PMU data to influence the operators into taking immediate control actions which may be damaging.

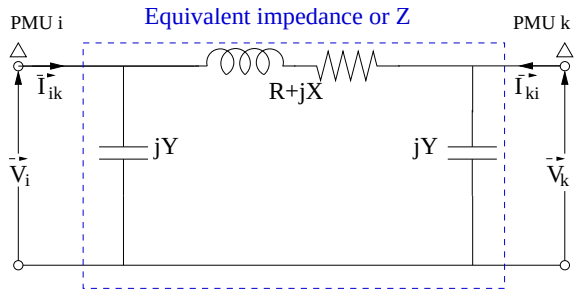

Fig. 1. Example topology.

IV. Basis of Detection Scheme

In this section, the various line parameters are first discussed followed by the basis of the scheme for differentiating normal PMU data from manipulated ones.

A. Transmission Line Parameters

Electrical transmission lines are represented by four electrical parameters: resistance (R), inductance (L), capacitance (C) and conductance (G). The resistance is affected by three factors, namely, temperature, frequency, and spiraling, and accounts for the thermal losses in the line. The inductance is due to the voltage induced by the magnetic flux changes caused by the changing conductor current. It depends on the line geometry, cable size and configuration and is the most dominant line parameter. The capacitance in transmission lines is present due to the potential difference between the conductors. For short transmission lines (length less than 50 miles), the effect of capacitance is negligible. The conductance is caused by the leakage current over the surface of the insulators. Since the leakage at insulators of overhead lines is negligible, the conductance is usually neglected.

The four parameters are uniformly distributed along transmission lines. But in most cases, they may be modeled using a lumped parameter configuration, without much loss of accuracy [22]. For a short line (length less than 50 miles), the series impedance, which is the resistance and the inductive reactance in series, in the lumped form is a good approximated model for the total line length. A medium line (between 50 and 150 miles) is modeled as a nominal π-circuit, wherein in addition to the series impedance, half of the capacitance to neutral of the line is lumped at each end. For long transmission lines (longer than 150 miles), the parameters may be considered to be distributed uniformly along the length of the line, if high level of accuracy is required. However, a nominal π-circuit may represent it sufficiently [22].

The estimation of the individual parameters of the approximated π-circuit model of a transmission line is an iterative process. The estimation results are not satisfactory when measurement errors or noise are present, when the load is unbalanced, or when mutual couplings exist on untransposed lines. Hence, detection based on individual line parameters may be unreliable. Thus in the proposed detection methodology, instead of estimating the individual line parameters, the equivalent impedance of the line is directly computed using the voltage and current measurements at the two end buses.

B. Detection Based on Equivalent Impedance Monitoring

Consider a power system with N buses, labeled as $i = 1, 2, \cdots, N$. The PMUs measure the voltage and all the currents incident on the bus where it is installed. Using these measurements on either side of each transmission line, the regional PDC or Super PDC can compute the equivalent impedance of the line. As shown in Figure 1, consider a transmission line which has PMUs at its two end buses (buses i and k). The measured bus voltage magnitudes and phase angles are denoted by V_x and θ_x, respectively, where $x = i, k$. The magnitude and the phase angle of the current flowing from bus i to bus k (respectively k to i) are denoted by I_{ik} and δ_{ik} (I_{ki} and δ_{ki}), respectively. In order to compute the equivalent impedances, we first arbitrarily pick one of the buses as the reference bus. The phase angle of the reference bus is then subtracted from all the phase angle measurements to obtain their phase angles with respect to the reference bus. The vector value of the equivalent impedance of the line as seen from bus i, $\vec{z_{ik}}$, is then calculated using:

$$\vec{V_i} = V_i(\cos\theta_i + j\sin\theta_i) \tag{1}$$

$$\vec{I_{ik}} = I_{ik}(\cos\delta_{ik} + j\sin\delta_{ik}) \tag{2}$$

$$\vec{z_{ik}} = (\vec{V_i} - \vec{V_k})/\vec{I_{ik}}. \tag{3}$$

Similarly, $\vec{z_{ki}}$ can be computed using the voltages and the current as measured at bus k.

Although the magnitudes of $\vec{z_{ik}}$ and $\vec{z_{ki}}$ may be slightly different due to instrumentation errors in the current transformers, potential transformers or PMUs, the trend observed in both should be the same. Similarly, although the angles of the equivalent impedances computed at both ends are not exactly the same, they too follow a similar pattern. Although slight changes are expected in the equivalent impedance due to changes in temperature, sagging etc., they occur gradually and typically remains within an expected range. If the magnitudes or angles of the equivalent impedances as seen from both ends show significant variation in trend, data manipulation may be suspected with a high level of confidence. This observation is the key to the proposed attack detection mechanism.

The ratio of the equivalent impedance magnitudes of a line between buses i and k, r_{ik}, and the difference in the angles of their equivalent impedances, d_{ik}, are given by:

$$r_{ik} = z_{ik}/z_{ki} \tag{4}$$

$$d_{ik} = arg(\vec{z_{ik}} - \vec{z_{ki}}). \tag{5}$$

Sample plots of the magnitudes of the equivalent impedances calculated at the two ends of a transmission line, and their ratio, are provided in Figure 2. These are based on real PMU values measured in the transmission system of New York, USA. It can be seen that r_{ik} has minimal variation when the measurements are not manipulated. A sample plot of the equivalent impedance angles at both ends of the line and their difference are shown in Figure 3. It can seen that d_{ik} too varies minimally under normal conditions. Thus, the equivalent impedance magnitudes, angles, and the quantities r_{ik} and d_{ik}, provide viable means of detecting manipulation of PMU data.

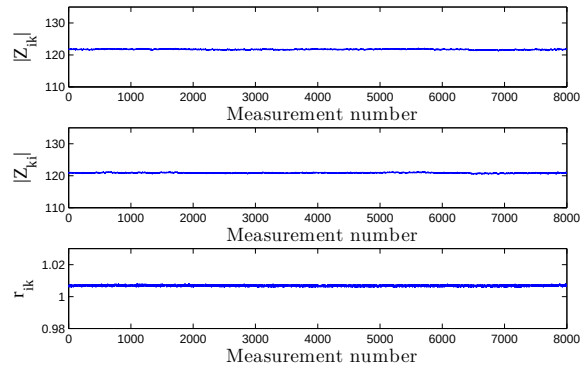

Fig. 2. Magnitude of equivalent impedances calculated from both ends and their ratio.

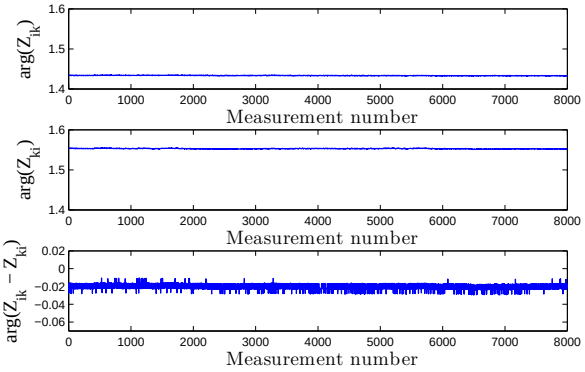

Fig. 3. Angles of equivalent impedances calculated from both ends and angle of difference of equivalent impedances.

V. Detection Mechanism Description

This section describes the proposed detection mechanism in detail. First, various features observed when different measurement quantities are manipulated are shown mathematically. Based on these features, a mechanism for detecting and classifying the various kinds of attacks is presented.

A. Data Manipulation and Deviations in Monitored Quantities

As mentioned in Section III, a data manipulation attack may modify the current magnitude, current angle or voltage angle. Based on the type of measurement manipulated, a specific subset of the monitored impedance-related quantities will show corresponding deviations, as described below.

1) Current Magnitude: The current magnitude may vary widely in normal situations due to changes in load, generation, and routes. Thus in this case, the difficulty in distinguishing between normal system variations and variations caused by an attacker's manipulation is the greatest. Let us assume that the attacker changes the magnitude of current measured at bus i, I_{ik}, by a factor p, so that the manipulated current is given by,

$$I'_{ik} = pI_{ik}. \tag{6}$$

Hence the computed equivalent impedance magnitude as seen from bus i will be impacted while that seen from bus k will remain unchanged. The observed impedances are:

$$z'_{ik} = \frac{|\vec{V_i} - \vec{V_k}|}{I'_{ik}} = \frac{z_{ik}}{p} \qquad \text{and} \qquad z'_{ki} = z_{ki} \tag{7}$$

The ratio of the magnitudes of the equivalent impedances or r_{ik} will therefore deviate from 1 and is given by:

$$r'_{ik} = \frac{z'_{ik}}{z'_{ki}} = \frac{z_{ik}/p}{z_{ki}} = \frac{1}{p}. \tag{8}$$

Similarly, if I_{ki} is changed by a factor of p, r_{ik} will become:

$$r'_{ik} = \frac{z'_{ik}}{z'_{ki}} = \frac{z_{ik}}{z_{ki}/p} = p. \tag{9}$$

With current magnitude manipulation, the angles of the equivalent impedances calculated on both sides remain unaffected. However, the angle of the difference of the equivalent impedances, d_{ik}, changes. Let the true equivalent impedances at buses i and k respectively be given by $\vec{z_{ik}}$ and $\vec{z_{ki}}$. The complex expression of the equivalent impedance difference is:

$$\vec{z_{ik}} - \vec{z_{ki}} = \frac{(I_{ik}e^{j\delta_{ik}} + I_{ki}e^{j\delta_{ki}})(V_i e^{j\theta_i} - V_k e^{j\theta_k})}{I_{ik}I_{ki}e^{(\delta_{ik} + \delta_{ki})}} \tag{10}$$

and d_{ik} is the argument of the above expression. Let the attacker inject a value x to the current magnitude measured at bus i. Therefore, the manipulated current value becomes:

$$I'_{ik} = I_{ik} + x. \tag{11}$$

Let $\vec{z'_{ik}}$ be the changed equivalent impedance calculated at bus i. The equivalent impedance calculated at bus k is unaffected and given by $\vec{z_{ki}}$ as before. As in (10), the equivalent impedance difference after the manipulation is given by:

$$\vec{z'_{ik}} - \vec{z_{ki}} = \frac{(I'_{ik}e^{j\delta_{ik}} + I_{ki}e^{j\delta_{ki}})(V_i e^{j\theta_i} - V_k e^{j\theta_k})}{I'_{ik}I_{ki}e^{\delta_(ik+\delta_{ki})}}. \tag{12}$$

The changed angle of equivalent impedance difference after the current magnitude manipulation is then given by:

$$d'_{ik} = arg(\vec{z'_{ik}} - \vec{z_{ki}}). \tag{13}$$

Let us represent the deviation in the value of d_{ik} by Δd_{ik}. Thus, this can be expressed as:

$$\Delta d_{ik} = d'_{ik} - d_{ik} = arg\left(\frac{\vec{z'_{ik}} - \vec{z_{ki}}}{\vec{z_{ik}} - \vec{z_{ki}}}\right). \tag{14}$$

Substituting (10) and (12) in (14) and simplifying, we get:

$$\Delta d_{ik} = arg\left(1 + \frac{xe^{j\delta_{ik}}}{I_{ik}e^{j\delta_{ik}} + I_{ki}e^{j\delta_{ki}}}\right). \tag{15}$$

Since I_{ik} and I_{ki} are the currents measured at either ends of the same line, their magnitudes will be very close in value. The small difference will be due to the capacitive current flowing between the lines. So for simplicity of analysis, if we assume $I_{ik} = I_{ki} = I$, the expression becomes:

$$\Delta d_{ik} = arg\left(1 + \frac{x}{I(e^{j(\delta_{ki} - \delta_{ik})} + 1)}\right). \tag{16}$$

Let $(\delta_{ki} - \delta_{ik})$ be given by α. Expressing the argument in (16) as the difference of the arguments of the numerator and the

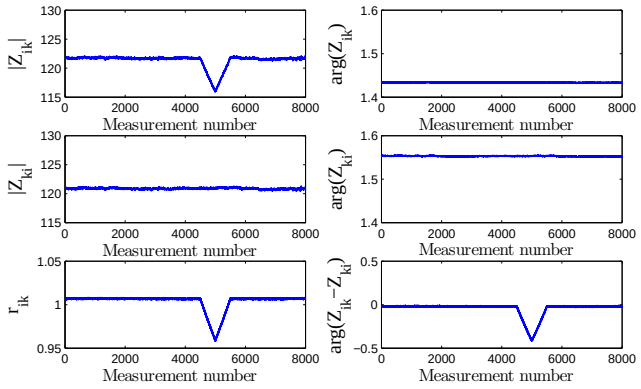

Fig. 4. Computed quantities when ramp attack is executed on current magnitude measured at bus i.

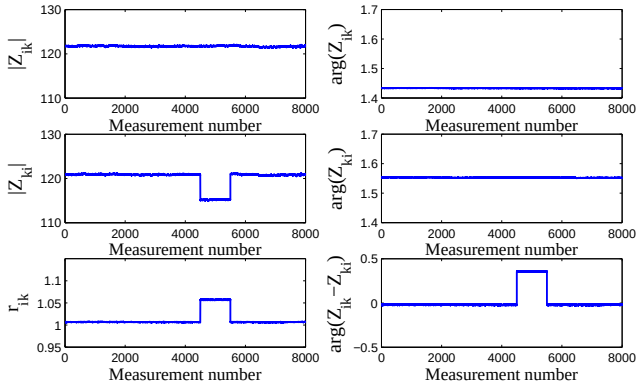

Fig. 5. Computed quantities when step attack is executed on current magnitude measured at bus k.

denominator and converting to inverse tan expression, we get:

$$\Delta d_{ik} = \tan^{-1}\left(\frac{I\sin\alpha}{I+x+I\cos\alpha}\right) - \tan^{-1}\left(\frac{I\sin\alpha}{I+I\cos\alpha}\right), \tag{17}$$

and after a series of algebraic operations, we have:

$$\Delta d_{ik} = \tan^{-1}\left(\frac{-\tan(\alpha/2)}{\frac{2I_{ki}}{x}+1}\right). \tag{18}$$

Similarly, if the current magnitude measured on the other bus, i.e., I_{ki} is increased by x, the deviation in the angle of the equivalent impedance difference will be given by:

$$\Delta d_{ik} = d'_{ik} - d_{ik} = \tan^{-1}\left(\frac{\tan(\alpha/2)}{\frac{2I_{ki}}{x}+1}\right). \tag{19}$$

The impact of current magnitude modification on our monitored quantities is shown in Figures 4 and 5. In Figure 4, the magnitude of I_{ik} has been linearly increased upto 5% between measurements 4500 and 5000 and then decreased between 5000 and 5500, executing a ramp attack. In Figure 5, a step attack has been executed by abruptly increasing the magnitude of I_{ki} by 5%. The other measurements are not manipulated. It can be seen that there are deviations in the magnitude of the equivalent impedance of the manipulated bus, the ratio of the impedance magnitudes, and the angle of the equivalent impedance difference, corresponding to the manipulation.

2) Current Angle: In the second case, the attacker may change the current angle in order to mislead the control center regarding the power factor of the load. Let the current angle at one of the end buses, say δ_{ik}, be changed by a factor of q:

$$\delta'_{ik} = q\delta_{ik}. \tag{20}$$

Let γ be the angle of phasor $(\vec{V_i} - \vec{V_k})$. Then, the original angle will be given by,

$$arg(\vec{z_{ik}}) = \gamma - \delta_{ik}. \tag{21}$$

After the attacker's manipulation, the changed angle of the equivalent impedance as seen from bus i will be given by,

$$arg(\vec{z'_{ik}}) = \gamma - q\delta_{ik}. \tag{22}$$

The angle of the equivalent impedance seen from the other bus, as well as the magnitudes of the equivalent impedances calculated on both sides and their ratio remain unchanged.

As in the current magnitude attack, the angle of the difference between the equivalent impedances will change from the true value. Let the current angle measured at bus i be increased by x, i.e.,

$$\delta'_{ik} = \delta_{ik} + x. \tag{23}$$

The true value of the equivalent impedance difference is given by (10). After the manipulation, the equivalent impedance calculated at bus i will change to $\vec{z'_{ik}}$, while that seen from bus k remains unaffected. Hence, after current angle manipulation, the difference of the equivalent impedances is given by:

$$\vec{z'_{ik}} - \vec{z_{ki}} = \frac{(I_{ik}e^{j\delta'_{ik}} + I_{ki}e^{j\delta_{ki}})(V_i e^{j\theta_i} - V_k e^{j\theta_k})}{I_{ik}I_{ki}e^{(}\delta'_{ik} + \delta_{ki})}. \tag{24}$$

Therefore, the deviation observed in the value of d_{ik} will be given by (14) and can be expressed as:

$$\Delta d_{ik} = -x + arg\left(\frac{I_{ik}e^{j\delta'_{ik}} + I_{ki}e^{j\delta_{ki}}}{I_{ik}e^{j\delta_{ik}} + I_{ki}e^{j\delta_{ki}}}\right). \tag{25}$$

As in the previous case, if we substitute $(\delta_{ki} - \delta_{ik})$ by α and for simplicity approximate $I_{ik} = I_{ki} = I$ in the above expression and simplify, we get,

$$\Delta d_{ik} = \frac{-x}{2}. \tag{26}$$

Similarly, it can be shown that if the current angle measured at bus k, δ_{ki}, is increased by x, a similar deviation but of opposite sign is observed in the value of d_{ik}, and:

$$\Delta d_{ik} = \frac{x}{2}. \tag{27}$$

The impact of current angle manipulation is shown in Figures 6 and 7. It can be seen that during the attack, there are deviations in the monitored quantities conforming to our derived expressions.

3) Voltage Angle: The attacker may manipulate the voltage angle measurements to mislead the control center about the phase angle states at those buses. When all measurements are true, let $\vec{V_d}$ be the phasor $(\vec{V_i} - \vec{V_k})$. Then the magnitude of

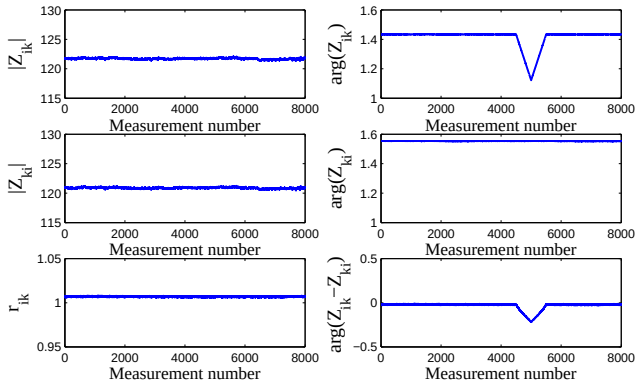

Fig. 6. Computed quantities when ramp attack is executed on current angle measured at bus i.

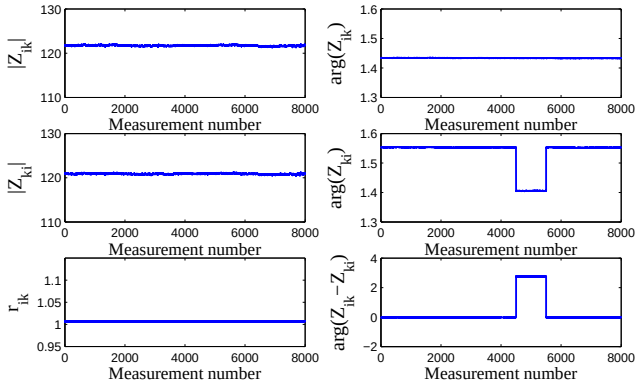

Fig. 7. Computed quantities when step attack is executed on current angle measured at bus k.

the equivalent impedance as seen from bus i is given by,

$$z_{ik} = \frac{V_d}{I_{ik}}, \tag{28}$$

and the angle of the equivalent impedance is given by,

$$arg(\vec{z_{ik}}) = arg(\vec{V_d}) - \delta_{ik}. \tag{29}$$

Let one of the voltage angles (say θ_i) be changed as follows:

$$\theta_i' = r\theta_i. \tag{30}$$

When the voltage angle θ_i is changed, the corresponding voltage difference phasor, i.e. $\vec{V_i'} - \vec{V_k}$, changes. Let it be denoted by $\vec{V_d'}$. It is quite apparent that the magnitude as well as the angle of this phasor deviates when any of the voltage angles is manipulated, i.e.,

$$V_d' \neq V_d \qquad \text{and} \qquad arg(\vec{V_d'}) \neq arg(\vec{V_d}).$$

Hence, the magnitude of the equivalent impedance changes to,

$$z_{ik}' = \frac{V_d'}{I_{ik}}. \tag{31}$$

The angle of the corresponding equivalent impedance also changes to,

$$arg(z_{ik}') = arg(V_d') - \delta_{ik}. \tag{32}$$

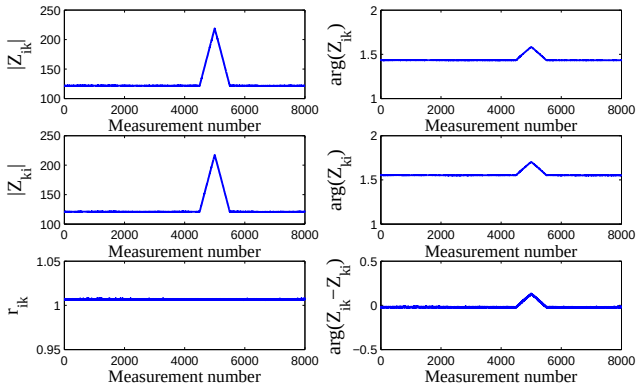

Fig. 8. Computed quantities when ramp attack is executed on voltage angle at bus i.

As in the previous cases, the angle of the equivalent impedance difference or d_{ik} also changes. If the attacker adds a value of x to the voltage angle, that is,

$$\theta_i' = \theta_i + x, \tag{33}$$

then the true value of the equivalent impedance difference is as given by (10). After the manipulation, the equivalent impedances calculated at buses i and k change. Therefore, the changed value of their difference is given by:

$$z_{ik}' - z_{ki}' = \frac{(I_{ik}e^{j\delta_{ik}} + I_{ki}e^{j\delta_{ki}})(V_i e^{j\theta_i'} - V_k e^{j\theta_k})}{I_{ik}I_{ki}e^{(}\delta_{ik} + \delta_{ki})}. \tag{34}$$

Then, the deviation in the value of d_{ik} can be expressed as:

$$\Delta d_{ik} = d_{ik}' - d_{ik} = arg\left(\frac{V_i e^{j\theta_i'} - V_k e^{j\theta_k}}{V_i e^{j\theta_i} - V_k e^{j\theta_k}}\right). \tag{35}$$

Since the voltages are maintained close to 1 p.u. and the difference in voltage magnitudes measured at the two ends of a line is small (dependent on the voltage regulation), for simplicity of analysis, we assume them to be approximately equal. That is, $V_i \simeq V_k = V$. Also, let us take $\gamma = \theta_k - \theta_i$. Then (35) simplifies to,

$$\Delta d_{ik} = arg(e^{jx} - e^{j\gamma}) - arg(1 - e^{j\gamma}). \tag{36}$$

After algebraic manipulations, Δd_{ik} can be shown to be,

$$\Delta d_{ik} = \tan^{-1}\left(-\cot\left(\frac{x+\beta}{2}\right)\right) - \tan^{-1}\left(-\cot\left(\frac{\beta}{2}\right)\right). \tag{37}$$

A similar deviation of opposite sign is observed in the value of d_{ik} if the voltage angle at the other end of the line is changed.

Thus, when there is voltage angle modification, both the equivalent impedance magnitudes, angles, as well as the angle of the difference between the equivalent impedances change. The effect of modifying the voltage angles is shown in Figures 8 and 9 respectively for the ramp and step attacks.

B. Attack Detection and Classification

Based on the dependencies derived in Section V-A, a truth table can be developed for detecting the presence of data manipulation and determining which measurement has been

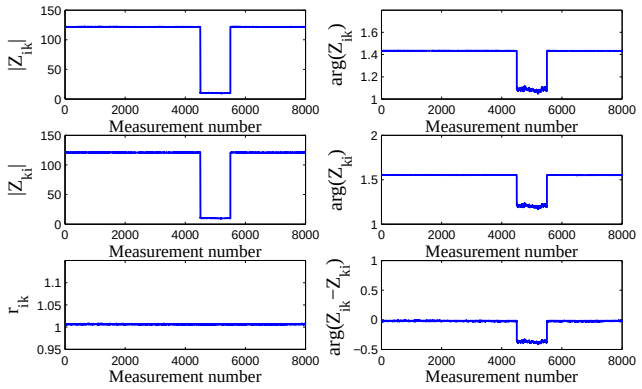

Fig. 9. Computed quantities when step attack is executed on voltage angle at bus k.

manipulated. It is also possible to point out whether the manipulation is positive or negative in nature.

If PMU measurements at two end buses (say i and k) of a line are available, for every set of measurement data, the following six quantities are computed and monitored to check for inconsistencies or manipulations:

1) z_{ik} or equivalent impedance magnitude from bus i
2) $arg(z_{ik})$ or equivalent impedance angle from bus i
3) z_{ki} or equivalent impedance magnitude from bus k
4) $arg(z_{ki})$ or equivalent impedance angle from bus k
5) r_{ik} or ratio of equivalent impedance magnitudes
6) d_{ik} or angle of difference between equivalent impedances.

From the deviations in Section V-A, the changes in the six quantities for various data manipulation may be summarized as follows:

1) If the current magnitude at bus i, I_{ik}, is increased by the attacker, it results in a corresponding decrease in z_{ik}, r_{ik} and d_{ik}. If the current magnitude at the other bus, I_{ki}, is increased, the value of z_{ki} decreases while r_{ik} and d_{ik} increase. The other quantities do not show any abnormal deviation. Also, if the current magnitudes are decreased, deviations of opposite sign are observed.
2) If the current angle measured at bus i, δ_{ik} is increased, it results in a corresponding decrease in the values of $arg(z_{ik})$ and d_{ik}. If the current angle at bus k or δ_{ki} is increased, the value of $arg(z_{ki})$ decreases while the value of d_{ik} increases. Deviations of opposite sign are observed when the current angles are decreased.
3) If the voltage angle measured at bus i, θ_i, is increased by the attacker, there is a corresponding increase in the values of z_{ik}, $arg(z_{ik})$, z_{ki}, $arg(z_{ki})$ and change (either increase or decrease) in the value of d_{ik}. If, however, the voltage angle at bus k or θ_{ki} is increased, the values of z_{ik}, $arg(z_{ik})$, z_{ki}, $arg(z_{ki})$ decrease while the value of d_{ik} changes (either increases or decreases). Deviations of opposite sign are observed when the voltage angles are decreased.

Thus, when any measurement data is maliciously increased or decreased, deviations of predictable sign are observed in a specific subset of the six monitored features depending on the type of data manipulated. The deviations in the monitored features for manipulation of different measured quantities is indicated in Table I. Therefore, based on the scheme, five different types of attacks can be accurately classified:

1) I_{ik} manipulation attack denoted by c_1
2) I_{ki} manipulation attack denoted by c_2
3) δ_{ik} manipulation attack denoted by c_3
4) δ_{ki} manipulation attack denoted by c_4
5) θ_{ik} or θ_{ki} manipulation attack denoted by c_5

If the states of z_{ik}, $arg(z_{ik})$, z_{ki}, $arg(z_{ki})$, r_{ik} and d_{ik} are represented by the variables $a_k \quad k = 1, \cdots, 6$ respectively, where 1 represents a change while 0 represents no change, then the truth table shown in II is obtained.

The proposed attack detection scheme is shown in Algorithm 1. Initially all a's and c's are at zero where the a's denote the state of the monitored features and the c's denote the five possible attack classifications. On receiving measurement data, the PDCs calculate the six quantities for each of the lines. Anomaly detection is performed on each of these monitored quantities. Let x_i be the current estimated value of any the six quantities for the i^{th} measurement and let μ_i be the mean and σ_i be the standard deviation as observed over a window of n preceding measurements, i.e.,

$$\mu_i = \frac{1}{n} \sum_{j=i-n}^{i-1} x_j \tag{38}$$

$$\sigma_i^2 = \frac{1}{n} \sum_{j=i-n}^{i-1} (x_j - \mu_{i-1})^2. \tag{39}$$

TABLE I
IMPACT OF DIFFERENT ATTACKS ON FEATURES OF CLASSIFICATION. ↑: INCREASE, ↓: DECREASE, ↕: INCREASE/DECREASE, $\gamma = arg(z)$

Data	Change	z_{ik}	γ_{ik}	z_{ki}	γ_{ki}	r_{ik}	d_{ik}
I_{ik}	↑	↓				↓	↓
	↓	↑				↑	↑
I_{ki}	↑			↓		↑	↑
	↓			↑		↓	↓
δ_{ik}	↑		↓				↓
	↓		↑				↑
δ_{ki}	↑				↓		↑
	↓				↑		↓
θ_i	↑	↑	↑	↑	↑		↕
	↓	↓	↓	↓	↓		↕
θ_k	↑	↓	↓	↓	↓		↕
	↓	↑	↑	↑	↑		↕

TABLE II
TRUTH TABLE FOR ATTACK TYPE CLASSIFICATION.

a_1	a_2	a_3	a_4	a_5	a_6	Attack Type
1				1	1	c_1
		1		1	1	c_2
	1				1	c_3
			1		1	c_4
1	1	1	1		1	c_5

If for any of the quantities,

$$|x_i - \mu_i| > |\sigma_i|, \tag{40}$$

then the difference of the LHS and RHS is stored in a variable (say Δx_i). The system monitors the cumulative sum of Δx_i for each these six variables. If any of the six cumulative sums crosses a particular threshold, that particular quantity is indicated to have deviated abnormally and the corresponding value of a is incremented to 1. The system then enters the alert mode and initializes a timer of t seconds. The abnormal deviations of the monitored quantities are compared with the possibilities as shown in Table II. If a match is found with any of the combinations, an alert is generated indicating the attack type.

Alarm clustering is performed in order to reduce false alarms. Before the expiry of the timer, if two alerts of the same type are generated then data manipulation attack on that measurement type is confirmed. However, if one or less alerts of a particular kind are generated then they are regarded as false positives and monitoring is continued in the usual manner. The Δx_i's and cumulative sums are reset after timer expiry.

Algorithm 1 Anomaly Detection and Alert Generation

loop
  for arrival of measurement number j **do**
    Update measurements from bus i: V_i, θ_i, I_{ik}, δ_{ik};
    Update measurements from bus k: V_k, θ_k, I_{ki}, δ_{ki};
    $\vec{V_i} = V_i(cos\theta_i + i\sin\theta_i)$;
    $\vec{I_{ik}} = I_{ik}(cos\delta_{ik} + i\sin\delta_{ik})$;
    Similarly calculate $\vec{V_k}$ and $\vec{I_{ki}}$;
    Calculate eq. imp. seen from bus i: $z_{ik}\vec{(j)} = (V_i\vec{(j)} - V_k\vec{(j)})/I_{ik}\vec{(j)}$;
    Calculate $z_{ki}\vec{(j)}$ similarly;
    Compute $x_1(j) = z_{ik}(j)$;
    Compute $x_2(j) = arg(z_{ik}(j))$;
    Compute $x_3(j) = z_{ki}(j)$;
    Compute $x_4(j) = arg(z_{ki}(j))$;
    Compute $x_5(j) = r_{ik}(j)$
    Compute $x_6(j) = d_{ik}(j)$
    Calculate $\mu_l(j)$, $\sigma_l(j)$, $l = 1, 2, ..6$;
    if $|x_l(j) - \mu_l(j)| > |\sigma_l(j)|$ **then**
      Calculate $\Delta x_l = |x_l(j) - \mu_l(j)| - |\sigma_l(j)|$;
      Calculate cumulative sum of deviations: $S_l = S_l + \Delta x_l$;
      if then $S_l > \eta_l$
        ALERT(l)
        Reset S_l;
      end if
    end if
  end for
end loop when session is terminated

Algorithm 2 Features Classification and Final Detection

function ALERT(l)
  if then $\sum_{l=1}^{6} a_l + \sum_{p=1}^{5} c_p = 0$
    Start timer for t sec;
    Update $a_l = 1$;
  else
    Update $a_l = a_l + 1$;
    if $(a_1 > 0) \wedge (a_5 > 0) \wedge (a_6 > 0)$ **then**
      Alert: $c_1 = c_1 + 1$;
      Update $a_l = 0$ for $l = 1, 5, 6$;
    else if $(a_3 > 0) \wedge (a_5 > 0) \wedge (a_6 > 0)$ **then**
      Alert: $c_2 = c_2 + 1$;
      Update $a_l = 0$ for $l = 3, 5, 6$;
    else if $(a_2) > 0 \wedge (a_6 > 0)$ **then**
      Alert: $c_3 = c_3 + 1$;
      Update $a_l = 0$ for $l = 2, 6$;
    else if $(a_4) > 0 \wedge (a_6 > 0)$ **then**
      Alert: $c_4 = c_4 + 1$;
      Update $a_l = 0$ for $l = 4, 6$;
    else if $(a_1 > 0) \wedge (a_2 > 0) \wedge (a_3 > 3) \wedge (a_4 > 0) \wedge (a_6 > 0)$ **then**
      Alert: $c_5 = c_5 + 1$;
      Update $a_l = 0$ for $l = 1, 2, 3, 4, 6$;
    end if
  end if

  if Timer has not Expired **then**
    if $c_1 \geq 2$ **then**
      Generate alarm "I_{ik} manipulation attack";
      Reset timer and c_1;
    else if $c_2 \geq 2$ **then**
      Generate alarm "I_{ki} manipulation attack";
      Reset timer and c_2;
    else if $c_3 \geq 2$ **then**
      Generate alarm "δ_{ik} manipulation attack";
      Reset timer and c_3;
    else if $c_4 \geq 2$ **then**
      Generate alarm "δ_{ki} manipulation attack";
      Reset timer and c_4;
    else if $c_5 \geq 2$ **then**
      Generate alarm "θ_{ik} or θ_{ki} manipulation attack";
      Reset timer and c_5;
    end if
  else
    Reset timer, a's and c's;
  end if
end function

VI. Simulation Results

In this section we present simulation results to verify the proposed detection mechanism. Real PMU data collected from the power grid in the state of New York, USA have been used to verify the mechanism. Data available from 4 PMUs on the grid that are located at buses on both sides of two specific lines have been used. The PMU reporting rate is 30 samples/second. Five sets of PMU data containing 9000 measurement samples each, and collected on different days were used for evaluating the proposed detection method. The operators have verified the data to be unmodified and we

have simulated the attacks by modifying the original data. Three sets of simulations have been performed: first with modification in current magnitude, second with modification in current angle, and third with modification in voltage angle. For each set, five levels of modifications have been simulated, that is, 0% or no change, 1%, 1.5%, 2% and 5% changes. In case of step attacks, a manipulation or modification level of x% means that the attacker-manipulated value differs from the true measurement value by x%. In case of ramp attacks, the manipulation is done gradually and a manipulation of x% means that the maximum deviation of the manipulated value from the true value achieved at any time during the attack is x%. Here, we gradually increase the manipulation to x% at the middle of the attack period and then again bring it back to normal.

In order to evaluate the effectiveness of the proposed data manipulation attack detection mechanism, we consider two types of attacks that may be executed on PMU data: ramp and step. We evaluate the performance of the proposed detection mechanism in terms of its accuracy (ACC), false positive (FP) rates, and false negative (FN) rates. The accuracy is the probability that detection outcome is correct. The false positive rate is the probability that a data manipulation attack alarm is raised when in reality there was no data manipulation. The false negative rate is the probability that a data manipulation attack goes undetected. The overall results of detecting data manipulation attack are provided in Table III.

TABLE III
OVERALL DETECTION RESULTS USING PROPOSED DATA MODIFICATION ATTACK DETECTION ALGORITHM.

Modification percentage	Ramp change			Step change		
	A	FP	FN	A	FP	FN
0	100	0	0	100	0	0
1	85	0	15	93	0	7
1.5	98	0	2	98	0	2
2	100	0	0	100	0	0
5	100	0	0	100	0	0

The accuracy of the detection scheme is above 85% with almost zero false alarms when there is no attack. The accuracy increases with increase in the modification level or if the modification is in the form of step-change. The maximum possible detection delay is 3 seconds, since it is the timer value used for the simulations. However, in most cases the threshold is exceeded much before the timer expiry, making the detection delay much smaller.

A. Comparison with Bad Data Detection

Conventional state estimators are equipped with mechanisms for bad data detection. In this section, we use arguments based on existing results to show that there proposed mechanism can successfully detect data manipulation attacks in scenarios where bad data detection techniques fail to do so.

Let $x = (x_1, x_2, \cdots, x_n)^T$ and $z = (z_1, z_2, \cdots, z_m)^T$ denote state variables and PMU measurements respectively, where n is the number of state variables and m is the number of PMU measurements and $m \geq n$. The bad data detector will flag a measurement if $||z - Hx|| > \tau$, where H is an $m \times n$ system matrix and τ is the detection threshold. Let the attack vector be denoted by $a = a_1, a_2, \cdots, a_m$ and thus the vector of observed measurements is $z_a = z + a$. It has been shown in [5] that if a is a linear combination of the column vectors of H (i.e. $a = Hc$), then the manipulated measurements can pass the bad data detection test.

For simplicity consider a system with three states and four measurements whose system matrix is given by

$$H = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Now consider the attack vector obtained using $c = [0.5, 0, 0]$. The observed measurements in this case will be given by

$$z_a = z + \begin{bmatrix} 0.5 \\ -0.5 \\ 0 \\ 0 \end{bmatrix}$$

which is equivalent to a step attack on the first two measurements. As demonstrated by our results in this section, the proposed detection methodology will be able to detect such attacks (with higher accuracy if the change of 0.5 is large compared to the original value of z) while from [5], it will be missed by traditional bad data detection.

B. Discussion

In the proposed detection mechanism, false positives may occur due to sensor problems or system disturbances. Also, false negatives may occur when the data manipulation level is low and in the range of the possible sensor variation. Note that the proposed detector is quicker and more accurate for attacks with quick changes in the manipulated values (e.g. step attacks). If the measurements are manipulated slowly (e.g. ramp attack), the gradually drifting mean and standard deviation values of the monitored parameters makes it more difficult to detect the attacks.

When an attacker changes any measurement of one PMU then the corresponding features will show a deviation. Now there are two constraints for the attacker to satisfy in order to evade detection. Firstly, the equivalent impedance phasor values for all the lines connected to this compromised node after the attack should be similar in value to the equivalent impedance phasor values computed before the attack was initiated. Only this will ensure that the first four features in our detection mechanism will not indicate an anomaly. Secondly, the equivalent impedance phasors for all the connected lines as seen from the bus with the compromised PMU should be similar in value to the equivalent impedances computed from the other ends of all these lines as well. This will ensure that the last two features of our detection mechanism will indicate no abnormal deviations. In order to satisfy this constraint, the attacker has to have access to all the PMUs at the other ends of all the lines connected to the compromised node and change their values too. Now to evade detection, the attacker

has to satisfy the two constraints for all those PMUs as well, and thus a chain of dependencies is created. This dependency on the measurements thus iteratively includes all the PMUs in the grid. Therefore, for changing the PMU measurements and evade detection, the attacker will need to change the measurements of all the PMUs. The attacker needs unlimited access to all the PMUs in the network for the attack to be successful, which is highly unlikely. If the attacker, however, is successful in changing the measurements of all the PMUs while satisfying the constraints mentioned above, it may not be possible to detect that attack.

VII. Conclusions

This paper proposed a mechanism for detecting data manipulation attacks on PMU data. The method is reliable even in the presence of instrumentation errors. It does not require any iterative computations and hence is comparatively faster. The effectiveness of the detection mechanism has been verified using simulations. Also, unlike many existing mechanisms for detecting bad data, it does not require the assumption that either some of the PMUs are absolutely secure or that some of the states are verifiable.

References

[1] H. Sandberg, A. Teixeira, and K. Johansson, "On Security Indices for State Estimators in Power Networks", *Proc. of Workshop on Secure Control Systems,* Stockholm, Sweden, 2010.

[2] A. Abur and A. G. Expsito, Power System State Estimation: Theory and Implementation, Marcel Dekker, 2004.

[3] M. Baran and A. Abur, "Power System State Estimation," *Wiley Encyclopedia of Electrical and Electronics Engineering,* 1999.

[4] E. Handschin, F. Schweppe, J. Kohlas and A. Fiechter, "Bad data analysis for power system state estimation," *IEEE Transactions on Power Apparatus and Systems,* vol. 94, no.2, pp. 329-337, March 1975.

[5] Y. Liu, P. Ning and M. Reiter, "False data injection attacks against state estimation in electric power grids", *Proc. of ACM CCS,* Chicago, IL, November 2009.

[6] R. B. Bobba, K. M. Rogers, Q. Wang, and H. Khurana, "Detecting false data injection attacks on DC state estimation", *Proc. of Workshop on Secure Control Systems,* Stockholm, Sweden, 2010.

[7] S. Bi and Y. J. Zhang, "Defending mechanisms against false-data injection attacks in the power system state estimation," in GLOBECOM Workshops (GC Wkshps), 2011 IEEE, Dec. 2011, pp. 1162 1167.

[8] O. Kosut J. Liyan, R. Thomas and T. Lang, "Malicious Data Attacks on Smart Grid State Estimation: Attack Strategies and Countermeasures," *Proc. of IEEE SmartGridComm,*, pp. 220-225, Gaithersberg, MD, Oct. 2010.

[9] T. Kim and H. V. Poor, "Strategic Protection Against Data Injection Attacks on Power Grids," *IEEE Transactions on Smart Grid,* vol. 2, no. 2, pp. 326-333, June 2011.

[10] C. Shuguang, H. Zhu, S. Kar, T. Kim, H. V. Poor and A. Tajer, "Coordinated data-injection attack and detection in the smart grid: A detailed look at enriching detection solutions,"*IEEE Signal Processing Magazine,* vol. 29, no. 5, pp. 106-115, Sept. 2012.

[11] A. Giani, R. Bent, M. Hinrichs, M. McQueen and K. Poolla, "Metrics for assessment of smart grid data integrity attacks,"*Proc. of IEEE Power and Energy Society General Meeting,*pp. 1-8, San Diego, CA, July 2012.

[12] A. Teixeira, S. Amin, H. Sandberg, K. Johansson, and S. Sastry, "Cyber security analysis of state estimators in electric power systems," in Decision and Control (CDC), 2010 49th IEEE Conference on, pp. 59915998., 2010.

[13] W. Niemira, R. B. Bobba, P. Sauer, W. H. Sanders, "Malicious data detection in state estimation leveraging system losses and estimation of perturbed parameters," *IEEE International Conference on Smart Grid Communications (SmartGridComm) 2013,* pp. 402-407, Vancouver, BC, Canada, 21-24 October 2013.

[14] S. Pal, B. Sikdar and J. Chow, "Detecting Malicious Manipulation of Synchrophasor Data," *Proc. of IEEE SmartGridComm,* Miami, FL, Nov. 2015.

[15] S. Mousavian, J. Valenzuela and J. Wang, "A Probabilistic Risk Mitigation Model for Cyber-Attacks to PMU Networks," in IEEE Transactions on Power Systems, vol. 30, no. 1, pp. 156-165, Jan. 2015.

[16] Y. Fan, Z. Zhang, M. Trinkle, A. D. Dimitrovski, J. B. Song and H. Li, "A Cross-Layer Defense Mechanism Against GPS Spoofing Attacks on PMUs in Smart Grids," in IEEE Transactions on Smart Grid, vol. 6, no. 6, pp. 2659-2668, Nov. 2015.

[17] K. Manandhar, X. Cao, F. Hu and Y. Liu, "Detection of Faults and Attacks Including False Data Injection Attack in Smart Grid Using Kalman Filter," inIEEE Transactions on Control of Network Systems, vol. 1, no. 4, pp. 370-379, Dec. 2014.

[18] D. Wang , X. Guan , T. Liu , Y. Gu , C. Shen and Z. Xu, "Extended distributed stateestimation: A detection method against tolerable false data injection attacksin smart grids",Energies, vol. 7, no. 3, pp. 1517-1538, 2014.

[19] L. Xie, Y. Mo and B. Sinopoli, "False Data Injection Attacks in Electricity Markets,"Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on, Gaithersburg, MD, 2010, pp. 226-231.

[20] O. Kosut, Liyan Jia, R. J. Thomas and Lang Tong, "Limiting false data attacks on power system state estimation,"Information Sciences and Systems (CISS), 2010 44th Annual Conference on, Princeton, NJ, 2010, pp. 1-6.

[21] O. Kosut, L. Jia, R. J. Thomas and L. Tong, "On malicious data attacks on power system state estimation,"Universities Power Engineering Conference (UPEC), 2010 45th International, Cardiff, Wales, 2010, pp. 1-6.

[22] J. Grainger and W. Stevenson, *Power System Analysis,* New York: McGraw-Hill, 1994.