Stealth False Data Injection using Independent Component Analysis in Smart Grid

Mohammad Esmalifalak‡, Huy Nguyen†, Rong Zheng†, and Zhu Han‡
†CS Department, University of Houston, Houston, TX
‡ECE Department, University of Houston, Houston, TX

***Abstract*—In smart grid, the strong coupling between cyber and physical operations makes power systems vulnerable to cyber attacks. In this paper, stealth false data attacks are investigated where the attackers** *without prior knowledge of the power grid topology*, **try to make inferences through phasor observations. We show that when the system dynamics are small and can be approximated linearly, linear independent component analysis (ICA) can be applied to estimate the Jacobian matrix multiplied by the eigenvectors of the covariance matrix of the state variables. The inferred structural information can then be used to launch unobservable attacks. As demonstrated by the simulation results using data generated by MATPOWER, the proposed scheme can indeed inject false data with low detectability.**

I. Introduction

State estimation is a key function in building real-time models of electricity networks in Energy Management Centers (EMC) [1]–[6]. A real-time model is a quasi-static mathematical representation of the current conditions in an interconnected power network [1]. This mathematical representation is usually obtained from measured and telemetered data every few seconds to the Energy Control Center (ECC). Real-time models of the network can be used by Independent System Operator (ISO) to make optimal decisions with respect to technical constraints such as transmission line congestion, voltage and transient stability [7]. In practice, it is not economical or even feasible to measure all possible states in the network; and thus state estimation is a useful tool for estimating those quantities from a limited set of measurements. Two kinds of information are usually used for state estimation in power systems [5]: i) Analog data of the system such as Megavar flows on all major lines, P and Q loading of generators and transformers, and voltage magnitudes at most of the buses of the system; ii) The on/off status of switching devices such as circuit breakers, disconnect switches, and transformer taps that determines the network topology[1].

Due to the importance of state estimation, the effects of injecting false measurement data have been studied in literature [8]–[12]. False data may be due to unintended measurement abnormalities or topology errors, or injection by malicious attacks. In [10], the authors demonstrate the feasibility of carrying out undetectable false data injection attacks with the objective of manipulating pricing of the electrical market. In [8], [9], the authors consider two categories of attacks, namely, observable and unobservable attacks and propose an algorithm to find the minimum number of compromised input sources

[1]In power system, On/Off switches change their state from Off to On (or viceversa) for removing the fault from the system so, in normal condition, these switches don't operate (network topology is static).

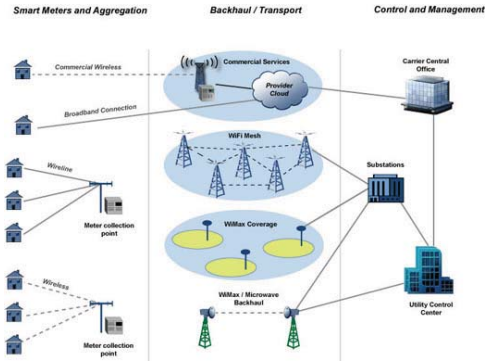


Fig. 1. Distributed Topology for Future Smart Grid

that are needed to carry out an unobservable attack. Also proposed in [8] are several counter-measures for observable attacks. [11] shows that an attacker can carry out stealth attacks by corrupting the power flow measurements through attacking the RTUs, tampering with the heterogeneous communication network or breaking into the SCADA system through the control center office LAN. Most of existing works assume attackers have knowledge regarding the network topology.

In this paper, we study a category of stealth false data injection attack with low detectability. The proposed stealth attack assumes no knowledge of the network topology and makes inferences from the correlations of the line measurements. Though distributed implementation is possible, we focus on centralized attack carried out by breaking into the SCADA system in the utility control center (Fig. 1). At the core of the proposed attacks, independent component analysis is implemented to infer the linear structure of the power flow measurements. Simulation studies using the data generated by MATPOWER [13] demonstrate its effectiveness. Our findings reveal the potential vulnerability of smart grid, and make a case for more advanced methods to detect and protect power systems from data injection and manipulation by intruders.

The remainder of this paper is organized as follows. The system model is given in Section II. False data injection is studied in Section III. The proposed scheme – ICA-based false data injection is presented and analyzed mathematically in Section III-C. The numerical results are provided in Section IV, and the conclusion is given in Section V.

II. System Model

Power systems generally have three main parts: 1) generation 2) transmission and 3) distribution. In power systems, transmission lines are used to transfer generated power to

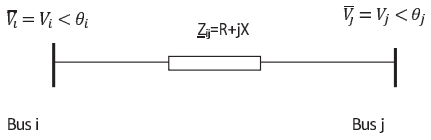


Fig. 2. Transmission Line Model

consumers [14]. Theoretically, transmitted complex power between bus i and bus j depends on the voltage difference between these two buses, and it is also a function of impedance between these buses, as shown in Figure 2.

In general, transmission lines have high reactance over resistance (i.e. X/R ratio), and one can approximate the impedance of a transmission line with its reactance. Transmitted active power form bus i to bus j can be written as [5]:

$$P_{ij} = \frac{V_i V_j}{X_{ij}} \sin(\theta_i - \theta_j), \tag{1}$$

where V_i is the voltage magnitude, θ_i is the voltage phase angle in bus i, and X_{ij} is the reactance of transmission line between bus i and bus j. In DC power flow studies it is usually assumed that voltage phase difference between two buses are small, and the amplitudes of voltages in buses are near to the unity. Therefore, further simplification gives a linear relation between voltage phase angles and lines reactance as [15]:

$$P_{ij} = \frac{\theta_i - \theta_j}{X_{ij}}. \tag{2}$$

The state-estimation problem is to estimate n phase angles θ_i, by observing m real-time measurements. In power flow studies, the voltage phase angle (θ_i) of the reference bus are fixed and known, and thus only $n-1$ angles need to be estimated. We define the state vector as

$$\mathbf{x} = [\theta_1, \ldots, \theta_n]^T. \tag{3}$$

The control center observes a vector $\mathbf{z}$ for m active power measurements. These measurements could be either transmitted active power from bus i to j (P_{ij}), or injected active power to bus i ($P_i = \sum_j P_{ij}$). The observation can be described as follows:

$$\mathbf{z} = \mathbf{P}(\mathbf{x}) + \mathbf{e}, \tag{4}$$

where $\mathbf{z} = [z_1, \cdots, z_m]^T$ is the vector of measured active power in transmission lines, $\mathbf{P}(\mathbf{x})$ is the nonlinear relation between measurement z and, state x that is the vector of n bus phase angles θ_i, and $\mathbf{e} = [e_1, \cdots, e_m]^T$ is the Gaussian measurement noise vector with covariant matrix $\mathbf{\Sigma}_e$.

Define the Jacobian matrix $\mathbf{H} \in \mathbb{R}^m$ as

$$\mathbf{H} = \frac{\partial \mathbf{P}(\mathbf{x})}{\partial \mathbf{x}} \mid_{\mathbf{x}=\mathbf{0}} . \tag{5}$$

If the phase differences ($\theta_i - \theta_j$) in (2) is small, then the linear approximation model of (4) can be described as:

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{e}. \tag{6}$$

Note that $\mathbf{H}$ is unknown to the attackers now but known to the independent system operator (ISO).

III. Stealth False Data Injection

We focus on the problems of identifying and mitigating the impact of malicious cyber attacks on state estimation, by recognizing the key role of state estimation as the interface between cyber and physical operations in a smart grid as shown in Figure 1. In this section, we first study how to detect the false data injection, and then we explain the concept of stealth attack.

A. Detection of false data injection

Given the power flow measurements $\mathbf{z}$, the estimated state vector $\hat{\mathbf{x}}$ can be computed as:

$$\hat{\mathbf{x}} = (\mathbf{H}^T \mathbf{\Sigma}_e^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{\Sigma}_e^{-1} \mathbf{z}. \tag{7}$$

Thus, the residue vector $\mathbf{r}$ can be computed as the difference between the measured quality and the calculated value from the estimated state:

$$\mathbf{r} = \mathbf{z} - \mathbf{H}\hat{\mathbf{x}}.$$

Therefore, the expected value and the covariance of the residual are:

$$E(\mathbf{r}) = 0 \text{ and } cov(\mathbf{r}) = (\mathbf{I} - \mathbf{M})\mathbf{\Sigma}_e, \tag{8}$$

where $\mathbf{M} = \mathbf{H}(\mathbf{H}^T \mathbf{\Sigma}_e^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{\Sigma}_e^{-1}$.

False data detection due to faulty sensors and topological errors can be performed using a threshold test [16]. Therefore, the hypothesis of not being attacked is accepted if

$$\max_i |r_i| \leq \gamma, \tag{9}$$

where γ is the threshold and r_i is the component of $\mathbf{r}$.

B. Stealth (unobservable) attack

From the discussion of the false data detection, we observe that if the attacker has knowledge on the topology $\mathbf{H}$, it can add $\delta\mathbf{x}$ to $\hat{\mathbf{x}}$. As a result,

$$\mathbf{z}' = \mathbf{H}(\mathbf{x} + \delta\mathbf{x}) + e, \tag{10}$$

then the hypothesis test would fail in detecting the attacker, since the control center believes that the true state is $\mathbf{x} + \delta\mathbf{x}$. This is called *stealth false data injection*.

Now the question is if the topology is not available to the attacker, can the attacker still successfully launch stealth false data injection. Our answer is, somewhat surprisingly, *yes*. The main idea is when system parameters (e.g., active or passive loads) vary in a small dynamic range, the structure (topology) information is in fact embedded in the correlations among power flow measurements. To get some intuition, let us consider an example. Let $\mathbf{z}(t), \mathbf{x}(t)$ be the measurements and state vectors at time t, where $\mathbf{x}(t)$ is unknown. For a particular t, it is impossible to infer $\mathbf{H}$ from $\mathbf{z}(t)$ alone. However, over time, if we have knowledge on some stochastic properties of the random process $\mathbf{x}(t)$, then we may be able to infer $\mathbf{H}$.

In power systems, the state variables are generally a (nonlinear) function of the loads $\mathbf{y}$ and the topology $\mathbf{H}$, namely, $\mathbf{x} = f(\mathbf{y}, \mathbf{H})$. While the topology is likely to be static over a period of time, loads can be modeled as varying *independently*.

If such variations are sufficiently small, we can approximate f with $\mathbf{x} = \mathbf{A}\mathbf{y}$, where $\mathbf{A}$ is the first-order coefficient matrix of the Taylor expansion at $\mathbf{y}$, i.e., $\mathbf{z} = \mathbf{H}\mathbf{A}\mathbf{y} + \mathbf{e}$.

With $\mathbf{H}\mathbf{A}$ and $\mathbf{y}$, we can carry out the false data injection attack by modifying the measurement data as $\mathbf{z}' = \mathbf{z} + \mathbf{H}\mathbf{A}\delta\mathbf{y}$, where $\delta\mathbf{y}$ can be arbitrarily chosen. At the ISO, from (7), we have $\hat{\mathbf{x}} = (\mathbf{H}^T\mathbf{\Sigma}_e^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{\Sigma}_e^{-1}\mathbf{z}'$. Let $\delta\mathbf{x} = A\delta\mathbf{y}$. Since $\mathbf{r} = \mathbf{z}' - H\hat{x} = \mathbf{z} + \mathbf{H}(\hat{\mathbf{x}} + \delta\mathbf{x})$, $E(\mathbf{r}) = 0$, $cov(\mathbf{r}) = (\mathbf{I} - \mathbf{M})\mathbf{\Sigma}_e$. In other word, the mean and variance of $\mathbf{r}$ is the same as the case without attackers. As a result, the attack cannot be detected.

C. ICA Algorithm

To infer $\mathbf{H}\mathbf{A}$ and $\mathbf{y}$, we adopt the linear independent component analysis (ICA) technique. Linear ICA [17] is a recently developed method in which the goal is to find a linear representation of the data so that components are as statistically independent as possible. It is a special case of blind source separation formulated as follows,

$$\mathbf{u} = \mathbf{G}\mathbf{v}, \tag{11}$$

with $\mathbf{u} = [u_i, i = 1, 2, \ldots, m]$ is the observable vector containing observation from m signal monitors, $\mathbf{G} = [g_{ij}, i = 1, 2, \ldots, m, j = 1, 2, \ldots, n]$ is the unknown mixing matrix, and $\mathbf{v} = [v_i, i = 1, 2, \ldots, m]$ is the source vector of n independent latent variables. Given the model and realizations of $\mathbf{u}$, ICA infers both the mixing matrix $\mathbf{G}$ and the source vector $\mathbf{v}$ by adaptively calculating the weight vector $\mathbf{w}$ and setting up a cost function that either maximizes the nongaussianity of the calculated $\mathbf{s} = (\mathbf{w}^T\mathbf{u})$ or minimizes the mutual information.

The following theorem establishes the identifiability of ICA up to scaling and permutation.

Theorem 1: (**Comon** [20]) Let no noise be present in $y = Mx$ and $y = Fz$, where x are independent random variables with at most one being Gaussian. The $F = M\Lambda P$, where Λ is an invertible diagonal matrix and P a permutation.

As discussed earlier, as long as the dynamics is small, we can linearize the mapping between the measurements and the state vector as $\mathbf{z} = \mathbf{H}\mathbf{x}$. We observe through simulated data that the state vector is typically highly correlated, namely, varying the load on one bus in the system may change the value of states on all buses. Thus, conceptually, we need to further project the state vector $\mathbf{x}$ to a space where the resulting vector $\mathbf{y}$ is independent. If the state vector $\mathbf{x}$ follows Gaussian distributions, then the principle component analysis (PCA) [21] can be used. In this case, let $\mathbf{x} = \mathbf{A}\mathbf{y}$, where $\mathbf{y}$ are independent random vectors, and $\mathbf{A}$ is the eigenvectors of $\mathbf{x}$. Thus, we have $\mathbf{z} = \mathbf{H}\mathbf{A}\mathbf{y} = \mathbf{G}\mathbf{y}$. For general distributions of $\mathbf{x}$, we can apply the ICA.

FASTICA [17] is an efficient and popular algorithm for ICA that iteratively finds the direction the for weight vector $\mathbf{w}$ to maximize the nongaussianity of the projection $\mathbf{w}^T\mathbf{z}$ for the data $\mathbf{z}$. Entries in G satisfying $\mathbf{w}^T\mathbf{G} = \mathbf{I}$ that are too small (compared to the threshold ε) will be removed, where $\mathbf{I}$ is the identity matrix. Finally, the quasi state vector $\mathbf{y}$ can be estimated by $\mathbf{w}^T\mathbf{z}$.

The algorithm is summarized in Algorithm 1. Line 2 verifies if $\mathbf{z}$ follows a linear model. If the linearity assumption holds then $\max(\mathbf{z} - \mathbf{G}\mathbf{y})$ should be small.

Algorithm 1: Stealth false data injection

input : $\mathbf{z}$ = data matrix;

1 [$\mathbf{G}$ and $\mathbf{y}$] = `FastICA` ($\mathbf{z}$);
2 **if** $\max(\mathbf{z} - \mathbf{G}\mathbf{y}) > \epsilon$ **then**
  exit;
3 Generate $\delta\mathbf{y} \sim N(0, \sigma^2)$;
4 $\mathbf{z}' = \mathbf{z} + \mathbf{G}(\mathbf{y} + \delta\mathbf{y})$;

output: false data $\mathbf{z}'$

IV. NUMERICAL RESULTS

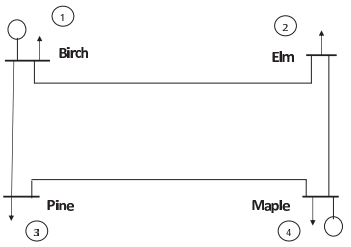
In this section, we evaluate the performance of the proposed methodology through extensive simulations using different network topologies. The state vector $\mathbf{x}$ represents the phase bus angle differences generated by MATPOWER [13]. MATPOWER is a Matlab simulation tool for solving power flow and optimal power flow problems. Using the data generated by MATPOWER reflects a more realistic simulated environment. The algorithm is implemented on Matlab. The presented results are experiment results conducted on 4-Bus test system [5], IEEE 14-Bus and 30-bus [22] smart grid models (as shown in Figure 3) with different number of measurements. In the simulation, we randomly vary the load on Load Buses (also known as PQ buses), which gives rise to (correlated) variations of the state vector. We also evaluate the impact of measurement noise on the detection probability.

A. Validation of linearity in ICA

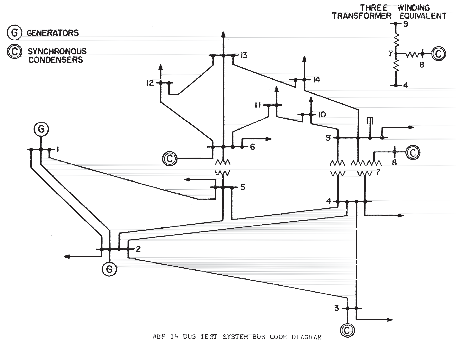
In this section, we evaluate the validity of linearity assumptions in ICA and its performance with different levels of noises and the number of measurements.

Figure 4 shows the mean square error of $\Delta\mathbf{z} = \mathbf{z} - \mathbf{G}\hat{\mathbf{y}}$, where $\mathbf{G}$ and $\hat{\mathbf{y}}$ are the estimation by ICA under different topologies. In the experiments, we vary the level of measurement noises ($\mathbf{e}$) indicated by the Signal Noise Ratio (SNR) of the true signal and the measurement noise. As shown in Figure 4, with the increase of SNR, the mean square error (MSE) decreases linearly in log-log scale. Furthermore, for different types of buses, the MSEs coincide even though the 14-bus has a more complex structure. When SNR is high (~ 40dB), the MSE is as low as 10^{-4}. This implies that the power flow can indeed be characterized using a linear model, and the ICA can successfully identify the underlying structure for different topologies as long as the noise is not too significant.

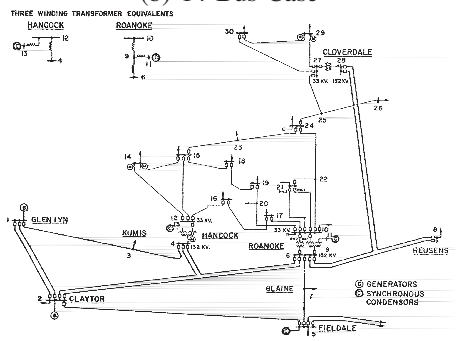
Next, we evaluate the speed of convergence for ICA. Figure 5 gives the MSEs with different number of observations under different SNR (5dB and 30dB) in the 14-bus topology. Similar results have been observed for other topologies and are thus omitted. As shown in Figure 5, as the number of measurements per bus increases from 10 to 30, the MSEs decreased drastically. However, when the number of measurements per bus increases beyond 30, there is little change in MSEs for



(a) 4 Bus Case



(b) 14 Bus Case



(c) 30 Bus Case

Fig. 3. System Model for Power Line

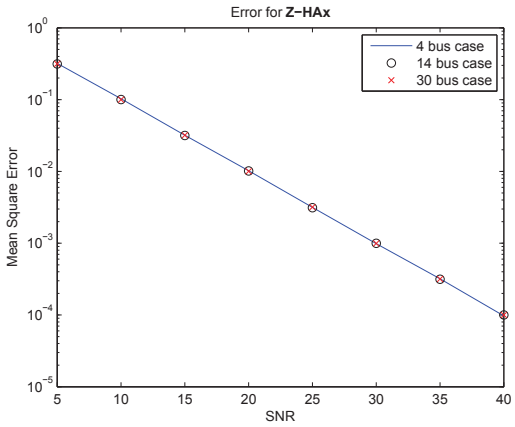


Fig. 4. MSE of ICA inference ($\mathbf{z} - \mathbf{G}\mathbf{y}$) vs. SNR.

both SNR levels. Therefore, we can conclude that the ICA algorithm can achieve high accuracy with a small number of observations. This implies that the attacks can be launched in almost real time.

In the third experiment, we study the independence of the state vector $\mathbf{x}$. We compute the eigenvalues of the covariance

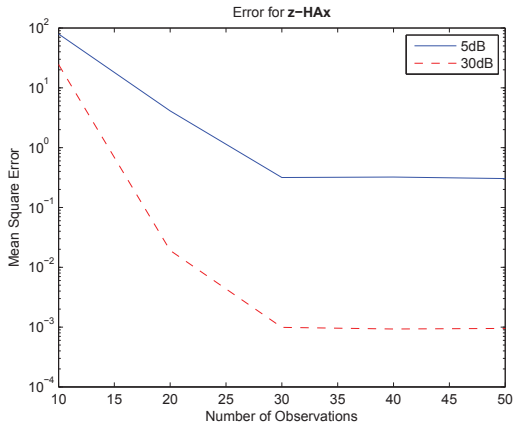


Fig. 5. MSE of ICA inference ($\mathbf{z} - \mathbf{G}\mathbf{y}$) vs. the number of Observations (14-bus case).

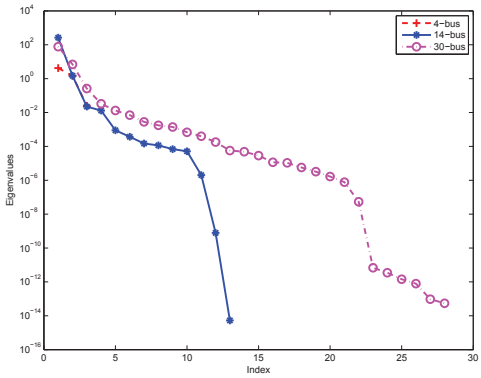


Fig. 6. Eigenvalues of the state vector of different bus topologies

sort them in descending order. As show in Figure 6, the state vector is clearly highly correlated. In fact, for the 14-bus and 30-bus, there are only 8 and 12 main components (with eigenvalues greater than 10^{-4}). Since ICA gives independent components, the resulting $\mathbf{y}$ are naturally independent. A key take-away from this set of experiments is that more sophisticated detection mechanisms can be devised if the correlation structure of the state vector can be utilized (namely, the 2nd order statistics). In this case, even when an attacker knows $\mathbf{H}$, if it naively injects random data to the measurement as $\mathbf{H}(\mathbf{x} + \delta\mathbf{x})$, then as long as $\mathbf{x}$ does not exhibit the same correlation structure $\mathbf{x}$, sophisticated detection mechanisms may still be able to detect the bad data injection. In contrast, since we decouple the dependency among $\mathbf{x}$'s by projecting them to a low-dimension space of independent vector, the proposed attacks are harder to detect.

B. Performance of attacks

In the previous section, we demonstrate that the ICA algorithm can successfully identify the linear structure of the power flow measurements. Next, the strength of the ICA-based attack is evaluated. As a baseline, we consider a naive attack that randomly injects false data (following a Gaussian distribution with zero mean and the same variance, 10dB higher than the noise level, as the stealth attack) without knowledge on $\mathbf{H}$.

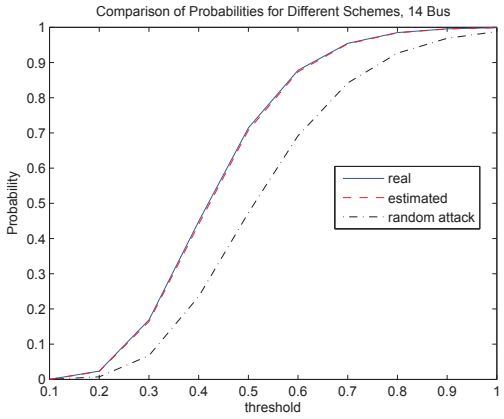
We further compare the proposed attack to the case without any false data injection.

The null hypothesis (no attack) is accepted with the probability that (9) holds. The probability is an increasing function of the threshold. To compute the probability, we assume the residual error $\mathbf{r}$ follows Gaussian distribution with mean and variance in (8), respectively.

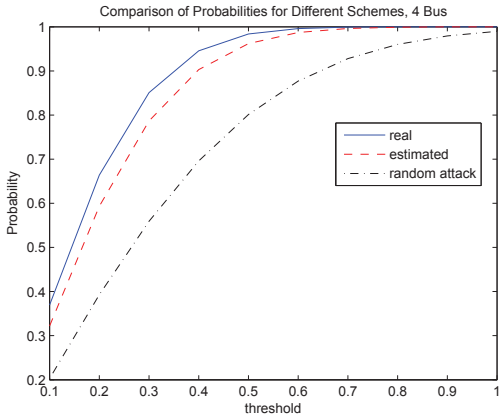
From Fig. 7(a), we can see the proposed stealth attack has an almost identical probability as the no-attack case in the 14-bus topology. So it is basically indistinguishable for the proposed attack. On the other hand, the random attack has very different characteristics. The gap between no-attack and random attack cases is even more pronounced in the 4-bus topology. However, as shown in Fig. 7(b), there is a small gap between the no-attack and proposed attack in the probabilities. This may be due to the higher dynamic loads resulting nonlinearity in the power flow measurements. We have also carried out simulation for the 30-bus topology and observe similar results as those in the 14-bus topology.

V. Conclusion

In this paper, we proposed an inference algorithm for smart grid system based on linear independent component analysis. Extensive simulation process demonstrates the methodology's potential to exploit the system vulnerability. We showed that an attacker can estimate both the system topology and power states just by observing the power flow measurements. Once the information is at hand, malicious attacks can be launched without triggering the detection system. Independent component analysis algorithm is proposed to obtain the information. Under the simulated data from MATPOWER, we compare the proposed algorithm with the random attack scheme and showed that our attack methodology is undetectable.



(a) 14 Bus Case



(b) 4 Bus Case

Fig. 7. Probability for miss detection of attacks

References

[1] A. Monticelli, "Electric Power System State Estimation," *Proceedings of the IEEE* , vol. 88, pp. 262–282, Feb. 2000.

[2] A. Bose and K. A. Clements, "Real-time modeling of power networks," *Proc. IEEE*, vol. 75, pp. 1607–1622, Dec. 1987.

[3] F. F. Wu, "Power system state estimation: A survey," *Int. J. Elect. Power Eng. Syst.*, vol. 12, pp. 80–87, Jan. 1990.

[4] L. Holten, A. Gjelsvik, S. Aam, F. F. Wu, and W. H. E. Liu, "Comparison of different methods for state estimation," *IEEE Trans. Power Syst.*, vol. 3, pp. 1798–1806, Nov. 1988.

[5] J. J. Grainger and W. D. Stevenson Jr, *Power system analysis*, vol. 621, 1994, McGraw-Hill

[6] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*, Marcel Dekker, Inc., 2004.

[7] Z. Alaywan and J. Allen, "California electric restructuring; A broad description of the development of the California ISO," *IEEE Trans. Power Syst.*, vol. 13, pp. 1445–1452, Nov. 1998.

[8] O. Kosut, L. Jia, R.J. Thomas, and L. Tong, "Malicious Data Attacks on Smart Grid State Estimation: Attack Strategies and Countermeasures," *First IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 220 – 225 , Nov. 2010.

[9] O. Kosut, L. Jia, R.J. Thomas, and L. Tong, "Limiting false data attacks on power system state estimation" *44th Annual Conference on Information Sciences and Systems (CISS)*, pp. 1 – 6 , May. 2010.

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

[11] Y. Liu, M. K. Reiter, and P. Ning, "False data injection attacks against state estimation in electric power grids," *the 16th ACM conference on Computer and communications security.*, Nov. 2009.

[12] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Conference on Computer and Communications Security*, Chicago, IL, Nov. 2009.

[13] R. D. Zimmerman, C. E. Murillo-Snchez, and R. J. Thomas, *MATPOWER Steady-State Operations, Planning and Analysis Tools for Power Systems Research and Education*, Power Systems, IEEE Transactions on, vol. 26, no. 1, pp. 12–19, Feb. 2011.

[14] J. Casazza and F. Delea, *Understanding Electric Power Systems*, IEEE Press Understanding Science and Technology Series, A John Wiley and Sons, Inc., 2010.

[15] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*, Wiley New York et al., 1996.

[16] F. F. Wu and W. E. Liu, "Detection of topology errors by state estimation", *IEEE Transactions on Power Systems*, vol.4, no.1, p.p.176–183, February 1989.

[17] J. Himberg and A. Hyvarinen, "Independent component analysis for binary data: An experimental study," in Proceedings of *the 3rd International Conference on Independent Component Analysis and Blind Signal Separation (ICA)*, pp. 552-556, 2001.

[18] H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval Research Logistic Quarterly*, vol. 2, pp. 83-97, 1955.

[19] H. Nguyen and R. Zheng, *Binary independent component analysis with or mixtures*, preprint (2011), available at *http://arxiv.org/abs/1007.0528*.

[20] Pierre Comon, "Independent component analysis, a new concept?", Signal Process. 36, 3 (April 1994), 287-314.

[21] Jolliffe I.T. *Principal Component Analysis, Series: Springer Series in Statistics*, 2nd ed., Springer, NY, 2002, XXIX, 487 p. 28 illus. ISBN 978-0-387-95442-4.

[22] University of Washington, "Power Systems Test Case Archives ," *http://www.ee.washington.edu/research/pstca/* ,[Apr. 2011].