Cyber Attacks on AC State Estimation: Unobservability and Physical Consequences

Jingwen Liang, Oliver Kosut and Lalitha Sankar

***Abstract*—An algorithm is developed to construct unobservable attacks for an AC state estimator (SE). It is shown that unobservability of the attack, in the absence of noise, is guaranteed when the attacker exploits its local network knowledge to perform AC SE locally than the simpler DC SE often assumed in the literature. Finally, the consequences of such an unobservable attack are highlighted via a scenario in which the physical system is changed due to false data injection.**

***Index Terms*—State estimation, false data injection, consequence.**

I. INTRODUCTION

Advances in sensing, communications, and computing are enabling a smart grid with an intelligent cyber layer that is tightly integrated with the physical layer and is capable of real-time monitoring control, and actuation. This, however, also makes the system more vulnerable to cyber-attacks. Such cyber-attacks can potentially affect the physical system by intelligently faking the system operating conditions without detection.

Cyber-attacks on state estimation from the perspectives of both the system and the attacker have been the subject of recent papers. In [1], Liu *et al.* are the first to study unobservable attacks on a DC state estimator and show that the conventional bad data detector is incapable of detecting such attacks. From a systems perspective, in [2], [3], [4], the authors quantify the vulnerability of an unobservable attack on a DC state estimator and illustrate how the the result can be used either to launch an attack or to devise a protection strategy. Focusing on unobservable attack, in [5], Kosut *et al.* define a security index for attacks on DC state estimation; furthermore, they also present a new detection algorithm for observable attacks. Distributed DC state estimation with attacks is considered in [6].

In contrast to such DC state estimation based attack models, [7] considers attacks on AC state estimation; the authors introduce an algorithm to identify a subgraph that suffices to construct an unobservable attack. They show that an attacker can execute unobservable attack on an AC state estimator if it has knowledge of both system topology and the system states (complex voltages) of its subgraph; this is contrast to the DC attack models for which it suffices for the attacker to have topology knowledge.

Using the subgraph algorithm of [7], this paper takes a step further and develops both DC and AC attack models for an attacker restricted to such a subgraph, having access only to local measurements and no direct state information. Specifically, we show that an attacker capable of using an AC state estimator locally within its subgraph can always launch an unobservable attack. We also introduce an operational definition of unobservability for attacks on general non-linear measurement models, including AC state estimation. It has not yet been demonstrated how attacks against either AC or DC state estimation. In this paper, we give examples in which an unobservable false data attack on AC state estimation can impact the physical system via the control center re-dispatching generation by using ACOPF from the (corrupted) state estimate.

II. MODEL FOR STATE ESTIMATION AND ATTACK

A. State estimation

State estimation is used to acquire the system operating conditions through noisy measurements, such as line power flows, bus voltage and line current magnitudes. All raw measurements are first passed through an observability analysis phase. If there are enough measurements to be used to do state estimation, the system will be observable; otherwise, the system will be unobservable and an unobservable island will be identified. State will then be estimated and passed to bad data detection, typically a χ^2 test. Measurements with too large error will be eliminated.

Since full power flow, i.e. AC power flow, follows non-linear mathematical dependencies, the AC measurement model is given by:

$$z = h(x) + e \tag{1}$$

where z, e and x are $m \times 1$, $m \times 1$ and $n \times 1$ vectors with entries z_i, e_i and x_k, respectively $i \in \{1, ..., m\}$ and $k \in \{1, ..., n\}$. The function $h(\cdot)$ denotes a non-linear relationship between the states and measurements , and e_i is assumed to be independent and Gaussian distributed with 0 mean and σ_i^2 covariance.

In AC state estimation, the state variables are solved as a least square problem with objective function [8]:

$$\min J(x) = (h(x) - z)^T R^{-1} (h(x) - z) \tag{2}$$

The solution to (2) satisfies:

$$g(\hat{x}) = \frac{\partial J(\hat{x})}{\partial x} = H^T(\hat{x}) \cdot R^{-1} \cdot (h(\hat{x}) - z) = 0 \tag{3}$$

where $H = \frac{\partial h(x)}{\partial x} \mid_{x=\hat{x}}$. This non-linear equation can be solved by iterative method.

978-1-4799-6415-4/14/$31.00 ©2014 IEEE

B. Bad data detection

The bad data detector filters noisy measurement and guarantees the accuracy of estimation. One of these methods of detection is the χ^2 test. To pass a χ^2 test, the estimated state should satisfy:

$$\sum_{i=1}^{m} \frac{(z_i - h_i(\hat{x}))^2}{\sigma_i^2} \leq \chi^2_{(m-n),p} \tag{4}$$

where $\hat{x}$ is the estimated state, $h_i(\hat{x})$ is the estimated measurements, p is the detection confidence of probability and $\chi^2_{(m-n),p}$ denotes the value in χ^2 distribution table corresponding to p and the degree of freedom $m - n$.

C. Attack model

We first assume that the attacker has following capabilities:

1) Attacker has access to all measurements and topology information of a small area S bounded by buses. The set of all measurement indices in S is denoted as I_S and the set of all state indices in S is denoted as K_S;
2) Attacker can change or replace all measurements in S, possibly by compromising communication channels between the RTUs in this area and control center;
3) Attacker has computational capability.

According to (1), suppose the i-th measurement prior to attack is $z_i = h_i(x) + e_i$, the attack model changes the i-th measurement z_i to $z_i^{(a)}$ such that:

$$z_i^{(a)} = \begin{cases} z_i & \text{if } i \notin I_S \\ \tilde{z}_i & \text{if } i \in I_S \end{cases} \tag{5}$$

where $\tilde{z}_i$ is chosen by attacker.

D. Unobservable attack model

Definition 1. We say an attack is *unobservable* for a measurement model $h(\cdot)$ if, in the absence of measurement noise, there exists a $c \neq 0$ such that $z_i^{(a)} = h_i(x + c)$ for all i.

Therefore, again assuming no measurement noise, (5) becomes:

$$z_i^{(a)} = \begin{cases} z_i & \text{if } i \notin I_S \\ h_i(x + c) & \text{if } i \in I_S \end{cases} \tag{6}$$

When the model $h(\cdot)$ is linear, i.e. $h(x) = Hx$, (6) becomes $z^{(a)} = Hx + a$, where $a = Hc$ and a_i is non-zero only if $i \in I_S$. Note that this condition is identical to the attack originally identified in [1].

E. Topology analysis

From (6), if the k-th state x_k is required to compute $h_i(x)$ for any $i \notin I_S$, then for any unobservable attack the corresponding k-th entry in attack vector must satisfy $c_k = 0$. Therefore, for a feasible attack, the attack region S must be chosen such that c is a non-zero vector. To identify such a collection of one or more buses in S, as in [7], we distinguish between two types of buses based on the presence of power injection. We henceforth identify buses with injection as *injection buses* and buses without injection as *non-injection buses*. An attacker can attack either type of bus. However, attacking a non-injection bus implies that the measurements at the closest injection buses also need to be changed due to the power balance requirement. For our analysis, we focus on an injection bus attack. In [7], a method is introduced to identify a subgraph of the network that allows an attacker to perform an unobservable attack. We use the same method, as summarized below.

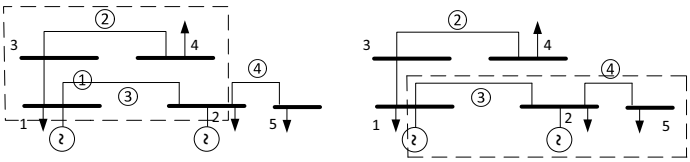

Figure 1. Examples of attack subgraphs. The figure on the left shows the subgraph with target bus 1, and on the right the subgraph with target bus 2.

Let k be a target injection bus. The corresponding *attack subgraph* S_k is constructed as follows:

1) Include bus k in S_k.
2) Extend S_k from bus k by including all buses that are connected to bus k;
3) If there is a non-injection bus on the boundary of S_k, extend S_k to include all adjacent buses of such a boundary bus;
4) Repeat (3) until all buses on the boundary are injection buses.

Fig. 1 illustrates a simple example system with 5 buses and 4 branches (circled numbers). The attack subgraph S_1, with target bus 1, is given by buses 1-4 and branches 1-3. Similarly, buses 1, 2, and 5, along with branches 3 and 4, form S_2.

III. Attack Strategy

A. Choice of attack vector c and attack subgraph S

As discussed in II-C and II-E, the choice of attack area S is equivalent to the choice of non-zero entries in c. For a target injection bus k, c_k can be chosen to be any desired value. From power balance requirements, the change for the non-injection buses in S_k are directly obtained.

Thus, the protocol for attack subgraph S is:

$$S = \bigcup_{\text{injection bus } k \,:\, c_k \neq 0} S_k \tag{7}$$

This choice of attack subgraph results in estimated load changes within S while no net load changes in the system.

B. DC attack model

Since (6) is nonlinear and generally hard to solve, it is reasonable for attacker first consider a simplified DC model to launch an attack. As [1] demonstrated, by knowing system Jacobian matrix H, an attacker can intelligently construct an unobservable attack vector $a = Hc$ such that $\tilde{z}_i = z_i + a$.

Thus, (6) becomes:

$$z_i^{(a)} = \begin{cases} z_i & \text{if } i \notin I_{SP} \\ z_i + H_{(i,:)}c & \text{if } i \in I_{SP} \end{cases} \quad (8)$$

where, I_P denotes the sets of indices of active power measurements, $I_{SP} = I_S \cap I_P$ and $H_{(i,:)}$ denotes the i-th row of H.

Though DC attack model is easy to construct, it is not an unobservable attack to AC state estimator. Without taking reactive power flow into account, a DC attack will be detected when c is too large. Indeed, we determine numerically in Section. IV that DC attacks modifying a bus angle by 0.1 degrees are typically detectable by AC state estimation. Thus, an ambitious attacker may want to use AC model to hide the attack completely.

C. AC attack model

From (6), in contrary to DC attack, we have that the attacker must know all x that appears in h_i, for all $i \in I_S$, to construct z_i precisely. However, this information is not available to the attacker. Thus, attacker uses the following model instead:

$$z_i^{(a)} = \begin{cases} z_i & \text{if } i \notin I_S \\ h_i(\hat{x}_k + c) & \text{if } i \in I_S \end{cases} \quad (9)$$

where $k \in K_S$ and $\hat{x}_k$ is a state estimate found by the attacker using local AC state estimation from its available measurements in S. In particular, the attacker uses the following procedure to find $z_i^{(a)}$:

1) The attacker uses protocol in III-B to choose S for desired c.
2) Perform state estimation in S using measurements available and get $\hat{x}_k$. The slack bus may be chosen arbitrarily among all injection buses.
3) For all injection buses j, set $\tilde{x}_j = \hat{x}_j + c_j$. The power balance requirements at non-injection buses suffice to compute $\tilde{x}_j$, for all non-injection bus $j \in S$.
4) The attacker then create false measurements as $z_i^{(a)} = h_i(\tilde{x})$.

Remark 2. The attack above is unobservable for AC measurement model because in the absence of noise, local state estimation perfectly recovers the states, albeit with constant phase shift due to a different slack bus. Thus, $z_i^{(a)} = h_i(\hat{x}_k + c)$ for all $i \in I_S$.

IV. Simulation Results

We use the IEEE RTS 24 bus system as test system. We assume that both active and reactive power flows are measured at two ends of each line and both active and reactive injection are measured at each injection bus, which makes 192 measurements in total. All measurements are assumed to have an error $e_i \sim N(0, 10^{-4})$ and the χ^2 detector threshold is set to be 174.1 with 95% confidence of detection. In our simulation, we use MATPOWER to generate measurements and perform state estimation. The default setting of IEEE RTS 24 bus system in MATPOWER is assumed.

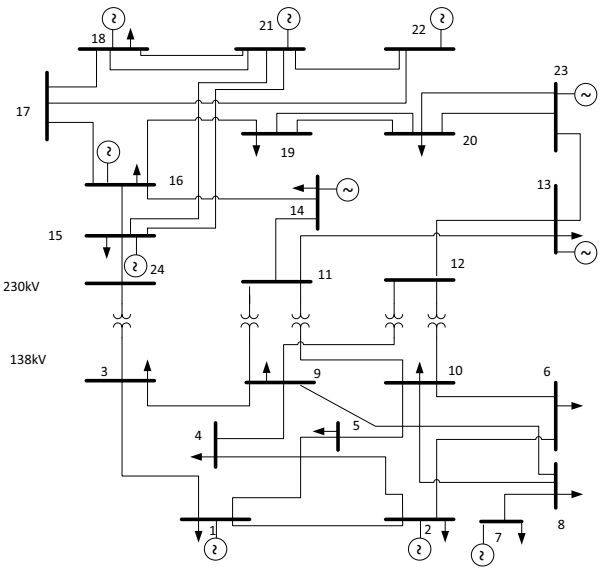

Figure 2. IEEE RTS 24 bus system

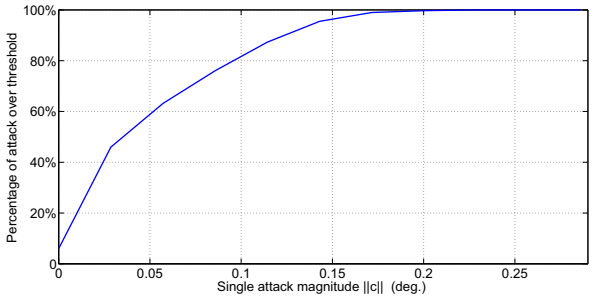

Figure 3. The percentage of attacks above threshold as the the single state error c_i increases

A. DC and AC attack

In this simulation, we focus on the case $\|c\|_0 = 1$ and the only non-zero entry c_k corresponds to the voltage angle of bus k. Note that there are 19 injection buses; we consider an attack centered at each one. To evaluate the performance of AC and DC attacks, we vary the value of c_k and compare the residual of the AC state estimator with χ^2 threshold.

1) DC Attack: We summarize results for DC attack model in Table I. The table gives the size of attack subgraph for each attack scenario, as well as the value of c_k at which the mean residual crosses the χ^2 threshold. In Fig. 3, over 100 attack simulations per attack bus, we plot the percentage of attacks above the threshold as the function of attack magnitude $\|c\|$. Observe that percentage above threshold increases quickly as $\|c\|$ increases; in fact, for $\|c\|$ as small as 0.2 degrees, virtually all DC attacks are detectable. Specifically, for buses 4 and 10 we plot the residual as a function of c in Figs. 4 and 5, respectively. Target buses 4 and 10 are representative of attacks on buses with relatively larger and smaller subgraph, respectively.

2) AC Attack: Also plotted in Fig. 4 and 5 are the residuals when the attacker uses an local AC state estimation for the same values of c_k. As expected, the residuals resulting from

Table I
SUMMARY OF DC ATTACK

Center bus k	Number of buses in S_k	Number of branches in S_k	Post-threshold c_k (deg.)
1	4	3	0.0228
2	4	3	0.0228
3	5	4	0.0915
4	3	2	0.1145
5	3	2	0.0915
6	3	2	0.0915
7	2	1	0.1087
8	4	3	0.0743
9	10	13	0.0572
10	10	13	0.0457
14	6	5	0.0401
15	5	5	0.0228
16	7	6	0.0170
18	5	6	0.0170
19	3	3	0.0228
20	3	4	0.0170
21	4	5	0.0170
22	5	6	0.0858
23	6	7	0.0228

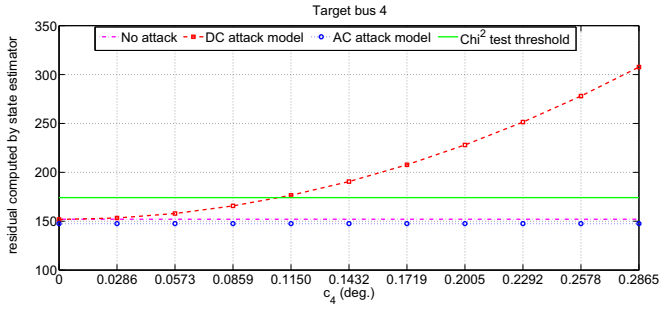

Figure 4. Residual for DC and AC attacks as the attacker increases the voltage angle of bus 4

the AC attack model are always below the χ^2 threshold irrespective of the value of c_k. More interestingly, we note that the average residual is even smaller than the no attack case because the false data injections resulting from the AC attack model are noise-free.

B. Consequence for unobservable attack

We now describe a physical consequence of the AC attack model. We assume that the system is operating at normal state

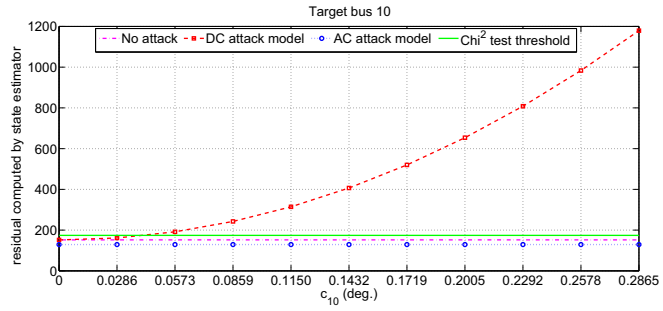

Figure 5. Residual for DC and AC attacks as the attacker increases the voltage angle of bus 10

prior to the attack. After the attack is launched, it may trick the operator into thinking that the normal state has moved to an emergency state (some operational limits are violated) or a restorative state (partial or total blackout). Either way, it is possible that such an attack leads to additional control actions that changes the physical system, including topology, generation dispatch, load shedding schedule, and so forth.

For instance, suppose that the attacker injects false data such that the estimated voltage angle at bus 7 is increased by 4.01 degrees. The absolute power flow measured at bus 7 side of branch 7-8 is increased from 89.40 MVA to 196.19MVA, which exceeds its long term rating of 175 MVA. The control center observes this abnormality and considers the system to be in emergency state in need of corrective control. If the attacker has knowledge about emergency control procedure, then it is possible for the attacker to influence the dynamics of the physical system.

We simulate this emergency response at the control center as follows:

1) The system is modeled as operating at an optimal power flow situation and the load of the system is constant during the attack period.
2) An ACOPF with a minimum cost objective function is applied as an emergency or corrective control procedure to re-dispatch generation when the operator monitors any line limit violation.
3) The system is assumed to have congestion. Long term ratings of all lines are degraded proportionally to let one line be congested just prior to the attack, specifically line 6-10. This assumption is made because IEEE 24 bus RTS is a system with redundant transmission capacity.

Suppose the state estimator runs every time unit. At time $t = 1$, the attacker constructs an unobservable attack vector c such that $\|c\|_0 = 1$ and $c_7 = 2.865$ degree. The absolute power flow of line 7-8 (measured at bus 7 side) increases from 89.40 MVA to 165.00 MVA, which is 101.11% of the long-term rating. This attack causes the estimated load at bus 7 to decrease and estimated load at bus 8 to increase. It triggers an alarm and the emergency control is involved. Then, as a result, the control center re-dispatches the generation via ACOPF to eliminate the false line rating violation. Following this initial attack, the attacker continues to use the same strategy and injects the same c into the system at subsequent estimation intervals. As shown in Fig. 8, at time $t = 1$, the generation level at bus 7 reduces and that of bus 13 increases. Fig. 8 also shows that after time 1, changes in active power generation are minor and caused only by measurement errors. Thus, an unobservable attack on a single bus led to a physical generational re-dispatch. Specifically, the generators at bus 7 reduce generation to decrease line flow from bus 7. To ensure the power balance, generators at bus 13 increase generation. Fig. 6 and 7 show some of the real states evolutions during the attack. Fig. 9 to 11 show a different case at bus 2 with $c_2 = 1.719$ degree.

V. CONCLUSION

In this paper, we have introduced attack models on AC state estimation. We have shown that, with partial information

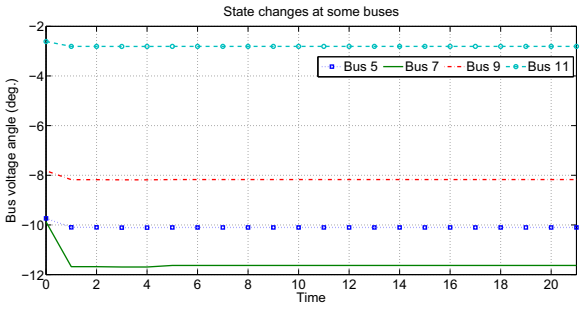

Figure 6. For attack centered at bus 7, real state evolution at bus 2, 7, 9 and 11: Angle (degree)

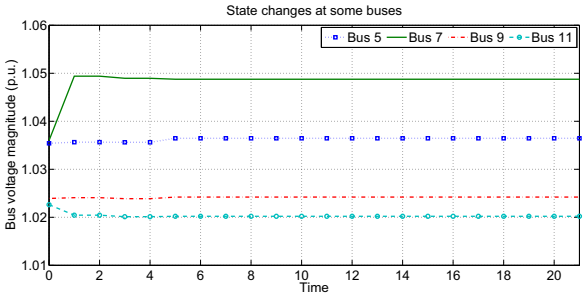

Figure 7. For attack centered at bus 7, real state evolution at bus 2, 7, 9 and 11: Magnitude (p.u.)

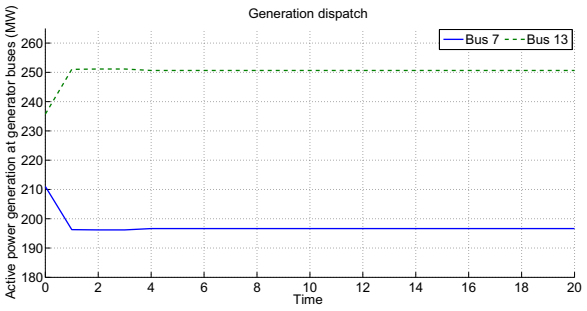

Figure 8. For attack centered at bus 7, active power generation dispatch before and after attack. Attack starts at $t = 1$

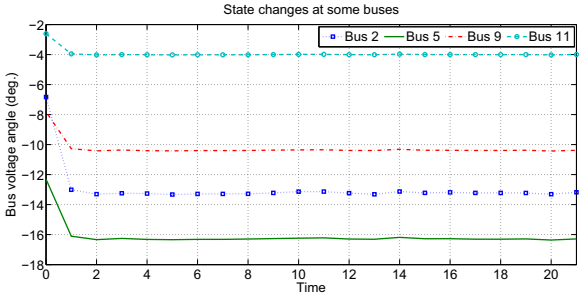

Figure 9. For attack centered at bus 2, real state evolution at bus 2, 5, 9 and 11: Angle (degree)

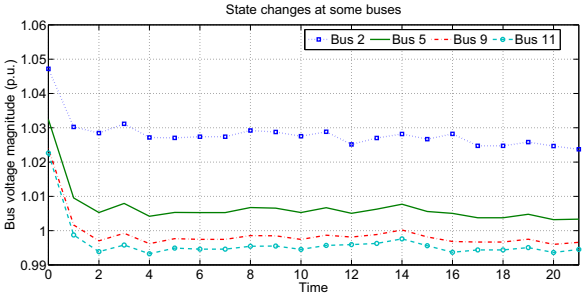

Figure 10. For attack centered at bus 2, real state evolution at bus 2, 5, 9 and 11: Magnitude (p.u.)

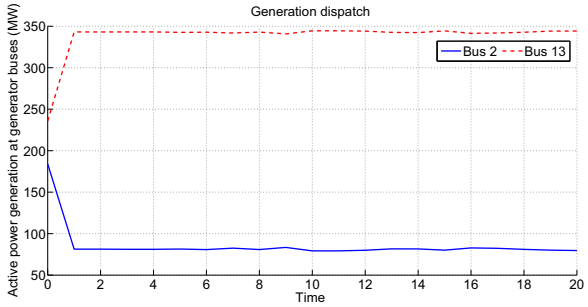

Figure 11. For attack centered at bus 2, active power generation dispatch before and after attack. Attack starts at $t = 1$

of system configuration and measurements, an attacker can bypass the bad data detector and inject false data into the AC state estimator. In contrast, attacks based on DC models can be detected using AC state estimation even if the attack magnitude is quite small. We have also highlighted a physical consequence of such attack. Future work will include developing robust detection mechanisms; this will require a better understanding of the physical consequences of attacks.

REFERENCES

[1] Y. Liu, M. K. Reiter, and P. Ning, "False data injection attacks against state estimation in electric power grids," in *in Proc. 16th ACM Conference on Computer and Communication Security*, 2009, pp. 21–32.

[2] H. Sandberg, A. Teixeira, and K. H. Johansson, "On security indices for state estimators in power networks," in *Proc. 1st Workshop Secure Control Syst.*, 2010.

[3] G. Dan and H. Sandberg, "Steath attacks and protection shcemes for state estimators in power systems," in *Proc. 1st IEEE SmartGrid Comm.*, 2010.

[4] A. Teixeira, S. Amin, H. Sandberg, K. Johansson, and S. Sastry, "Cyber security analysis of state estimators in electric power systems," in *2010 49th IEEE Conference on Decision and Control (CDC)*, 2010, pp. 5991–5998.

[5] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on smart grids," *IEEE Transactions on Smart Grid*, vol. 2, pp. 645–658, 2011.

[6] F. Pasqualetti, R. Carli, and F. Bullo, "A distributed method for state estimation and false data detection in power networks," in *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2011, pp. 469–474.

[7] G. Hug and J. Giampapa, "Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1362–1370, 2012.

[8] A. Abur and A. G. Expódsito, *Power system state estimation: theory and implementation*. CRC, 2000.