False-data Injection Attack to Control Real-time Price in Electricity Market

Suzhi Bi, *Student Member, IEEE*, Ying Jun (Angela) Zhang, *Senior Member, IEEE*
Department of Information Engineering, The Chinese University of Hong Kong,
Shatin, New Territories, Hong Kong. Email: {bsz009, yjzhang}@ie.cuhk.edu.hk

***Abstract*—The normal operation of electricity market requires accurate state estimation of the power grids. However, recent research shows that carefully synthesized false-data injection attacks can easily introduce errors to state estimates without being detected by the current security systems. In this paper, we analyze and formulate an effective false-data injection attack to control real-time electricity price at any tagged bus. It is observed that an adversary capable of false-data injection attack can induce false real-time electricity price by fabricating biased transmission congestion pattern. From the strategic level, we propose a simple algorithm that finds the effective congestion pattern with minor distortion to the normal system operation. For practical implementation, load redistribution attack, a special false-data injection attack which produces biased load estimates, is used to achieve the derived congestion pattern. We also propose a cost-aware neighborhood load redistribution attack, which only compromises limited measurements around the tagged bus. From theory to practice, our results here reveal the potential cyber vulnerabilities of current electricity market and would help to build more secure smart power grid in the future.**

I. Introduction

The deregulation of electricity market eliminates price control and encourages market entry of entities in all aspects of power generation, transmission and distribution, etc. To enhance the market efficiency, independent system operators (ISOs) are introduced as the third-party market regulators apart from energy consumers and suppliers [1]. One major responsibility of ISOs is to set fair electricity prices for market clearing. Currently, the dominant pricing method adopted by a number of ISOs in the United States is locational marginal pricing (LMP) [2]. Real-time LMP is calculated based on the power system state estimation produced by EMS/SCADA (Energy Management System and Supervisory Control and Data Acquisition) systems that monitor and control the power grids. Evidently, the normal operation of electricity market requires accurate state estimation that truthfully reflects the real-time power system operating conditions.

The integrity of state estimation is under mounting threat as we gradually transform the current electricity infrastructures to smart power grids. Smart power grids are more open to and physically accessible by the outside networks, such as office local area networks and smart meters that allow two-way communications between energy consumers and suppliers. With new entry points introduced to the power system, potential complex and collaborating malicious attacks are brought in as well. Liu *et al*. [3] showed that a new false-data injection attack could circumvent bad data detection (BDD) in todays SCADA system and introduce arbitrary errors to state estimates without being detected. Being aware of the imminent threats of false-data injection attack to power system, a number following studies are devoted to both understanding its attacking patterns and providing effective countermeasures [4]–[7]. For instance, [4] formulated optimized attacks with different resource constraints to compromise the state estimates. Besides, [6] proposed efficient defending mechanisms to protect any set of state estimates from false-data injection attacks.

The potential cost of false-data injection attack to power markets has also attracted much research interests. [8] showed that load redistribution attack, a special false-data injection attack that fabricates biased real-time load patterns, is capable of causing higher power generation cost or even regional blackout. Meanwhile, [9] showed that attackers can profit from virtual bidding investment by pushing line flow estimates below their security limits, inducing equal real-time LMPs at two tagged buses. The method to control real-time LMP at a specific bus using false-data injection attack is first studied in [10]. It showed that, by compromising the actual congestion pattern in the network, an adversary such as a power supplier can increase the real-time LMP to maximize its total revenue. Nevertheless, the proposed method to obtain the optimal congestion pattern and attack vector requires exhaustive evaluation of the entire vector space, which is of prohibitively high in computational complexity and lacking of intuitions.

In this paper, we derive specific guidelines to formulate efficient false-data injection attacks to control the real-time LMP at any tagged bus. Our contributions are detailed below:

1) We derive the condition for a congestion pattern to compromise the real-time LMP at a specific location. We also propose a simple "add-then-remove" procedure that finds the effective congestion pattern which causes minor distortion to the true system operating condition.
2) A practical load redistribution (LR) attack is used to realize the obtained congestion pattern. In particular, we propose a cost-aware "neighborhood" load redistribution (NLR) attack that only compromises limited measurements around the tagged bus.
3) We show that LR attack to real-time LMPs also has impact on the future electricity prices due to its distortion of load prediction.

This research was supported, in part, by the General Research Fund (Project No. 419509) established under the University Grant Council of Hong Kong.

978-1-4799-1353-4/13/$31.00 ©2013 IEEE

II. System Model

A. Power system state estimation

We consider a steady-state power system with n buses interconnected by t transmission lines. For the simplicity of expositions, we assume that each bus has at least one generator and load. Notice that our main results still hold even without this assumption. The ISO monitors the power system through the measurements received by the SCADA host. Denoted by $\mathbf{z}$, the received measurements are

$$\mathbf{z} = \begin{pmatrix} \bar{\mathbf{p}} \\ \bar{\mathbf{d}} \\ \bar{\mathbf{f}} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{G} & -\mathbf{G} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{p} \\ \mathbf{d} \end{pmatrix} + \mathbf{e}, \tag{1}$$

where $\mathbf{p}$ and $\mathbf{d}$ represent the actual generation outputs and loads at the n buses. $\bar{\mathbf{p}}$ and $\bar{\mathbf{d}}$ denote the noisy measurements. $\bar{\mathbf{f}}$ represents the directional power flow measurements in the t transmission lines. $\mathbf{I}$ is an identity matrix. $\mathbf{G}$ is the $t \times n$ generating shift factor matrix, indicating the sensitivity of directional power flows against power injection at the buses. $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ contains independent measurement noises, where $\mathbf{R}$ is the diagonal covariance matrix. In general, the above linear measurement model can be written as

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{e}, \tag{2}$$

where $\mathbf{H}$ is the measurement Jacobian matrix and $\mathbf{x}$ is the state variable. The maximum likelihood (ML) estimate of $\mathbf{x}$, including generation outputs and loads, is

$$\hat{\mathbf{x}} = \begin{pmatrix} \hat{\mathbf{p}} \\ \hat{\mathbf{d}} \end{pmatrix} = \left(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{z} \triangleq \mathbf{P}\mathbf{z}. \tag{3}$$

Accordingly, the estimate of line flows is

$$\hat{\mathbf{f}} = (\mathbf{G}, -\mathbf{G})\,\hat{\mathbf{x}} \triangleq \mathbf{Q}\hat{\mathbf{x}}. \tag{4}$$

The estimates $\hat{\mathbf{x}}$ could be wrong due to random errors in measurements. Current power systems use BDD to compare the squared l_2-norm of measurement residual with a threshold τ, where it identifies bad data measurements if

$$||\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}||^2 > \tau. \tag{5}$$

Otherwise, $\mathbf{z}$ is considered as a normal measurement. In general, a random attack to normal measurements is likely to be detected by BDD. However, as we will introduce in Section IV, carefully synthesized false-data injection attacks can bypass BDD and compromise the power market.

B. Real-time LMP calculation

A combined ex-ante and ex-post method is widely adopted by major US power markets to calculate real-time LMP [2]. In the ex-ante stage, the ISO obtains the optimal generation dispatch instruction and ex-ante LMPs by solving a security constrained economic dispatch problem based on projected system load levels [1]. Then, the optimal generation dispatch instruction and ex-ante LMPs are sent to all generators as a reference, in the hope that generators would follow the dispatch instruction to mitigate transmission congestions.

At the end of dispatch interval, the ISO obtains the actual system response through state estimation. Assuming loads are not dispatchable, the ISO solves the following problem

$$\begin{aligned} & \underset{\triangle\mathbf{p}}{\text{minimize}} && \triangle C = \textstyle\sum_{i=1}^{n} c_i \cdot \triangle p_i \\ \text{s. t.} \quad & (\lambda) && \textstyle\sum_{i=1}^{n} \triangle p_i = 0 \\ & (\alpha_i^{min}, \alpha_i^{max}) && \triangle p^{min} \le \triangle p_i \le \triangle p^{max},\ \forall i \\ & (\mu_l^{max}) && \textstyle\sum_{i=1}^{n} G_{l,i} \cdot \triangle p_i \le 0,\ l \in \mathcal{C}_+ \\ & (\mu_l^{min}) && \textstyle\sum_{i=1}^{n} G_{l,i} \cdot \triangle p_i \ge 0,\ l \in \mathcal{C}_-. \end{aligned} \tag{6}$$

Here $\mathcal{C}_+$ and $\mathcal{C}_-$ denote the sets of observed positively congested lines and negatively congested lines, defined as

$$\mathcal{C}_+ \triangleq \left\{l : \hat{f}_l \ge f_l^{max}\right\} \quad \mathcal{C}_- \triangleq \left\{l : \hat{f}_l \le f_l^{min}\right\},$$

where $\hat{f}_l$ is the estimate of power flow of line l, f_l^{max} and f_l^{min} are the line flow limits. $\triangle p^{min}$ and $\triangle p^{max}$ are fixed parameters. The ex-post LMP is a function of the optimal Lagrangian multipliers of (6) [2], given by

$$\begin{aligned} LMP_i &= -\lambda^* - \sum_{l \in \mathcal{C}_+} \mu_l^{max*} G_{l,i} + \sum_{l \in \mathcal{C}_-} \mu_l^{min*} G_{l,i} \\ &= c_i + \alpha_i^{max*} - \alpha_i^{min*}, \quad i = 1, \cdots, n, \end{aligned} \tag{7}$$

where the last equality holds by KKT conditions. Ex-post LMP is also the real-time electricity price for market clearing. When load prediction is accurate and generators follow the generation dispatch, the ex-ante and ex-post LMPs are equal [11]. We see from (6) that real-time LMPs depend only on the ISO's congestion observations, i.e. $\{\mathcal{C}_+, \mathcal{C}_-\}$. Thus, an adversary capable of fabricating fake congestion pattern also has the ability to control electricity prices.

C. Adversary model

We assume that an adversary's objective is to reduce the real-time LMP at a tagged bus where at least one generator is located. A successful attack will either reduce its power consumption payment or increase its profit from certain financial derivative that favors lower LMP. Its capabilities include

1) the knowledge of the generation cost of the tagged bus, the topology of the power grid (the $\mathbf{G}$ matrix), the current load estimates $\hat{\mathbf{d}}$ and line flow estimates $\hat{\mathbf{f}}$ (thus the actual congestion pattern $\{\mathcal{C}_+, \mathcal{C}_-\}$).
2) able to perform load redistribution (LR) attack [8]. That is, it can compromise a given set of line flows and load measurements. The detailed characterization of LR attack will be presented in Section IV.A.

With the knowledge of the current system operating conditions, the adversary can choose to either perform LR attack immediately or postpone its attack until a more favorable time, e.g. lower detection probability or higher financial profit.

III. Malicious Congestion Pattern Derivations

In this section, we analyze and derive the effective congestion pattern to control the real-time LMP at a tagged bus. It is shown in the next section that the derived congestion pattern can be realized with false data injection attacks.

A. Attack analysis

Without loss of generality, we assume that the adversary's objective is to reduce the real-time LMP at the n^{th} bus.

To avoid confusions, we use the subscript "at" to denote variables in the presence of attack. For instance, $\triangle p_n^*$ denotes the optimal solution of (6) at the n^{th} bus when no attack is present, while $\triangle p_n^{*at}$ denoting the value after attack. By complementary slackness, real-time LMP_n in (7) is

$$LMP_n = \begin{cases} c_n - \alpha_n^{min*}, & \triangle p_n^* = \triangle p^{min} \\ c_n, & \triangle p^{min} < \triangle p_n^* < \triangle p^{max} \\ c_n + \alpha_n^{max*}, & \triangle p_n^* = \triangle p^{max}. \end{cases} \quad (8)$$

Since $\alpha_n^{min*}, \alpha_n^{max*} \geq 0$, LMP_n is an increasing step function with $\triangle p_n^*$. When $\triangle p_k^* = \triangle p^{max}$, attackers can efficiently decrease LMP_n through reducing $\triangle p_n^*$ by marginal amount, causing only minor distortion to the true measurements. For $\triangle p_n^* < \triangle p^{max}$, however, attackers will not be able to reduce LMP_n until $\triangle p_n^{*\,at} = \triangle p^{min}$. In this case, they may need to compromise a large number of measurements and induce large deviations from their true values. Since short-term load prediction is often accurate, such abnormal deviation is likely to be captured, even if the attack successfully passed BDD. Considering the overall efficiency and risk of detection, we assume that an rational adversary only performs attack when $\triangle p_n^* = \triangle p^{max}$, i.e. when the upcoming true $LMP_n > c_n$.

From the above analysis, an effective false congestion pattern, denoted by $\{\mathcal{C}_+^{at}, \mathcal{C}_-^{at}\}$, should cause the ISO to yield $\triangle p_n^{*\,at} \in \left[\triangle p^{min}, \triangle p^{max}\right)$. This is obtained by studying the following system that associated with the constraints of (6),

$$\begin{cases} \sum_{i=1}^{n-1} x_i = -\beta \\ \triangle p^{min} \leq x_i \leq \triangle p^{max}, \;\; i = 1, \cdots, n-1 \\ \sum_{i=1}^{n-1} G_{l,i} \cdot x_i \leq -G_{l,n} \cdot \beta, \; l \in \mathcal{A}_+ \\ \sum_{i=1}^{n-1} G_{l,i} \cdot x_i \geq -G_{l,n} \cdot \beta, \; l \in \mathcal{A}_-, \end{cases} \quad (9)$$

where $\mathcal{A}_+$ and $\mathcal{A}_-$ are two disjoint sets and β is a parameter. In can be inferred that a $\{\mathcal{A}_+, \mathcal{A}_-\}$ that renders (9) only feasible for some $\beta \in \left[\triangle p^{min}, \triangle p^{max}\right)$ can be used as $\{\mathcal{C}_+^{at}, \mathcal{C}_-^{at}\}$. This is because, such congestion pattern guarantees that the optimal solution of (6) is obtained at some $\triangle p_n^{*at} < \triangle p^{max}$. Intuitively, obtaining an effective $\{\mathcal{A}_+, \mathcal{A}_-\}$ needs to enumerate all possible sets. However, we propose in the following a "add-then-remove" heuristic to find a solution with low complexity. Conceptually, starting with $\{\mathcal{A}_+, \mathcal{A}_-\} = \{\mathcal{C}_+, \mathcal{C}_-\}$, we first add some inequalities to renders (9) infeasible when $\beta = \triangle p^{max}$, then remove other inequalities to restore its feasibility at some $\beta < \triangle p^{max}$.

B. Congestion pattern derivation

Depending on their effects to the feasibility of (9) when β decreases, all entries of $\{\mathcal{A}_+, \mathcal{A}_-\}$ can be classified into either a "tightening" set $\mathcal{A}_t$ or a "loosening" set $\mathcal{A}_l$, defined as

$$\begin{aligned} l \in \mathcal{A}_t &: (l \in \mathcal{A}_+ \wedge G_{l,n} \leq 0) \vee (l \in \mathcal{A}_- \wedge G_{l,n} > 0), \\ l \in \mathcal{A}_l &: (l \in \mathcal{A}_+ \wedge G_{l,n} > 0) \vee (l \in \mathcal{A}_- \wedge G_{l,n} \leq 0). \end{aligned} \quad (10)$$

Intuitively, the inequalities that correspond to $\mathcal{A}_t$, called tightening inequalities, become harder to satisfy when β decrease, and thus "tighten" the feasibility requirements of (9), while the loosening inequalities in $\mathcal{A}_l$ relax the feasibility requirements. Let $\{\mathcal{A}_t^0, \mathcal{A}_l^0\}$ denote the actual tightening and loosening sets when no attack is present, i.e. when $\{\mathcal{A}_+, \mathcal{A}_-\} = \{\mathcal{C}_+, \mathcal{C}_-\}$.

Algorithm 1: Procedure to find an attacking congestion pattern $\{\mathcal{A}_t^1, \mathcal{A}_l^1\}$ to compromise LMP_n.

input : true congestion pattern $\{\mathcal{A}_t^0, \mathcal{A}_l^0\}$, target bus n;
output: congestion pattern $\{\mathcal{A}_t^1, \mathcal{A}_l^1\}$ used for attack;
Initialize: $\{\bar{\mathcal{A}}_t, \bar{\mathcal{A}}_l, \mathcal{D}, \bar{\beta}\} = \{\mathcal{A}_t^0, \mathcal{A}_l^0, \emptyset, \triangle p^{max}\}$;
repeat
  $l^* = \arg\max_{l \in \{\mathcal{L} \setminus \{\bar{\mathcal{A}}_t \bigcup \bar{\mathcal{A}}_l\} \bigcup \mathcal{D}\}} |G_{n,l}|$;
  If (9) is infeasible when $\{\mathcal{A}_t, \mathcal{A}_l, \beta\} = \{\emptyset, \bar{\mathcal{A}}_l \bigcup l^*, \triangle p^{min}\}$, we let $\mathcal{D} = \mathcal{D} \bigcup l^*$. Otherwise, we update $\bar{\mathcal{A}}_l = \bar{\mathcal{A}}_l \bigcup l^*$.
until *(9) is infeasible when* $\{\mathcal{A}_t, \mathcal{A}_l, \beta\} = \{\bar{\mathcal{A}}_t, \bar{\mathcal{A}}_l, \bar{\beta}\}$;
repeat
  $\{\bar{\beta}, \bar{\mathcal{A}}_t\} = \{\bar{\beta} - \delta, \mathcal{A}_t^0\}$;
  while $\bar{\mathcal{A}}_t \neq \emptyset$ **do**
    $l^* = \arg\max_{l \in \bar{\mathcal{A}}_t} |G_{n,l}|$, $\bar{\mathcal{A}}_t = \bar{\mathcal{A}}_t \setminus l^*$;
    if *(9) is feasible when* $\{\mathcal{A}_t, \mathcal{A}_l, \beta\} = \{\bar{\mathcal{A}}_t, \bar{\mathcal{A}}_l, \bar{\beta}\}$ **then**
      output: $\{\mathcal{A}_t^1, \mathcal{A}_l^1\} = \{\bar{\mathcal{A}}_t, \bar{\mathcal{A}}_l\}$;
    end
  end
until $\bar{\beta} = \triangle p^{min}$;

It can be seen that (9) is feasible when $\{\mathcal{A}_t, \mathcal{A}_l, \beta\} = \{\mathcal{A}_t^0, \mathcal{A}_l^0, \triangle p^{max}\}$, since $\triangle p_n^* = \triangle p^{max}$ is part of the optimal solution to (6) when no attack is present. We first add a set of loosening inequalities to cause (9) infeasible. By doing this, the feasibility of (9) can be restored later through decreasing β and removing tightening inequalities. In practice, we select the line l^* on which the n^{th} bus influences the most, i.e.

$$l^* = \arg \max_{l \in \{\mathcal{L} \setminus \{\mathcal{A}_t \bigcup \mathcal{A}_l\}\}} |G_{l,n}|, \quad (11)$$

where $\mathcal{L}$ is the set of all transmission lines. Then, we update $\mathcal{A}_l$ by letting $\mathcal{A}_l = \mathcal{A}_l \cup l^*$ while keeping $\mathcal{A}_t$ intact. We continue to add other inequalities to $\mathcal{A}_l$ in sequence following (11) until (9) becomes infeasible. The obtained loosening set is denoted by $\mathcal{A}_l^1$. A point to notice is that $\mathcal{A}_l^1$ must ensure that (9) is feasible when $\{\mathcal{A}_t, \mathcal{A}_l, \beta\} = \{\emptyset, \mathcal{A}_l^1, \triangle p^{min}\}$. Otherwise, (9) cannot be converted feasible by removing tightening inequalities.

We then aim to restore the feasibility of (9) at some $\beta < \triangle p^{max}$. Since tightening inequalities become even harder to satisfy as β decreases, some of them may need to be removed. In each iteration, β is decreased by a small value δ. We immediately reach our objective if (9) becomes feasible. Otherwise, we check whether (9) can be converted into feasible by removing certain tightening inequalities. Similar to (11), we first remove the line $l^* = \arg\max_{l \in \mathcal{A}_t} |G_{n,l}|$. Then, we update $\mathcal{A}_t = \mathcal{A}_t \setminus l^*$ while keeping $\mathcal{A}_l$ intact. We continue to remove tightening inequalities until (9) becomes feasible. If (9) is still infeasible even if $\mathcal{A}_t = \emptyset$, we further relax the system by decreasing β and resetting $\mathcal{A}_t = \mathcal{A}_t^0$. The above iteration repeats until a feasible system is obtained.

It is worth noting that a feasible system can always be obtained, since (9) is feasible at $\{\emptyset, \mathcal{A}_l^1, \triangle p^{min}\}$ by construction. Meanwhile, it is unlikely for the obtained system to be feasible when $\beta = \triangle p^{max}$, since the feasibility of system is now dominated by loosening inequalities, while increasing β to $\triangle p^{max}$ will make all the loosening inequalities harder to satisfy. Denoted by $\mathcal{A}_t^1$, the obtained tightening set and $\mathcal{A}_l^1$ can be mapped back to $\{\mathcal{A}_+, \mathcal{A}_-\}$, which will then be used as the malicious congestion pattern to formulate attacks. A pseudo-

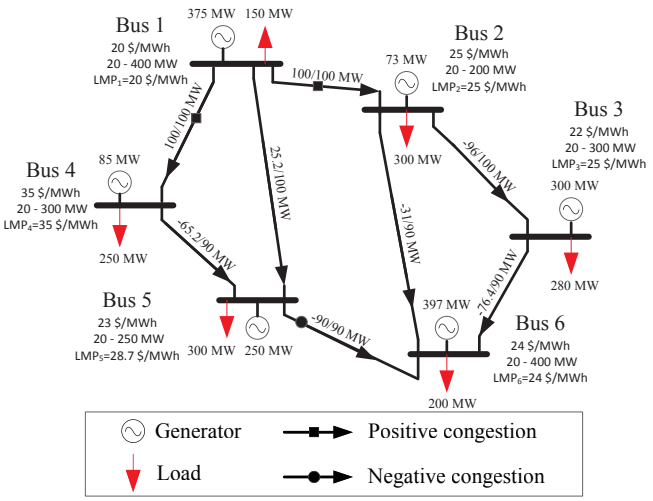

Fig. 1. Illustration of the 6-bus system: for instance, bus 1's generation cost is 20$/MWh, generation output range is $20-400$ MW. In the given operating condition, the actual load at bus 1 is 150 MW and the actual generation output is 375 MW. As a result, 100 MW power flows from bus 1 to 2, reaching the line capacity 100 MW. A negative power flow value means the power flows in the opposite direction of the defined line orientation, e.g. 96 MW power flows from bus 3 to 2 in line l_{23}, whose capacity is 100 MW.

code for the "add-then-remove" algorithm is Algorithm 1, where the lines 2 to 5 correspond to adding loosening constraints while lines 6 to 14 for removing tightening constraints.

C. A case study and insights

To bring out intuitions, we implement the proposed algorithm in an example $6-$bus system in Fig. 1, where $\mathbf{G}$ [1] is

$$\mathbf{G} = \begin{pmatrix} 0.4 & -0.4 & -0.4 & 0.2 & 0.8 & 0 \\ 0.2 & 0.15 & 0.15 & -0.4 & -0.4 & 0 \\ 0.4 & 0.25 & 0.25 & 0.2 & -0.4 & 0 \\ 0.25 & 0.4 & -0.6 & 0.15 & 0.5 & 0 \\ 0.15 & 0.2 & 0.2 & 0.05 & 0.3 & 0 \\ 0.25 & 0.4 & 0.4 & 0.15 & 0.5 & 0 \\ 0.2 & 0.15 & 0.15 & 0.6 & -0.4 & 0 \\ 0.6 & 0.4 & 0.4 & 0.8 & 0.2 & 0 \end{pmatrix}. \quad (12)$$

In the given operating condition, the congestion pattern is $\mathcal{C}_+ = \{l_{12}, l_{14}\}$ and $\mathcal{C}_- = \{l_{56}\}$. Suppose that bus 5 is the tagged bus, where actual $LMP_5 > c_5$. The corresponding $\mathcal{A}_t = \{l_{14}, l_{56}\}$ and $\mathcal{A}_l = \{l_{12}\}$. By Algorithm 1, attackers add l_{45} to $\mathcal{A}_l$, then remove l_{56} from $\mathcal{A}_t$. In this case, the congestion pattern of attack is $\mathcal{C}_+^{at} = \{l_{12}, l_{14}\}$ and $\mathcal{C}_-^{at} = \{l_{45}\}$.

With closer observation, we see that the selected lines are indeed those connect the tagged bus with its one-hop neighboring buses. To reduce the LMP at the tagged bus, attackers congest transmission lines of which the tagged bus is the immediate sending end, i.e. l_{45}, and decongest lines of which the tagged bus is the immediate receiving end, i.e. l_{56}. In general, attackers may also need to congest or decongest necessary lines between buses of multiple hops. The combined actions fabricate a "radial" power flow pattern centered at the tagged bus. This conveys a false message to the ISO that excessive power is flowing out of the tagged bus, which is the root cause of congestions. The ISO therefore reacts by "hypothetically"[2] reducing the generation at the tagged bus 5 in the ex-post LMP problem, i.e., $\triangle p_5^{*at} < \triangle p_5^*$. As a result, LMP at the tagged bus is reduced as well.

IV. Attack Implementations

In this section, we realize the malicious congestion pattern using LR attack, where different attacks are formulated depending on the attackers's control over measurements. Interestingly, we also find that LR attack to real-time LMP also has impact to future electricity market.

A. LR attack to real-time LMP

Suppose that a false data injection attack injects malicious data $\mathbf{a} = (a_1, a_2, .., a_m)'$ into valid measurements $\mathbf{z}$ in (1). Then, the received measurements are

$$\tilde{\mathbf{z}} = \mathbf{H}\mathbf{x} + \mathbf{e} + \mathbf{a}. \quad (13)$$

From (3), the ML estimate of $\mathbf{x}$ now becomes $\tilde{\mathbf{x}} = \hat{\mathbf{x}} + \mathbf{P}\mathbf{a}$ and the estimate of line flows is $\tilde{\mathbf{f}} = \hat{\mathbf{f}} + \mathbf{Q}\mathbf{P}\mathbf{a}$. Compared with the actual state estimate in (3), the introduced errors are $\tilde{\mathbf{x}} - \hat{\mathbf{x}} = \mathbf{P}\mathbf{a}$. Meanwhile, the residual is

$$\tilde{\mathbf{r}} = \tilde{\mathbf{z}} - \mathbf{H}\tilde{\mathbf{x}} = (\mathbf{I} - \mathbf{H}\mathbf{P})\tilde{\mathbf{z}} \triangleq \mathbf{B}(\mathbf{z} + \mathbf{a}). \quad (14)$$

Load redistribution (LR) attack is first introduced in [8]. The attack vector of a LR attack is $\mathbf{a} = (\mathbf{a_p}; \mathbf{a_d}; \mathbf{a_f})^T$, where $\mathbf{a_p} = \mathbf{0}$ and $\mathbf{e}^T\mathbf{a_d} = 0$ always hold to balance the power generation and demand seen by the ISO. Notice that, LR attacks do not compromise generation measurements. This is because generation measurements are often physically well protected and can be verified through direct communications between ISOs and utility companies. In contrast, line flows and load measurements are widely distributed and fast varying, thus are more vulnerable to attacks.

1) LR attack without resource constraint: We first formulate LR attack to alter the real-time LMP assuming that the attackers have full control over all load and line flow measurements. We denote the set of positively-congested, negatively-congested and un-congested lines derived from Algorithm 1 by $\mathcal{C}_+^{at}, \mathcal{C}_-^{at}$ and $\mathcal{C}_n^{at}$, respectively. The attack objective is to achieve the malicious congestion pattern with minimum errors introduced to the state estimates, i.e. $\mathbf{P}\mathbf{a}$. This is because large deviations from predictions may raise suspicion of the ISO. In addition, the BDD alarm must not be triggered.

Accordingly, LR attack is formulated as follows

$$\underset{\mathbf{a}}{\text{minimize}} \quad ||\mathbf{P}\mathbf{a}||^2 \quad (15a)$$

$$\text{s. t.} \quad (\mathbf{0}, \mathbf{e}, \mathbf{0})^T\mathbf{a} = 0 \quad (15b)$$

$$(\mathbf{I}, \mathbf{0}, \mathbf{0}) \cdot \mathbf{a} = \mathbf{0} \quad (15c)$$

$$||\mathbf{B}(\mathbf{a} + \mathbf{z})||^2 \leq \tau \quad (15d)$$

$$\epsilon_l\hat{\mathbf{d}} \leq \mathbf{P_d}\cdot(\mathbf{a} + \mathbf{z}) \leq \epsilon_u\hat{\mathbf{d}} \quad (15e)$$

$$\mathbf{Q}_{l,\cdot}\mathbf{P}\mathbf{a} \geq f_l^{max} - \hat{f}_l, \quad l \in \mathcal{C}_+^{at} \quad (15f)$$

$$\mathbf{Q}_{l,\cdot}\mathbf{P}\mathbf{a} \leq f_l^{min} - \hat{f}_l, \quad l \in \mathcal{C}_-^{at} \quad (15g)$$

$$f_l^{min} - \hat{f}_l + \delta \leq \mathbf{Q}_{l,\cdot}\mathbf{P}\mathbf{a} \leq f_l^{max} - \hat{f}_l - \delta, \quad l \in \mathcal{C}_n^{at} \quad (15h)$$

[1]The rows of $\mathbf{G}$ correspond to lines $\{l_{12}, l_{14}, l_{15}, l_{23}, l_{26}, l_{36}, l_{45}, l_{56}\}$ and columns correspond to bus 1 to 6.

[2]This is because ex-post calculation does not physically re-dispatch power generation but only sets the real-time LMP based on current estimations.

where $\mathbf{Q}_{l,\cdot}$ is the row in $\mathbf{G}$ that corresponds to line l and δ is a small positive padding parameter. $\mathbf{P_d}$ is the rows that correspond to load estimates in $\mathbf{P}$. (15b) is to satisfy $\mathbf{e}^T\mathbf{a_d} = 0$. (15c) is to guarantee no error is introduced to generation measurements. (15d) is to ensure the alarm is not triggered after attack. (15e) is to guarantee that the load estimates after attack are within reasonable ranges, controlled by parameters ϵ_l and ϵ_u. For simplicity, we can set $\epsilon_l = 0$ and $\epsilon_u = \infty$. (15f)-(15h) are to achieve the derived congestion pattern. (15) is a convex minimization problem with linear constraints and can be solved with great efficiency.

In practice, however, the attack in (15) is costly, mainly because it needs to simultaneously compromise most of load and line flow measurements, and thus is difficult to implement in large scale power networks. Inspired by the insights obtained from Section III.C, we propose in the following a cost-aware "neighborhood" LR (NLR) attack.

2) Resource-constrained NLR attack: We limit the capacity of attackers to only compromise the load measurements of the tagged bus and its one-hop neighboring buses, and line flow measurements within k hops. Here, k is a tunable parameter. For simplicity, a valid attack vector is denoted by $\mathbf{a} \in \mathcal{P}^k$, which satisfies $\mathbf{a_p} = \mathbf{0}$ for generation measurements and $a_l = 0$ for a transmission line l that is more than k hops away from the tagged bus.

Suppose that bus i is the tagged bus. We denote the set of positively congested, negatively congested and un-congested lines within k-hop of the tagged bus by $\mathcal{C}_+^k$, $\mathcal{C}_-^k$ and $\mathcal{C}_n^k$, respectively. Notice that all these congestion observations are actual measurements. The idea of a NLR attack is to create a fake radial power flow pattern between the tagged bus and its one-hop neighboring buses. In particular, it achieves

1) the tagged bus is not the immediate receiving end of any congested lines,
2) the tagged bus is the immediate sending end of at least one congested line.

Mathematically, it is formulated as

$$\underset{\mathbf{a}\in\mathcal{P}^k}{\text{minimize}} \quad ||\mathbf{Pa}||^{\mathbf{2}} \tag{16a}$$

$$\text{s. t.} \quad (\mathbf{0}, \mathbf{e}, \mathbf{0})^T\mathbf{a} = 0 \tag{16b}$$

$$||\mathbf{B}(\mathbf{a}+\mathbf{z})||^2 \le \tau \tag{16c}$$

$$\epsilon_l\hat{\mathbf{d}} \le \mathbf{P_d}\cdot(\mathbf{a}+\mathbf{z}) \le \epsilon_u\hat{\mathbf{d}} \tag{16d}$$

$$\mathbf{Q}_{l,\cdot}\mathbf{Pa} \le f_l^{max} - \hat{f}_l - \delta, \ \ \forall l \in \mathcal{C}_+^k \text{ and } G_{l,i} \le 0 \tag{16e}$$

$$\mathbf{Q}_{l,\cdot}\mathbf{Pa} \ge f_l^{min} - \hat{f}_l + \delta, \ \ \forall l \in \mathcal{C}_-^k \text{ and } G_{l,i} \ge 0 \tag{16f}$$

$$\mathbf{Q}_{l,\cdot}\mathbf{Pa} \le f_l^{min} - \hat{f}_l, \ \ \forall l \in \mathcal{C}_n^{at} \text{ and } G_{l,i} \le 0 \tag{16g}$$

$$\mathbf{Q}_{l,\cdot}\mathbf{Pa} \ge f_l^{max} - \hat{f}_l, \ \ \forall l \in \mathcal{C}_n^{at} \text{ and } G_{l,i} \ge 0. \tag{16h}$$

Constraints (16e) and (16f) are to decongest the transmission lines that the tagged bus is the immediate receiving end. (16g) and (16h) are to create new congestion that the tagged bus is the immediate sending end. Here, $\mathcal{C}_n^{at} \subseteq \mathcal{C}_n^1$ is a set of transmission lines within one-hop of the tagged bus, chosen to be congested after attack. Since the size of $\mathcal{C}_n^1$ is generally very small, it is affordable to try out all $\mathcal{C}_n^{at} \subseteq \mathcal{C}_n^1$ and select the one with minimum objective value.

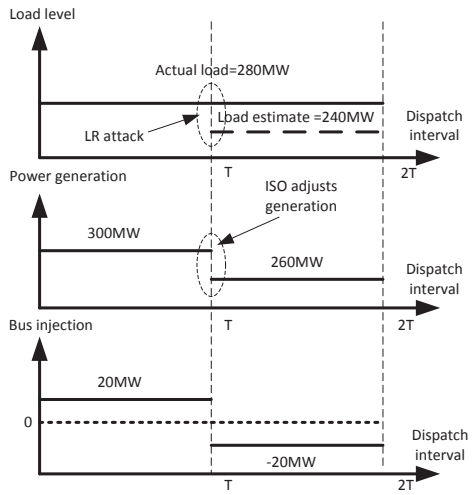

Fig. 2. Illustration of the impact of LR attack to future market operation.

A point to notice is that both the attacks in (15) and (16) could be infeasible. One cause of infeasibility is the random measurement error. Another cause is that LR attack is incapable of achieve arbitrary congestion pattern since it assumes no capability to compromise generation measurements. Currently, a straightforward solution is to re-run the algorithm until a feasible congestion pattern is found. Notice that, when we derive the malicious congestion pattern, the primary concern is its effectiveness of reducing the electricity price rather than implementation details. In future work, the congestion patter derivation and attack implementations may need to be jointly considered.

B. Impacts of LR attack to future market

Interestingly, we find that LR attack to real-time LMP can also affect future power market. This is because the current false load estimation can induce the ISO to produce biased load prediction in the next time slot, and thus issue wrong power generation dispatch. As a result, severe congestions could occur even if all entities follow dispatch orders and no attack is present. As a motivating example, the impact of LR attack to the future LMPs is illustrated in Fig. 2.

Fig. 2 shows the power generation, load level and power injection over two consecutive dispatch intervals at bus 3 in Fig. 1. The attackers launch LR attack at T. For simplicity, we assume that the ISO uses current load estimates as the load prediction for ex-ante calculation in the next time slot. When no attack is present, we assume that the prediction is accurate. Besides, all generators follow the ex-ante generation dispatch and the actual load consumptions do not change over the two dispatch intervals. At time epoch 0, the ISO dispatches generation order based on load prediction with $d_3 = 280MW$. The power generation at bus 3 is $300MW$, and thus the power injection is $20MW$. At time epoch T, attackers inject false data to reduce the electricity price LMP_3, causing the ISO to produce an biased estimate of $d_2' = 240MW$. The ISO foresees the decreasing load at bus 3, thus reacting by adjusting the power generation at bus 3 from $300MW$ to $260MW$ to

match the change of load in the next interval. However, the actual load does not change in the second dispatch interval and the actual power injection becomes $260 - 280 = -20MW$. From $20MW$ to $-20MW$, the sudden change of power injection may induce severe congestion at bus 3. As a result, even if no attack is present at $2T$, the ex-post LMPs at $2T$ could deviate significantly from the ex-ante LMP at T due to unexpected congestions. Specifically, the LMP at bus 3 tends to increase while LMPs at its neighboring buses tend to decrease. This conveys a potentially profitable market signal to the attackers if they are allowed to hold proper financial derivatives. Due to the page limit, we do not extend our discussions on attacking the future LMP in this paper.

V. Case Study

Using the power system in Fig. 1, we demonstrate the impacts of the LR attack to LMPs. We assume that the ISO's load prediction is perfect so that real-time LMP is consistent with ex-ante LMP if no attack is present. The actual congestion pattern in the network is $\mathcal{C}_+ = \{l_{12}, l_{14}\}$ and $\mathcal{C}_- = \{l_{56}\}$. The true real-time LMPs are $[20, 25, 25, 35, 28.7, 24]$ \$/MWh when no attack is present. Since the real-time LMP at bus 3 and 5 are higher than the generation costs ($c_3 = 22$ and $c_5 = 23$), they are considered as target buses for performing the proposed attack. Here, we assume $\triangle p^{max} = 0.1MW$ and $\triangle p^{min} = -2MW$ in (6). Using bus 2 and 3 as example, the attacking strategies of NLR attack to the two buses are

- attack bus 3: congest l_{23}, i.e. the congestion pattern is $\mathcal{C}_+^{at} = \{l_{12}, l_{14}\}$ and $\mathcal{C}_-^{at} = \{l_{23}, l_{56}\}$.
- attack bus 5: decongest l_{56} and congest l_{45}, i.e. the congestion pattern is $\mathcal{C}_+^{at} = \{l_{12}, l_{14}\}$ and $\mathcal{C}_-^{at} = \{l_{45}\}$.

The LMP variations at all 6 buses are plotted in Fig. $3(a)$. After the two attacks, LMPs at bus 3 and 5 are reduced by 3 \$/MWh and 5.7 \$/MWh, respectively.

Besides, we plot in Fig. $3(b)$ the LMPs of bus 3 and 5 against the time of dispatch intervals. If no attack is present, LMPs at two buses should be invariant across the two dispatch intervals. We assume that attackers use NLR attack to reduce LMP_5 at T. This is achieved by introducing errors to decrease the reading of d_5 and increase the readings of d_4 and d_6. Meanwhile, LMP_3 is increased as well. However, this only causes temporary price deviations from normal values. Once attackers stop injecting false data, electricity prices return to normal at $2T$. Nonetheless, this still indicates that attackers can manually create predictable market disturbance thus has the potential to make profit from both real-time and future markets.

VI. Conclusions

In this paper, we have provided detailed analysis and formulation of false-data injection attacks to control real-time electricity price in power market. A simple algorithm is proposed to find the effective congestion pattern to induce false LMP. We realized the obtained congestion pattern using load redistribution (LR) attack, where formulations of both resource-unconstrained attacks and resource constrained

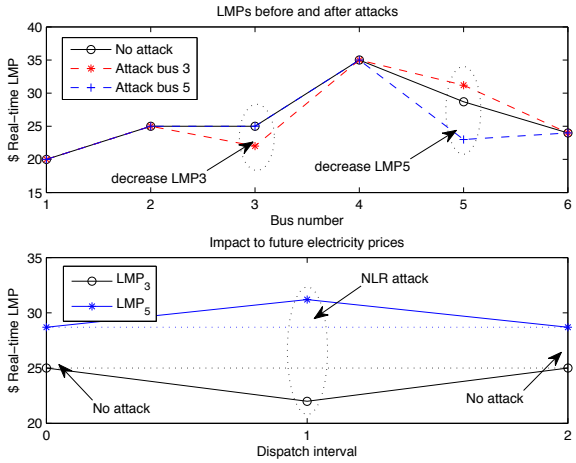

Fig. 3. Impacts of LR attack to real-time and future LMPs.

neighborhood LR attack have been derived. Interestingly, we find that LR attack which controls the real-time LMP also has impact on the future market.

References

[1] A. Ott, "Experience with PJM market operation, system design, and implementation", *IEEE Transcations on power systems*, Vol. 18, No.2, 528-534, May 2003.

[2] T. Zheng and E. Litvinov, "Ex Post Pricing in the co-optimized enegy and reserve market", *IEEE Transactions on power systems*, vol. 21, No. 4, 1528-1538, Nov. 2006.

[3] Y. Liu, P. Ning, and M. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proceedings of the 16th ACM conference on Computer and communications security*, Chicago, Illinois, 2009, pp. 21-32.

[4] O. Kosut, L. Jia, R. J. Thoma and L. Tong, "Malicious data attacks on the smart grid", *IEEE Trans. on smart grid*, Vol.2, no.4, pp. 645-658, Dec 2011.

[5] G. Dan and H. Sandberg, "Stealthy Attacks and Protection Schemes for State Estimators in Power Systems," in *SmartGridComms 2010*, pp. 214-219. 2010.

[6] S. Bi, Y. Zhang, "Defending mechanisms against false-data injection attacks in the power system state estimation," *IEEE SG-ComNeTs Workshop*, Houston, USA, 2011.

[7] R. Bobba, K. M. Rogers, Q. Wang, H. Khurana, K. Nahrstedt, and T. Overbye, "Detecting false data injection attacks on DC state estimation," in *Preprints of the First Workshop on Secure Control Systems, CPSWEEK 2010,* Apr. 2010.

[8] Y. Yuan, Z. Li, K. Ren, "Modeling Load Redistribution Attacks in Power Systems," *IEEE Transactions on Smart Grid*, vol.2, no. 2, pp. 382-390, June 2011.

[9] L. Xie, Y. Mo, and B. Sinopoli, "Integrity data attacks in power market operations," *IEEE Trans. on smart grid*, vol.2, no.4, pp. 659-665, Dec 2011.

[10] L. Jia, R. J. Thomas, and L. Tong, "Impacts of Malicious Data on Real-time Price of Electricity Market Operations", Hawaii Intl. Conf. on System Sciences, Jan 2012.

[11] F. Li, Y. Wei and S. Adhikari, "Improving an unjustified common practice in ex-post LMP calculation", *IEEE Transactions on power systems*, vol. 25, No. 2, 1195-1197, Nov. 2010.