False Data Injection Attacks Against Nonlinear State Estimation in Smart Power Grids

Md. Ashfaqur Rahman[†] and Hamed Mohsenian-Rad[‡]

[†]Department of Electrical and Computer Engineering, Texas Tech University, Lubbock, TX, USA [‡]Department of Electrical Engineering, University of California at Riverside, Riverside, CA, USA E-mails: md.rahman@ttu.edu and hamed@ee.ucr.edu.

Abstract—False data injection attacks are recently introduced as a class of cyber attacks against smart grid's monitoring systems. They aim to compromise the readings of grid sensors and phasor measurement units. Recent studies have shown that if the operator uses the DC, i.e., *linear*, state estimation to determine the current states of the power system, the attacker can adjust the attack vector such that the attack remains undetected and successfully passes the commonly used residue-based bad data detection tests. However, in this paper, we examine the possibility of implementing a false data injection attack when the operator uses the more practical AC, i.e., *nonlinear*, state estimation. We characterize such attacks when the attacker has *perfect* and *imperfect* knowledge of the current states of the system. To the best of our knowledge, this is the first paper to address false data injection attacks against non-linear state estimation.

Keywords: Smart grid security, false data injection attacks, non-linear state estimation, perfect and imperfect attacks.

I. INTRODUCTION

The recent advancements in the field of smart grid can potentially enhance efficiency and reliability in power systems. However, they may also create new vulnerabilities against cyber attacks. It fact, it has recently been shown that *false data injection attacks* (FDIAs) against *state estimation* can damage the grid and users equipment [6]. In an FDIA, an adversary hacks the readings of multiple sensors and phasor measurement units (PMUs) to mislead the grid operators. If the attack vector fulfils certain conditions, the adversary will be able to inject an arbitrary amount of error in state estimation and yet the FDIA will still pass the commonly used *residuebased bad data detection tests* [6], [7].

In [3], the authors showed that one can prevent FDIAs against state estimation by protecting a subset of sensors and PMUs. However, the number of sensors to be protected can be very large [2]. Another thread of research seeks to improve the existing residue-based bad data detection methods in state estimation such that they can also detect FDIAs. For example, some more advanced generalized likelihood ratio test and the adaptive cumulative sum control chart test to detect FDIAs are also proposed recently in [5] and [4], respectively.

Although FDIAs are widely studied over the past two years, most prior work, e.g., in [2]–[7], have focused only on a special class of FDIAs that target DC/linear state estimation. However, DC/linear state estimation is just a simplified version of a more general AC/nonlinear state estimation. There are several differences between linear and nonlinear state estimation. First, unlike in the linear case where the solution is obtained in closed-form, in non-linear state estimation the solution is obtained through iterations. Second, while linear state estimation is based on active power flow analysis, nonlinear state estimation is based on both active and reactive power flow analysis. Third, while the state variables in linear state estimation are only the voltage phase angles, nonlinear state estimation considers both voltage phase angles and magnitudes as states. These differences make nonlinear state estimation significantly more complicated. We believe that such complexity is the reason why FDIA against nonlinear state estimation has not been addressed before. However, since nonlinear state estimation is widely used in the power industry, understanding its vulnerability against FDIAs is of great practical importance.

In this paper, we develop a model for FDIAs against *nonlinear* state estimation. It requires the attacker to collect some *online* data from the grid while the attack is being implemented. This is in sharp contrast to the linear FDIA models, where the attacker only needs some *offline* data about the grid topology. Based on the accuracy of such online data gathering, we divide FDIAs against nonlinear state estimation into two classes: perfect and imperfect attacks. Simulation results show that our designed FDIAs can be successful in compromising the nonlinear power state estimation solutions.

Next, we briefly introduce FDIAs and nonlinear state estimation in section II. Perfect and imperfect nonlinear FDIAs are analyzed in section III. Simulation results are given in Section IV. The paper is concluded in section V.

II. SYSTEM MODEL AND BACKGROUND

A. Nonlinear State Estimation

Consider a power system such as the one in Fig. 1. Assume that S, with cardinality S, denote the set of buses. For the grid in Fig. 1, we have S = 30. Let z denote the vector of measurements which may include active and reactive power flows P_{km} and Q_{km} at each transmission line between any two buses $k, m \in S$ as well as active and reactive power injections P_k and Q_k at each bus k. These measurements are taken in such a way that the system becomes observable, i.e. it becomes possible to determine all state variables from measurements. In nonlinear state estimation, voltage magnitudes V_k and phase angles θ_k are usually taken as state variables at each bus $k \in S$. Since the phase angle for one of the buses is taken as reference angle, so the number of state variables becomes 2S - 1 which together form the state vector x. That is, we have

$$\mathbf{x} = \begin{bmatrix} \theta_2 & \theta_3 & \dots & \theta_S & V_1 & V_2 & \dots & V_S \end{bmatrix}^T.$$
(1)

The nonlinear power flow equations are the key to nonlinear state estimation. They indicate the relationships between the



Fig. 1. State estimation in an IEEE 30 bus system. The attacker collects local information to implement an FDIA attack against the phase angle at bus 25.

measurements and the state variables. For the transmission line between buses k and m, we have

$$P_{km} = V_k^2 g_{km} - V_k V_m g_{km} \cos(\theta_{km}) - V_k V_m b_{km} \sin(\theta_{km}),$$
(2)

$$Q_{km} = -V_k^2 b_{km} + V_k V_m b_{km} \cos(\theta_{km}) - V_k V_m g_{km} \sin(\theta_{km}).$$
(3)

Furthermore, for each bus k we have

$$P_{k} = V_{k} \sum_{m \in \mathcal{S}_{k}} V_{m}(-g_{km}\cos(\theta_{km}) - b_{km}\sin(\theta_{km})) + V_{k}^{2} \sum_{m \in \mathcal{S}_{k}} g_{km},$$

$$(4)$$

$$Q_{k} = V_{k} \sum_{m \in \mathcal{S}_{k}} V_{m}(-g_{km} \sin(\theta_{km}) + b_{km} \cos(\theta_{km})) - V_{k}^{2} \sum_{m \in \mathcal{S}_{k}} b_{km}.$$
(5)

Here, $S_k \subset S$ denotes the set of all buses that have lines to bus k. Furthermore, g_{km} and b_{km} are the conductance and susceptance of the line between buses k and m, respectively. Finally, $\theta_{km} = \theta_k - \theta_m$ denotes the voltage phase angle difference between buses k and m.

Given the power flow equations, the next step in nonlinear state estimation is to construct the Jacobian matrix \mathbf{J} , where the number of rows and columns are equal to the number of measurements and state variables, respectively. The entry in row *i* and column *j* of the Jacobian matrix denotes the derivative of the *i*th measurement with respect to the *j*th state variable based on their relationships in the power flow equations. For example, if the *i*th measurement is the amount of active power injection at bus *k* while the *j*th state variable is the voltage phase angle at bus $m \in S_k$, we have

$$\mathbf{J}_{ij} = \frac{\partial P_k}{\partial \theta_m} = V_k V_m (-g_{km} \sin(\theta_{km}) + b_{km} \cos(\theta_{km}))$$

There are different iterative methods for solving nonlinear state estimation problems such as Honest Gauss Newton method, Dishonest Gauss Newton method and Fast Decoupled State Estimator [7]. In Honest Gauss Newton method, the Jacobian matrix **J** is updated in each iteration, while in the Dishonest Gauss Newton method, it is assumed to be fixed. Let x[n] denote the most updated estimation of the state variables at iteration n. At flat start, all voltage magnitudes are set to 1 and all voltage phase angles are set to 0. That is,

$$\mathbf{x}[0] = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 \end{bmatrix}^T.$$
(6)

For a given x, let h(x) denote the measurement vector that is calculated based on the power flow equations. For example, if S = 10 and z_1 is the power flow from bus 1 to bus 2, then the 1st entry of vector h(x) becomes

$$\mathbf{h}_{1}(\mathbf{x}) = \mathbf{x}_{10}{}^{2}g_{1,2} - \mathbf{x}_{10}\mathbf{x}_{11}g_{1,2}\cos(-\mathbf{x}_{1}) - \mathbf{x}_{10}\mathbf{x}_{11}b_{1,2}\sin(-\mathbf{x}_{1}).$$
(7)

Recall from (1) that $\mathbf{x}_1 = \theta_2$, $\mathbf{x}_{10} = V_1$, and $\mathbf{x}_{11} = V_2$. After collecting all measurements \mathbf{z} and once matrix \mathbf{J} is constructed based on $\mathbf{x}[0]$, the following steps are repeated until the state vector converges to a fixed point:

Step 1: Set $\Delta \mathbf{x}[n] = (\mathbf{J}^T \mathbf{W} \mathbf{J})^{-1} \mathbf{J}^T \mathbf{W} (\mathbf{z} - \mathbf{h}(\mathbf{x}[n]))$. Step 2: Set $\mathbf{x}[n+1] = \mathbf{x}[n] + \Delta \mathbf{x}[n]$. Step 3: Update Jacobian \mathbf{J} based on $\mathbf{x}[n+1]$.

Note that Step 3 is performed only in the Honest Gauss Newton method [7]. In the Dishonest Gauss Newton method, matrix **J** is initiated based on the flat start state $\mathbf{x}[0]$ and it will not change through the iterations. Here, matrix **W** denotes the relative weight of the measurements based on the inverse of their noise variance. Measurements that have higher noise variance are given lower weight [7]. Finally, we note that there is no proof of convergence for the Gauss Newton algorithm above; however, experimental results have shown that the above algorithm almost always converges in practical nonlinear state estimation scenarios [1], [7].

B. Attacks Against DC State Estimation

The DC state estimation [6], [7] simplifies the nonlinear state estimation problem based on three assumptions. First, in a per unit system, all voltage magnitudes can be assumed to be fixed to 1. That is, $V_k = 1$, for all $k \in S$. Second, the resistance of the transmission lines is negligible. That is, $g_{k,m} = 0$, for all $k \in S$ and any $m \in S_k$. Three, the phase angle difference between any two neighboring buses is small. That is, $|\theta_{km}| \leq 5^\circ$, for all $k \in S$ and any $m \in S_k$. Applying these assumptions, the DC-equivalent/linear version of the power flow equations are obtained as

$$P_{km} = -b_{km}\theta_{km}, \qquad Q_{km} = 0. \tag{8}$$

The focus on DC power flow equations is only on active power injection at each bus. For each bus k, we have

$$P_k = \sum_{m \in \mathcal{S}_k} -b_{km} \theta_{km}.$$
(9)

Given the linear power flow equations in (8) and (9), the relationship between measurements and states becomes z = Hx, where the elements of matrix H are fixed and depend only on the grid topology and line admittances.

In FDIAs against linear state estimation [2]–[7], an adversary hacks the sensors such that the measurement vector \mathbf{z} is replaced by a compromised vector $\mathbf{z}_a = \mathbf{z} + \mathbf{a}$, where \mathbf{a} is false data vector. Given the false measurement vector \mathbf{z}_a , the state estimation solution becomes $\hat{\mathbf{x}}_a \neq \hat{\mathbf{x}}$. As shown in [6], FDIAs can sometimes be detected by using bad data detection methods based on evaluating the measurement residue:

$$\mathbf{z}_a - \mathbf{H} \hat{\mathbf{x}}_a, \tag{10}$$

and triggering an alarm if the residue is greater than a threshold. However, from [6], if the attacker selects vector **a** to be a linear combination of the rows in matrix **H**, i.e., $\mathbf{a} = \mathbf{H}\mathbf{c}$ for some arbitrary vector **c**, then residue-based bad data detection tests *cannot* detect the attack since the injected false data will no longer affect the residue:

$$\mathbf{r}_a = \mathbf{z}_a - \mathbf{H}\hat{\mathbf{x}}_a = \mathbf{z} + \mathbf{a} - \mathbf{H}(\hat{\mathbf{x}} + \mathbf{c}) = \mathbf{z} - \mathbf{H}\hat{\mathbf{x}} = \mathbf{r},$$
 (11)

where r denotes the residue in absence of an attack and

$$\hat{\mathbf{x}}_a = \hat{\mathbf{x}} + (\mathbf{H}^T \mathbf{W} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{W} \mathbf{H} \mathbf{c} = \hat{\mathbf{x}} + c.$$
 (12)

It is worth mentioning that matrix \mathbf{H} in DC state estimation is in fact part of the Jacobian matrix \mathbf{J} in nonlinear state estimation corresponding to the derivatives of real power flow and power injection equations with respect to the voltage phase angles, given the linear power flow equations in (8) and (9).

III. FALSE DATA INJECTION ATTACK FOR NONLINEAR STATE ESTIMATION

Let \mathbf{x} be the true state vector while \mathbf{x}_a is the false state vector that the attacker intends to inject into the nonlinear state estimation solution. The residue under attack becomes:

1 /

$$\mathbf{r}_{a} = \mathbf{z}_{a} - \mathbf{h}(\mathbf{x}_{a})$$

$$= \mathbf{z}_{a} - \mathbf{h}(\mathbf{x}_{a}) + \mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{x})$$

$$= \mathbf{z} + \mathbf{a} - \mathbf{h}(\mathbf{x}_{a}) + \mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{x})$$

$$= \mathbf{r} + \mathbf{a} - \mathbf{h}(\mathbf{x}_{a}) + \mathbf{h}(\mathbf{x}),$$
(13)

where the third equality is due to the definition of \mathbf{z}_a . From (13), in order to achieve $\mathbf{r}_a = \mathbf{r}$, such that the residue test does not reveal the attack, we must choose

$$\mathbf{a} = \mathbf{h}(\mathbf{x}_a) - \mathbf{h}(\mathbf{x}). \tag{14}$$

Thus, the attack vector in (14) can be a candidate for implementing an FDIA against nonlinear state estimation. Next, consider the case where no attack is attempted. Assuming that the three-step Dishonest Gauss Newton algorithm in Section II-A converges to the true state values, we have

$$\Delta \mathbf{x} = 0 \quad \Rightarrow \quad \mathbf{G}(\mathbf{z} - \mathbf{h}(\mathbf{x})) = \mathbf{0}, \tag{15}$$

where

$$\mathbf{G} = (\mathbf{J}^T \mathbf{W} \mathbf{J})^{-1} \mathbf{J}^T \mathbf{W}.$$
 (16)

Now assume that the attacker chooses the attack vector in (14). Also assume that the Dishonest Gauss Newton algorithm converges to a fixed point $\bar{\mathbf{x}}$. We have

$$\begin{aligned} \mathbf{\Delta x} &= 0 \; \Rightarrow \; \mathbf{G}(\mathbf{z}_a - \mathbf{h}(\bar{\mathbf{x}})) = \mathbf{0}, \\ &\Rightarrow \; \mathbf{G}(\mathbf{z} - \mathbf{h}(\mathbf{x})) + \mathbf{G}(\mathbf{h}(\mathbf{x}_a) - \mathbf{h}(\bar{\mathbf{x}})) = \mathbf{0}, \\ &\Rightarrow \; \mathbf{G}(\mathbf{h}(\mathbf{x}_a) - \mathbf{h}(\bar{\mathbf{x}})) = \mathbf{0}, \end{aligned}$$

where the last line is due to (15). From (17), either $\mathbf{h}(\mathbf{x}_a) - \mathbf{h}(\bar{\mathbf{x}})$ is in the null space of matrix **G** or we have $\mathbf{h}(\mathbf{x}_a) = \mathbf{h}(\bar{\mathbf{x}})$. Depending on the structure of vector function $\mathbf{h}(\cdot)$, the latter may indicate that $\mathbf{x}_a = \bar{\mathbf{x}}$. In fact, our simulation results in Section IV show that we indeed have $\mathbf{x}_a = \bar{\mathbf{x}}$ in most practical scenarios. From this, together with (13), we can conclude that selecting the attack vector (14) is *likely* to cause a successful attack which cannot be detected by residue tests. The *uncertainty* here is because of the Gauss Newton Method since not only its convergence is not guaranteed but also its exact fixed points cannot be analytically obtained either.

Next, we explain how an adversary may implement an FDIA based on (14). First, consider an example. For the power grid in Fig. 1, assume that an attacker aims to compromise state estimation for the voltage phase angle on bus 25. Let θ_{25} denote the true state value while c_{25} denotes the error that the attacker intends to inject into the state estimation solution. For the active power injection sensor on bus 25, the corresponding element in the attack vector **a** is obtained as

$$\begin{aligned} \mathbf{a}_{P_{25}} &= V_{25} \sum_{m \in S_{25}} V_m (-g_{25,m} \cos(\theta_{25} - \theta_m + c_{25})) \\ &\quad - b_{25,m} \sin(\theta_{25} - \theta_m + c_{25})) \\ &\quad - V_{25} \sum_{m \in S_{25}} V_m (-g_{25,m} \cos(\theta_{25} - \theta_m)) \\ &\quad - b_{25,m} \sin(\theta_{25} - \theta_m)) \\ &= 2V_{25} \sin(c_{25}/2) \sum_{m \in S_{25}} V_m (g_{25,m} \sin(\theta_{25} - \theta_m + c_{25}/2)) \\ &\quad - b_{25,m} \cos(\theta_{25} - \theta_m + c_{25}/2) \end{aligned}$$

where $S_{25} = \{24, 26, 27\}$. To implement the attack, the adversary needs to know the values of $V_{24}, V_{25}, V_{26}, V_{27}$ and $\theta_{24}, \theta_{25}, \theta_{26}, \theta_{27}$. This can be done either by directly measuring these quantities or estimating them using any available sensor in the region. Such sensors may include some of the operator's existing sensors that are hacked by the attacker



Fig. 2. Performance of a perfect nonlinear FDIA over iterations of nonlinear state estimation: (a) State estimation solution. (b) Residue norm.

and/or some new sensors that the attacker may temporarily deploy for the attack. Therefore, depending on the accuracy of estimating the state variables that the attacker needs for implementing the attack, false data injection attacks against nonlinear state estimation can be divided into two classes:

- *Perfect Attacks*: Attacks where the adversary can accurately obtain the needed state variables.
- *Imperfect Attacks*: Attacks where the adversary may obtain the needed state variables with error.

As we will see in Section IV, not only perfect but also some imperfect attacks can compromise nonlinear state estimation.

IV. SIMULATION RESULTS

A. Perfect Nonlinear FDIAs

Consider the power network in Fig. 1 and assume that the attacker aims to compromise the nonlinear state estimation solution for the phase angle on bus 25, i.e., θ_{25} . For this purpose, the attacker selects the attack vector according to (14). From Section III, this requires the attacker to estimate $V_{24}, V_{25}, V_{26}, V_{27}$ and $\theta_{24}, \theta_{25}, \theta_{26}, \theta_{27}$. Here, in this section, we assume that the attack is perfect and the attacker can accurately estimate all these states. The true value of the phase angle on bus 25 is assumed to be $\theta_{25} = 0.0202$ radians. The attacker aims to deviate the solution from this true value by $c_{25} = 0.03$ radians. The simulation results are shown in Fig. 2. Here, we assume that nonlinear state estimation is done using the Dishonest Gauss Newton method. We can see that state estimation quickly converges in less than 10 iterations. From Fig. 2(a) the attack has successfully deviated the nonlinear state estimation solution at the intended amount of c_{25} . From Fig. 2(b), the attack does not change in the residue norm.

B. Imperfect Nonlinear FDIAs

Next, we repeat the simulation in Section IV-A, but this time we assume that the attack is imperfect. In particular, we consider 100 different scenarios of slight (around 10%) inaccuracy in the attacker's obtained state estimation solutions



Fig. 3. Performance of an imperfect nonlinear FDIA for 100 different attack inaccuracy scenarios: (a) State estimation solution. (b) Residue norm.



Fig. 4. Changes in the residue corresponding to each measurement after implementing a *nonlinear* FDIA against *linear* state estimation.

for states V_{24} , V_{25} , V_{26} , V_{27} and θ_{24} , θ_{25} , θ_{26} , θ_{27} . From the results in Fig. 3(a), the attacker can no longer achieve a perfect attack as the deviations in the state estimation solutions are no longer equal to the intended amount of c_{25} . But the deviation is still close to the intended level. From the results in Fig. 3(b), imperfect attacks also make some changes in the residue. Interestingly, the residue norm has *decreased* due to the attacks. While we are not certain about the reason behind this observation, we conjecture that the *consistency* in the amounts of injected errors in the hacked sensors can be the cause to reduce the residue. However, regardless of the exact cause, the results in Fig. 3 show that even imperfect FDIAs can be successful in changing state estimation solutions while not being detected by residue-based bad data detection tests.

C. Nonlinear FDIAs on DC State Estimation

The proposed FDIA in this paper is designed against nonlinear state estimation. However, it works well even if the power grid operator uses linear state estimation. The changes



Fig. 5. Changes in the residue corresponding to each measurement after implementing a *linear* FDIA against *nonlinear* state estimation.



Fig. 6. Residue vs. measurement noise for perfect nonlinear FDIAs.

in the residue, i.e., the difference between the residues with and without attack, are shown in Fig. 4. We can see that the changes are minor (less than 6%), suggesting that the designed nonlinear FDIA can successfully pass the residue tests even if the operator implements a linear state estimation. The reason is that attacks against nonlinear state estimation are the generalizations of the attacks against linear state estimation.

D. DC FDIAs on Nonlinear State Estimation

In this section, we consider the opposite of the scenario in Section IV-C. That is, we assess the performance of the DC/linear FDIA in [6] when the grid operator uses nonlinear state estimation. As shown in Fig. 5, the changes in the residue due to the attack are very major. Therefore, if an attacker implements an FDIA intended for DC state estimation, but then the grid operator actually uses nonlinear state estimation, the attack will be detected. Note that, compared to Fig. 4, there are more residues in 5 because nonlinear state estimation uses measurements from both active and reactive power sensors.



Fig. 7. Residue vs. measurement noise for imperfect nonlinear FDIAs.

E. Impact of Measurement Noise

The impact of changes in the measurement noise on the residue norm for perfect and imperfect nonlinear FDIAs are shown in Figs. 6 and 7, respectively. We can see that perfect nonlinear FDIAs do not change the residue norm for all considered values of the measurement noise. The changes are still minor even for imperfect nonlinear FDIAs. On the other hand, the residue norm increases significantly when linear FDIAs are implemented. Furthermore, we can see in Fig. 7 that the residue norm drops when imperfect nonlinear FDIAs are implemented in presence of higher measurement noise.

V. CONCLUSIONS

This paper represents a first step towards understanding FDIAs against nonlinear state estimation. First, we developed a method to choose the attack vector, which requires the attacker to collect some online data from the grid while the attack is being implemented. This is in sharp contrast to linear FDIAs, where the attacker only needs some offline data about the grid topology. Then, based on the accuracy of such online data gathering, we classified nonlinear FDIAs into *perfect* and *imperfect* attacks. Simulation results showed that they are successful in changing the nonlinear state estimation solutions.

REFERENCES

- A. Abur and A. G. Exposito. Power System State Estimation : Theory and Implementation. CRC Press, New York, 2004.
- [2] S. Bi and Y. Zhang. Defending mechanisms against false-data injection attacks in the power system state estimation. In *Proc. of IEEE Globecom SG-COMNETS*, Houston, TX, Dec. 2011.
- [3] R. B. Bobba, K. M. Rogers, Q. Wang, H. Khurana, K. Nahrstedt, and T. J. Overbye. Detecting false data injection attacks on dc state estimation. In *Proc. of IEEE SCS*, Stockholm, Sweden, Apr. 2010.
- [4] Y. Huang, H. Li, K. A. Campbell, and Z. Han. Defending false data injection attack on smart grid network using adaptive cusum test. In *Proc. of IEEE CISS*, Baltimore, MD, Mar. 2011.
- [5] O. Kosut, L. Jia, R. J. Thomas, and L. Tong. On malicious data attacks on power system state estimation. In *Proc. of IEEE UPEC*, Aug. 2010.
- [6] Y. Liu, P. Ning, and M. K. Reiter. False data injection attacks against state estimation in electric power grids. In *Proc. of ACM CCS*, Chicago, IL, Oct. 2010.
- [7] A. Monticelli. State Estimation in Power Systems. Kluwer Academic Publishers, Boston, 1999.