Effect Of Stealthy Bad Data Injection On Network Congestion In Market Based Power System

Mohammad Esmalifalak[†], Zhu Han[†], and Lingyang Song[‡]
[†]ECE Department, University of Houston, Houston, TX, USA
[‡]School of Electrical Engineering and Computer Science, Peking University, Beijing, China

***Abstract*—In a smart grid, the strong coupling between cyber and physical operations makes power systems vulnerable to cyber attacks. Changing the traditional structure of power systems and integrating communication devices are beneficial for better monitoring and decisionmaking by System Operators but increases the chance of being maliciously attacked. The communication links can be hacked so that the attacker can alter the power flow and power injection measurements, which are used to estimate the states of power system. In this paper, we formulate an attack strategy that can change the congestion of transmission lines without being detectable. Moreover, the financial benefit in an Ex-Post market is also investigated. Simulation results on an IEEE 30-Bus test system shows both the changes of congestion and the potential financial benefit gained by an attacker.**

***Index Terms*— State Estimation; Stealthy Bad Data Injection Attack; Locational Marginal Price, Ex-Ante/Ex-Post Market.**

I. Introduction

Due to increase in electricity demand and limited energy resources, power systems become more and more complex in structure and should operate in a secure and optimal conditions. The difAfter deregulation, the physical and financial operations of electric power systems become significantly different and challenging. State estimation (SE), introduced to improve the operating point of the power system from both technical and economical viewpoints, helped engineers in energy control centers (ECC) to monitor all states of the power system, which was previously not practical (or economical) to monitor. Currently, SE plays an important rule in secure and optimal operation of power systems [1], [2]. Accuracy of state estimation can be affected by bad data in measurements. These bad data could be caused by unintended measurement abnormalities or topology errors, or even by injection of malicious attacks. Cyber-attacks can be increased by integrating more advanced cyber technologies into the energy management system (EMS) and can cause major technical problems such as blackouts in a power system. These attacks also can be designed for illicit financial benefit by changing the optimal operation of the system, thereby increasing the net cost of electricity for consumers.

A cyber attack in 2003, caused a blackout in the eastern US and Canada [3]. This attack alarmed cyber-security decision makers motivating them to define critical infrastructure that is vulnerable to cyber-attack. In 2007, researchers at the Idaho National Lab tried to attack a synchronous generator. This attack was successful and the generator was self-destroyed in a couple of minutes [4]. In [5], analysis of cyber-security vulnerabilities of a supervisory control and data acquisition (SCADA) center is done at three levels: system, scenario, and access points. In [6], an undetectable attack by bad data detectors (BDD) was first introduced, where the attacker knows the state estimation Jacobian matrix (H) and defines an undetectable attack using this matrix. In [7], it is shown that even if the attacker does not know this matrix, similar undetectable attack can be inserted by independent component analysis. In [8], a security measure is defined to quantify how hard attacks are to perform, and then an efficient algorithm is constructed to compute the security measure. Potential financial misconduct that may be induced from cyber attack was discussed in reference paper [9].

In this paper we show that attacker can compromise measurements and insert false data into state estimation. This attack will be undetectable and stealthy enough to defeat traditional bad data detection methods and able to change the transmitted power in transmission lines. This change can increase or decrease the congestion level in the attacked lines, which in a market-based environment will change the price of congestion and Locational Marginal Prices. Our contribution is that the optimization problems are formulated to link the stealthy attack, power congestion and market prices. The simulation results on IEEE 30-Bus test system shows the changing of the line congestion by a stealthy attack, which provides the financial benefits to attacker.

The rest of the paper is organized as follows. The system model and basic concepts of power system state estimation and false data injection are introduced in Section II. Section III describes the false data injection problem. A rational attack is formulated from the attacker point of view in Section IV. The structure of the electricity market and the incentive for formulated attack is presented in Section V. A formulated attack is tested on an IEEE-30 bus test system and the results are presented in Section VI. For the sake of clarity, we show the abbreviations and notations in Table I and Table II, respectively.

II. System Model

In power systems, transmission lines transfer complex power from generating units to load centers. As described by circuit theory, this power is a function of line impedance[1] and

[1] 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.

978-1-4673-0437-5/12/$31.00 ©2012 IEEE

TABLE I
ABBREVIATION

BDD	Bad Data Detection
ED	Economic Dispatch
ECC	Energy Control Center
EMC	Energy Management Center
IP	Integer Programming
ISO	Independent System Operator
LMP	Locational Marginal Price
OPF	Optimal Power Flow
PMU	Phasor Measurement Units
SE	State Estimation
WLS	Weighted Least Squares

TABLE II
NOTATIONS

P_{ij}	Transmitted Power from bus i to bus j (MW)
$\mathbf{P_G}$	the active power generation vector (MW)
$\mathbf{P_D}$	the active power consumption vector (MW)
V_i	Voltage at bus i $(PerUnit)$
X_{ij}	Line reactance between bus i and j (Ω)
$\mathbf{z} = [z_1, \cdots, z_m]^T$	Active power measurements (MW)
$\theta = [\theta_1, \ldots, \theta_n]^T$	State vector (Rad)
$\mathbf{e} = [e_1, \cdots, e_m]^T$	Measurement noise vector with covariance $\mathbf{\Sigma}_e$
$\mathbf{z}_0$	Measurement vector w/o attack (MW)
$\mathbf{z}_{att} = Ha$	Attack vector (MW)
$\mathbf{B}$	Jacobian Matrix of power balance (S)
$\mathbf{H}$	Jacobian Matrix of linearized line flows (S)
$\mathbf{a}$	False injection vector (Rad)
$\mathbf{r}$	Residue vector for measurements (MW)

voltage on both ends of the line and can be obtained by [10]

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

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 linear state estimation, the linear form of (1) should be used which, can be obtained by:

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

this approximation assumes that amplitudes of voltages in buses are near unity and the difference of phase angles for two buses is small [11].

In power systems, measuring active power in a line (or injected active power in buses) is much easier than measuring angles. On the other hand, a control center needs to monitor angles of each bus in order to make corrective decisions for real-time operation of the system. In this regard, control centers try to use as many measurements as possible (redundant measurement) to have a better estimation during system monitoring. The linear approximation model of measurement vector $\mathbf{z} = [z_1, \cdots, z_m]^T$ can be described as:

$$\mathbf{z} = \mathbf{H}\theta + \mathbf{e}, \tag{3}$$

where $\theta = [\theta_1, \ldots, \theta_n]^T$ is the state vector[2], $\mathbf{H}$ is the Jacobian matrix of the power system and $\mathbf{e} = [e_1, \cdots, e_m]^T$ is the Gaussian measurement noise vector with zero mean and covariant matrix $\mathbf{\Sigma}_e$.

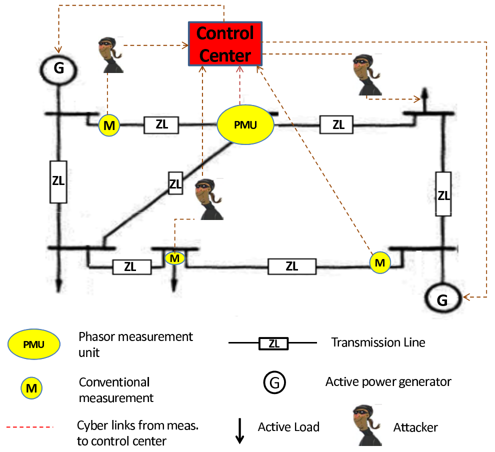

Fig. 1. Physical and cyber layer of a 6-bus power system

State estimation in power systems usually refers to finding best estimate of state vector θ from measurement vector $\mathbf{z}$. In (3), the number of equations (the number of measurements) are more than the number of unknown (state variables). These equations can be solved with Weighted Least Squares (WLS) estimator by [1]

$$\hat{\theta} = (\mathbf{H}^T \mathbf{\Sigma}_e^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{\Sigma}_e^{-1} \mathbf{z}. \tag{4}$$

Figure 1 illustrates the system model for state estimation of a power system, where PMU stands for phasor measurement unit. The measurements are obtained for each transmission line[3] and are reported to control center for the state estimation in (4).

III. Bad Data in State Estimation

The measured data can be faulty because of measurement abnormalities or malicious bad data injection [6]. These measurements have direct impact on state estimation functions in power system monitoring. In order to detect the false data or faulty measurements, the control center uses different routines such as the χ^2 method[4] and the maximum residue vector method [1] that we will discuss next.

A. Detection of False Data Injection

After estimation of angles using (4), the residue vector $\mathbf{r}$ can be computed as the difference between the measured quantity $\mathbf{z}$ and the calculated value $\hat{\mathbf{z}}$ from the estimated state:

$$\mathbf{r} = \mathbf{z} - \hat{\mathbf{z}} = \mathbf{z} - \mathbf{H}\hat{\theta}. \tag{5}$$

The hypothesis of not being attacked is accepted if

$$\max_i |r_i| \leq \gamma, \tag{6}$$

[2] In power flow studies, the voltage phase angle (θ_i) of the reference bus is fixed and known, and thus only $n-1$ angles need to be estimated.

[3] Power injection measurement is another type of measurements and is defined as super-composition of lines power which are connected to the measured bus. For example injected active power to bus i is $P_i = \sum_j P_{ij}$.

[4] which is used to check if the difference between expected and observed results is significant or not.

where γ is the threshold and r_i is the i^{th} component of $\mathbf{r}$. If the condition in (6) is not satisfied, the alarm is triggered for the false data in the i^{th} measurement.

B. Stealth (Unobservable) Attack

Recently, a certain type of attack was introduced by [6], which is not detectable by max residue vector r. In this attack, an attacker with knowledge of topology $\mathbf{H}$, can add $\mathbf{z}_{att} = \mathbf{Ha}$ to $\mathbf{z}_0$ (the measurement vector without attack). As a result,

$$\mathbf{z} = \mathbf{z}_0 + \mathbf{z}_{att} = \mathbf{z}_0 + \mathbf{Ha}. \tag{7}$$

Substituting (7) in (4) gives:

$$\hat{\theta} = \hat{\theta}_0 + \mathbf{a}, \tag{8}$$

where $\hat{\theta}_0 = \mathbf{M}\mathbf{z}_0$ and

$$\hat{\mathbf{z}} = \mathbf{H}\hat{\theta} = \mathbf{H}\hat{\theta}_0 + \mathbf{Ha}. \tag{9}$$

Substituting (7) and (9) in (5), we have

$$\mathbf{r} = \mathbf{z} - \hat{\mathbf{z}} = \mathbf{z}_0 - \mathbf{H}\hat{\theta}_0. \tag{10}$$

Equation (10) shows that residue r will not be affected by the injected attack vector to measurements, since the resulting $\mathbf{r}$ has the same mean and variance as before attack. As a result, the hypothesis test fails in detecting the attacker. In fact, the control center believes that the true state is $\theta_0 + \mathbf{a}$. Consequently, this is called *stealth false data injection.*

IV. Attack Against Voltage Angles in State Estimation

As described in Section III, if the control center cannot detect the attacked measurements, the results of state estimation will be incorrect and have negative effects on the control center decisions. In this section, we will show that the attacker can compromise the measurement vector and change the state of the system. Change in transmitted power can modify congestion levels, which are closely related to the price at which electricity trades in most markets. We formulate the problem to increase or decrease transmitted active power in the desired transmission lines by injecting bad data as follows:

$$\min_{\mathbf{a}} \quad \sum_{\{ij\}\in\mathcal{M}} P_{ij} - \sum_{\{ij\}\in\mathcal{N}} P_{ij} \tag{11}$$

$$= \left(\sum_{\{ij\}\in\mathcal{M}} \mathbf{H}_{ij} - \sum_{\{ij\}\in\mathcal{N}} \mathbf{H}_{ij} \right) \hat{\theta}$$

$$\text{s.t.} \begin{cases} \mathbf{Z_{min}} \leq \mathbf{Ha} \leq \mathbf{Z_{max}}, \\ C_{min} \leq \mathbf{C}_{att}^T \mathbf{Ha} \leq C_{max}. \end{cases}$$

The objective of the above optimization is to decrease and increase transmitted power, respectively, in group $\mathcal{M}$ and $\mathcal{N}$ of transmission lines represented by $\{ij\}$. $\mathbf{H}_{ij}$ are rows of $\mathbf{H}$ corresponding to each line $\{ij\}$, in which the attacker would like to decrease or increase transmitted power[5]. $\hat{\theta}$ is a function of attacking parameter $\mathbf{a}$ in (8). The two constraints impose limitations on injected measurement and total cost of attack, respectively, where $\mathbf{C}_{att}$ is a vector representing the cost of attacking each line. Optimization problem in (11) tries to change the estimated transmitted power in the system (without triggering the bad data detection alarm). In order to describe potential benefits for the attacker, the following section will briefly introduce the structure of electricity market.

V. Ex-Ante and Ex-Post Electricity Market

Optimal and secure operation of a power system is an important goal for ISO. In this regard, ISO needs to find the best schedule of generation and consumption[6]. This schedule considers transmission lines capacities, limitations in rate and capacity of generators, and security considerations. As a result, ISO may run different scheduling or optimization programs such as Power Flow (PF), Optimal Power Flow (OPF), Security Constrained Optimal Power Flow (SCOPF), Economic Dispatch, and Security Constrained Economic Dispatch (SCED) [12], [13], [14]. DC optimal power flow (DCOPF) is linear format of OPF that ISO usually uses for calculating the Locational Marginal Prices (LMPs)[7].

In the Ex-Ante model, the generation dispatches and LMPs are obtained from the same optimization model. In the Ex-Post model, the dispatch is performed at Ex-Ante, while the LMP is calculated after the cycle of the spot market, i.e., at Ex-Post such as after the 5-minute or hourly real-time market, using an incremental dispatch model [15]. The problem formulation of DCOPF is given by

$$\min_{P_G, P_D} \quad f = \mathbf{C_G}^T \mathbf{P_G} - \mathbf{C_D}^T \mathbf{P_D} \tag{12}$$

$$\text{s.t.} \begin{cases} \mathbf{B}\theta = \mathbf{P_G} - \mathbf{P_D}, \\ \mathbf{F_l^{min}} < \mathbf{H}\theta < \mathbf{F_l^{max}}, \\ \mathbf{P_G^{min}} < \mathbf{P_G} < \mathbf{P_G^{max}}, \\ \mathbf{P_D^{min}} < \mathbf{P_D} < \mathbf{P_D^{max}}, \end{cases}$$

where $\mathbf{C_G}$ is bid vector for supplying active power, $\mathbf{C_D}$ is bid vector for consumption of active power, $\mathbf{P_G}$ is the active power generation vector in Mega-Watts (MWs), $\mathbf{P_D}$ is the active power consumption vector in Mega-Watts (MWs), $\mathbf{F_L} = \mathbf{H}\theta$ is the vector of transmitted active power, and $\mathbf{B}$ is defined as $B_{ij} = -\frac{1}{X_{ij}}$ and $B_{ii} = \sum_{j=1}^{n} \frac{1}{X_{ij}}$. The first constraint of this optimization shows the balance between generation and consumption of active power in each bus. The second, third, and fourth constraints consider the thermal limitation in transmission lines, active power generation limits, and active power consumption limits, respectively. There are different uncertainties in actual operating states for all power systems such as load forecast uncertainty and unpredicted line or generator outages. These changes can be considered in a real-time market (Ex-Post market), which uses similar

[5] From (1), transmitted power in some of lines could be negative which means the direction of power flow is opposite of primary assumed direction. Congestion in these lines means the active power flow wants to be less than minimum value which is specified in the OPF program. For removing congestion in these lines, attacker needs to maximize the power flow in the primary assumed direction in (11).

[6] In market based power systems, generators and loads compete to supply and consume electricity respectively.

[7] LMP is the price of electricity in some electricity markets such as PJM Interconnection, New York, and New England.

optimization procedure to cover the latest changes in the power system.

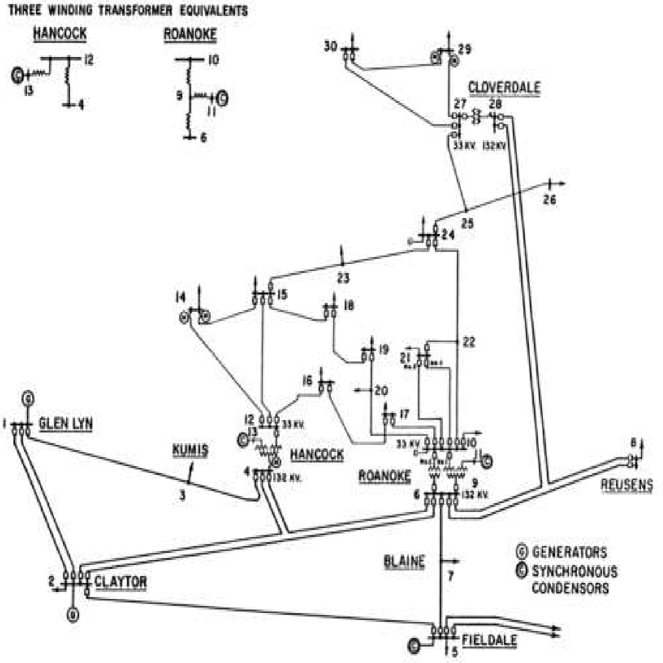

Fig. 2. IEEE 30-Bus Test System

We try to solve the linear optimization problem in (12) by the Lagrangian method. The Lagrangian multiplier for the first constraint in (12) is called the LMP, which could be obtained by:

$$\mathrm{LMP}_i = \frac{\partial f}{\partial P_i} \quad \forall i = 1, ..., n, \tag{13}$$

where $P_i = P_{g_i} - P_{d_i}$. LMP at bus i reflects the magnitude of change in cost function because of one unit change in the amount of active power in bus i. So, this value can be used as the price of electricity in each bus. In order to decrease the total cost of operation (f), the optimization problem in (12) will dispatch generators in an orderly fashion, starting with those with a low price and incrementally selecting those with higher prices. In this procedure, the last dispatched generator[8] will define LMP because the active power change in each bus will be provided by this generator so that the LMP will be the same in all buses. But in the case of congestion, this change will be covered by several marginal generators so that the LMP will be different in buses.

In addition to benefiting financially, an attacker could also cause power system blackout. Practically, a transmission line has different types of protecting relays. For example, in the case of overloading (e.g., the results of this paper's simulated attack) special relays will disconnect the line (to prevent overheating and physical damage). This line outage (especially during peak hours) would reduce the transmission capacity, increasing the chance of instability in the power system.

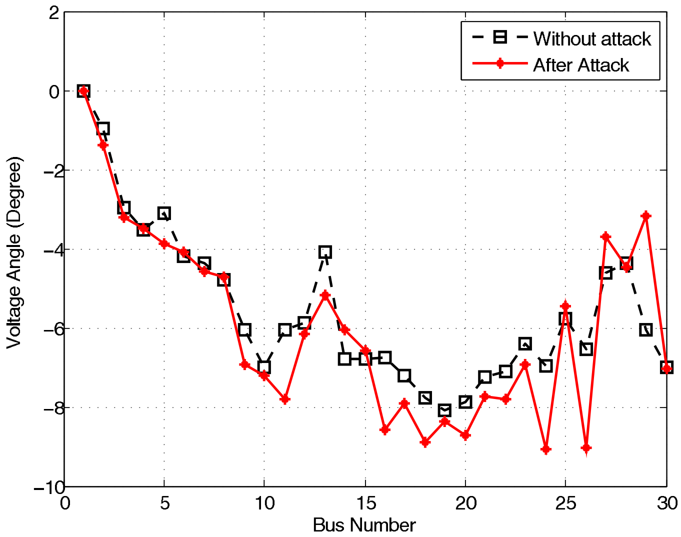

Fig. 3. Voltage angles in different buses

VI. Simulation Results

In this section, we show the effects of the presented attack on the IEEE 30-Bus test system shown in Fig. 2. In order to define LMPs in the Ex-Ante market, the control center solves DC optimal power flow in (12). This optimization is solved by Matpower, a package of MATLAB M-files for solving power flow and optimal power flow problems [16]. The attack targets line 29 in the power system.

In Fig. 3, we show the voltage angles for different buses with and without attack. We can see that the phase estimation is modified without being detected. Consequently, simulation results show that line 29 (from bus 21 to bus 22) is congested in Fig. 4. For comparison, we also show in this figure the case without attack and the thermal limits. As described in Section V, this congestion changes the marginal generator and as a consequence, the network will have different LMPs in its buses as shown in Fig. 5. The financial benefit is given by the following example.

Releasing congestion can change LMPs, so an attacker solves (11) and inserts an undetectable attack[9]. Due to the stochastic nature of loads, the control center believes that there is no congestion in the network (for example transmitted power through line 29, is less than its thermal limit). Based on these results, the control center will use the free (but fake) capacity in line 29 and run the Ex-Post program for the real-time market. This time, because of released congestion, LMPs will be the same in the network (Fig. 5). If, for example, attacker buys $P^{MW} = 10$ at bus 22 in the Ex-Ante market and sells it in the Ex-Post market in the same bus, the profit of this transaction will be:

[8] This generator is usually called "Marginal Generator"

[9] Practically attacker could insert false data to measurements by, changing the bias of measurements or hacking and sending the desired values to control center.

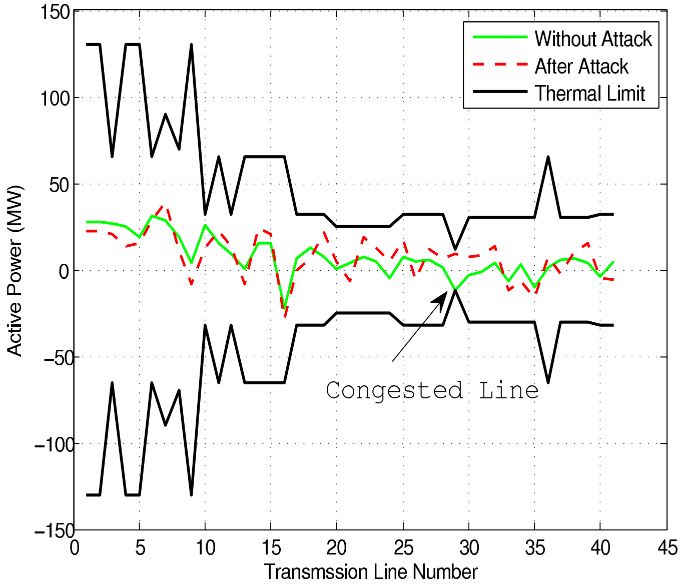

Fig. 4. Transmitted active power in transmission lines

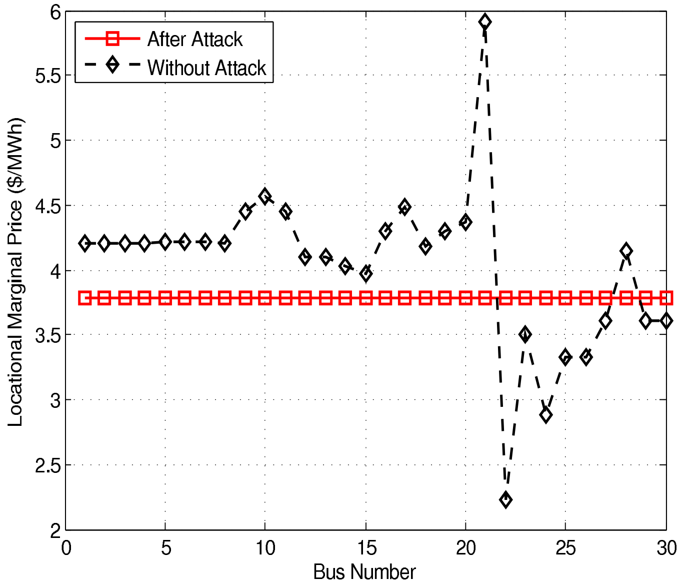

Fig. 5. Locational marginal price in buses

$$
\begin{aligned}
Profit_{att}^{(\$/h)} &= P^{(MW)}\left(LMP_{Ex-post}^{(\$/MWh)} - LMP_{Ex-Ante}^{(\$/MWh)}\right)\\
&= 10^{MW}(3.8^{\$/MWh} - 2.2^{\$/MWh}) = 16^{\$/h}.
\end{aligned}
\tag{14}
$$

In summary, the attacker can obtain the financial gain through changing power line congestion by stealthy bad data injection without being detected.

VII. Conclusion

In this paper, we demonstrate the effect of stealthy false data injection on power system congestion. The problem solution links stealthy attack, power congestion, and the resulting market price. We show that by changing congestion, the attacker can change LMPs and obtain financial benefit in an Ex-Post market. Besides financial misconduct, the attacker can also overload specific lines, which consequently increases the chance of line outages. We test our proposed attack in an IEEE 30-bus test system using MATPOWER and show the profitability of such an attack.

Acknowledge

This work is partially supported by US NSF CNS-1117560, ECCS-1028782, CNS-0953377, and CNS-0905556. We would like to thank Dr. Rong Zheng of the University of Houston computer science department for initial consultation.

References

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

[2] A. Monticelli, “Electric Power System State Estimation,” *Proceedings of the IEEE*, vol. 88, no. 2, pp. 262–282, Feb. 2000.

[3] D. White, A. Roschelle, P.Peterson, D.Schlissel, B. Biewald, and W. Steinhurst, “The 2003 Blackout: Solutions that Wont Cost a Fortune” *The Electricity Journal.*, vol. 16, no. 9, pp. 43–53, Oct. 2003

[4] Staged cyber attack reveals vulnerability in power grid, “CNN report” *http://www.cnn.com/2007/US/09/26/power.at.risk/index.html,*, [Sep. 26, 2007].

[5] C. Ten, C. Liu and G. Manimaran, “Vulnerability assessment of cybersecurity for SCADA systems” *IEEE Trans. Power Syst.*, vol. 23, no. 4, pp. 1836–1846, Nov. 2008

[6] Y. Liu, M. K. Reiter, and P. Ning, “False data injection attacks against state estimation in electric power grids,” *Proceedings of the 16th ACM conference on Computer and communications security*, pp. 21–30, Nov. 2009.

[7] M. Esmalifalak, H. Nguyen, R. Zheng, and Z. Han “Stealth False Data Injection using Independent Component Analysis in Smart Grid”, *Second IEEE second conference on smart grid Communications*, Brussels, Belgium, Oct. 2011.

[8] G. Dan, H. Sandberg. “Stealth attacks and protection schemes for state estimators in power systems” *First IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 214–219, Oct. 2010.

[9] 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.

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

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

[12] B.H. Chowdhury, S. Rahman, “A review of recent advances in economic dispatch”, *IEEE Transactions on Power Systems*, vol. 5, no. 4, p.p. 1248–1259, Nov. 1990.

[13] D. Sun, B. Ashley, B. Brewer, A. Hughes, and W. Tinney, “Optimal power flow by Newton approach”, *Power Apparatus and Systems, IEEE Transactions on Power Systems*, p.p. 2864–2880, Oct 1984.

[14] H. Dommel, and W. Tinney, “Optimal power flow solutions”, *IEEE Transactions on Power Systems*, vol. 87, no. 10, p.p. 1866–1876, Oct 1968.

[15] F. Li, Y. Wei, S. Adhikari, “Improving an unjustified common practice in Ex Post LMP calculation: An expanded version”, *Power and Energy Society General Meeting*, p.p.1–4, Jul 2010.

[16] 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”, *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, Feb. 2011.