Construct Data Integrity Attacks Against Real-Time Electrical Market in Smart Grid

Song Tan*, Wen-Zhan Song*, Michael Stewart†, and Lang Tong‡
Department of Computer Science, Georgia State University, GA, USA
School of Electrical and Computer Engineering, Cornell University, NY, USA
stan.songwz@cs.gsu.edu, mastewart@gsu.edu, ltong@ece.cornell.edu

Abstract—The normal operation of real-time electrical market requires accurate state estimation in Smart Grid. However, recent research shows that strategically designed data integrity attacks can easily introduce errors into state estimation without being detected. From the perspective of attackers, this increases the chances of controlling real-time electrical market operations through manipulations of meter measurements. In this paper, we first reveal the intrinsic relations between data integrity attacks and real-time electrical market, and explicitly characterize their complex interactions as a process simulator. Then a simulation-based global optimization problem is formulated, from which an attacker could maximize financial incentives through constructed data integrity attacks. More importantly, we further consider the construction of data integrity attacks when power network topologies and parameters are unknown. A systematic construction strategy is proposed based on recursive least square subspace estimation. As far as we know, our paper is the first attempt to attack real-time electrical market without network information. Finally, we evaluate the performance of the proposed attacking strategies through numerical simulations in IEEE test systems.

I. INTRODUCTION

As the deregulation of electric power industry, real-time electrical market mechanisms are widely adopted by major U.S. independent system operators (ISOs) and play key roles in balancing supply and load, clearing market prices, and maintaining grid stability [1]. The operations of real-time electrical market are built upon the data and results from state estimation process. However, recent research shows that strategically designed data integrity attacks can easily introduce errors into state estimation without being detected [2]. As a result, it becomes more and more likely for attackers to control real-time electrical markets through manipulations of meter measurements. Therefore, assessing the feasibility and impact of data integrity attacks against real-time electrical market is of paramount importance to cyber security in Smart Grid.

Quite a few of existing works have addressed this issue. In [3], Xie et al. firstly investigate the impact of integrity attacks on power market through virtual bidding. In [4], Kosut et al. evaluate the proposed data attacks by their generated market revenues and the work is further studied by Jia et al. in pursuit of maximizing the revenues [5]. Meanwhile, with the objective of controlling real-time Locational Marginal Prices (LMP) directly through data attacks, Tan et al. in [6] employ a control theory based approach to analyze the attack effect on pricing stability. More recently, the authors in [7] and [8] have respectively proposed formal analytic frameworks to further quantify the impact of data qualities on real-time LMP. However, all the above related works are based on the assumption that the attacker has full knowledge about the network information of targeted power systems, which includes network topologies and branch parameters, etc. In fact, in any given power system, the network information is huge and highly secured, and more importantly, these information are dynamic since the network topology could be reconfigured in both normal situations and contingencies. Therefore, from an attacker’s perspective, it is rather difficult to achieve complete awareness of network information in practice.

In this paper, we first reveal the intrinsic relations between data integrity attacks and real-time electrical market, and explicitly characterize their complex interactions as a process simulator. Then a simulation-based global optimization problem is formulated, from which an attacker could maximize financial incentives through constructed data integrity attacks. More importantly, we further consider the construction of data integrity attacks when the power network topology and parameters are unknown. A systematic construction strategy is proposed based on recursive least square subspace estimation. As far as we know, our paper is the first attempt to attack real-time electrical market without network information.

II. PRELIMINARIES AND SYSTEM MODEL

A. State Estimation and Bad Data Detection

In state estimation process, the control center collects real time measurements $z$ from the deployed sensors and combines the network topology and parameter information to calculate the real time estimates of the unknown system variables $x$. Mathematically [9], let $x = (x_1, x_2, ..., x_n)^T$ and $z = (z_1, z_2, ..., z_m)^T$ denote state variables and meter measurements, respectively, where $n$ is the number of unknown state variables, $m$ is the number of meters, and $m \geq n$. The state variables are related to the measurements by $z = Hx + e$, where $e$ is the Gaussian measurement noise with zero mean and a covariance matrix $\sigma^2 I$. Under DC power flow model [9], the measurement model can be represented as:

$$z = Hx + e$$

(1)
where \( z \) is the bus power injection (power generation or load) and branch power flow measurements, \( H \) is an \( m \times n \) full rank Jacobian matrix of the measurement model and \( x \) is the voltage phases at all buses. Then the estimated system states \( \hat{x} \) and branch power flows \( \hat{f} \) are given by:

\[
\hat{x} = (H^T H)^{-1} H^T z, \quad \hat{f} = F \hat{x}
\]

(2)

where \( F \) is the sensitivity matrix of branch flows with respect to the voltage phases. With DC power flow model, since there also exists a linear bijection between nodal power injections and voltage phases [10], then given a reference bus, we would have a \( l \)-by-\( n \) injection shift factor matrix \( S \) to stabilize the power system and calculate Locational Marginal Prices (LMP) [1] is solved to find the associated real-time LMP at bus \( i \). From (2),

\[
\hat{z} = H \hat{x} = K z, \quad \text{where} \quad K = H (H^T H)^{-1} H^T
\]

(4)

Then the measurement residual can be written as:

\[
r = z - \hat{z} = (I - K) z
\]

(5)

The detector fires an alarm when \( \| r \|_2 > \text{threshold} \).

**B. Real-Time Electrical Market**

A combined two-stage (day-ahead and real-time) market is widely adopted by major U.S. Independent System Operators (ISO) to stabilize the power system and calculate Locational Marginal Prices (LMP) [1]. In the day-ahead market, given the projected system load levels \( L \), the ISO obtains the optimal generation dispatch \( P^* \), the vector of predicted power generation at each bus. Then \( P^* \) are sent to all generators as generation reference, and day-ahead payments are collected from customers at all buses.

In the real-time stage, the ISO obtains the actual system response through state estimation, including the estimated power injections \( \hat{P}, \hat{L} \) and branch flows \( \hat{f} \). Then the following linear program [1] is solved to find the associated real-time LMP \( \lambda \), a vector whose \( i \)th element \( \lambda_i \) is the LMP at bus \( i \):

\[
\begin{align*}
\text{minimize} & \quad \sum_{i} C^G_i \Delta P_i - \sum_{j} C^L_j \Delta L_j \\
\text{s.t.} & \quad \sum_{i} \Delta P_i = \sum_{j} \Delta L_j \\
& \quad \Delta P_i^\text{min} \leq \Delta P_i \leq \Delta P_i^\text{max} \\
& \quad (\mu_b) : \sum_{i} S_{bi} \Delta P_i - \sum_{j} S_{bj} \Delta L_j \leq 0, \text{for } b \in \hat{C}
\end{align*}
\]

(6)

where \( \Delta P \) and \( \Delta L \) are the vectors of incremental generation dispatch and load dispatch at buses, with fixed cost \( C^G, C^L \) respectively. \( \Delta P_i^\text{min}, \Delta P_i^\text{max} \) are predefined lower and upper bounds, usually chosen as -2MW and 0.1MW in practice [1]. \( S_{bi} \) is element at \( b \)th row, \( i \)th column of matrix \( S \) in (3). Note that \( \hat{C} \) is called congestion pattern [8], which denote the sets of branches whose estimated power flow exceeds the flow limit \( f_{b}^\text{max} \).

\[
\hat{C} = \{ b : \hat{f}_b > f_{b}^\text{max} \}
\]

Then by solving (6), the real-time LMP at bus \( i \) is calculated:

\[
\lambda_i = \tau - \sum_{b \in \hat{C}} S_{bi} \mu_b
\]

(8)

where \( \tau, \mu_b \) are the corresponding dual variables in (6).

To clear the real-time market, the generator at bus \( i \) receives revenue \( \lambda_i (P_i - P_i^*) \), and the customer at bus \( j \) pays \( \lambda_j (L_j - L_i) \), where \( P_i \) and \( L_j \) are the estimated power generation and load at bus \( i \) and \( j \) from state estimation, respectively [1].

**III. PROBLEM FORMULATION**

Suppose a malicious party wants to generate revenues from the real-time electrical market by compromising a subset of meters \( \zeta_A \), such that only measurements from meters in \( \zeta_A \) can be modified. Note that the following strategies can also be applied to reducing customers’ payments.

**A. Constraints of Attacks**

Firstly, since the attacker can only modify the measurements from meters in \( \zeta_A \), then the perturbed measurements has to be in the form:

\[
z_a = z + a, \quad a \in \{ a \in \mathbb{R}^m | a = \Psi c, \forall c \in \mathbb{R}^m \}
\]

(9)

where \( a \) is the attack vector, and \( \Psi \) is the diagonal matrix:

\[
\Psi = \text{diag}(\psi_1, \ldots, \psi_m)
\]

(10)

where \( \psi_i \) is a binary variable and \( \psi_i = 1 \) iff meter \( i \in \zeta_A \).

Secondly, the attack should not be detected by the bad data detector. Based on (5), the new residual becomes \( \tilde{r} = r + (I - K) a \). Based on triangular inequality,

\[
\| \tilde{r} \|_2 \leq \| r \|_2 + \| (I - K) a \|_2
\]

(11)

Here we introduce a parameter \( \varepsilon \), such that \( \| (I - K) a \|_2^2 \leq \varepsilon \). The smaller \( \varepsilon \) is chosen, the less likely the attack will be detected. In the extreme case when \( \varepsilon = 0 \), the attack becomes unobservable [4].

**B. Objective of Attacks**

The objective of the attack is to maximize revenues from the real-time electrical market. From the end of Section II (B), we can see that the generator at bus \( i \) receives revenue \( \lambda_i (P_i - P_i^*) \) in normal situation. We analyze \( \lambda_i \) and \( \tilde{P}_i - P_i^* \) separately.

First, from (6) and (8), it suggests that given a shift factor matrix \( S \), the real-time LMP \( \lambda \) depends only on the ISO’s congestion pattern observation [8], i.e., \( \hat{C} \). Meanwhile, since the ISO determines \( C \) through the estimated branch flows \( \hat{f} \) as in (7), and \( \hat{f} \) are solely determined by the power injection measurements within \( z_a \) as in (3), therefore, we can see that \( z_a, \hat{f}, \hat{C} \) and real-time LMP \( \lambda \) form a Markov chain, such that given a tuple of \( (a, z, S) \), there is a single corresponding \( \lambda \). In other words, the LMP \( \lambda \) is essentially a function of \( (a, z, S) \).
So from now on, we denote LMP as \( \lambda(a, z, S) \). We abstract the complex routine of \( \lambda(a, z, S) \) as a simulator, and the flow chart of the simulator is shown in Figure 1.

![Locational Marginal Price Simulator](image)

**Fig. 1.** Locational Marginal Price Simulator \( \lambda(a, z, S) \)

Second, since the estimated power generation in real-time stage should match the optimal dispatch in day-ahead stage under normal situations [4] [5], then according to equation (4), for each bus \( i \),

\[
\tilde{P}_i - P_i^* = K_i(z + a) - P_i^* = K_i a
\]

where \( K_i \) is the corresponding row in matrix \( K \) to generation at bus \( i \). Therefore, assume the attacker wants to make revenues from generations at buses within a target set \( B \). Then the total revenue from attack vector \( a \) is:

\[
\mathcal{V}(a) = \sum_{i \in B} \lambda_i(a, z, S) K_i a
\]

**C. Construct Attacks Against Real-Time Electrical Market**

From all the above, the problem of constructing data integrity attacks against real-time electrical market can be formulated as a simulation-based global optimization problem P1:

\[
\text{(P1)}: \quad \max_a \mathcal{V}(a) = \sum_{i \in B} \lambda_i(a, z, S) K_i a
\]

s.t. \( a = \Psi c, \quad \forall c \in R^m \)

\[
\| (I - K) a \|_2^2 \leq \varepsilon
\]

**D. Remarks**

To solve P1, several facts are worth pointing out.

First, the objective function of P1 is based on complex simulator process in Figure 1. Only function values are available and there is no algebraic model to analyze differentiability and Lipschitz condition. Therefore, derivative-free optimization methods have to be employed [12]. The selection of optimization solvers is out of the scope of this paper.

Second, as in (3) (4) (6) (11) (13), the attacker will need the following knowledge to solve P1 accurately:

1) The meter measurement \( z \).

2) The elements of matrix \( H, S \) and \( K \).

The elements of matrix \( H, S \) and \( K \) depends on the detailed knowledge of network information, including network topology and branch parameters, such as the exact position of circuit breaker switches, transformer tap changers and power line admitances, etc. In fact, in any given power network, the network information is huge and highly secured, and these information could be dynamic since the topology can be reconfigured in both normal situations and contingencies. Therefore, it is rather difficult for the attackers to achieve complete awareness of network information in practice. A good question would be whether it is possible to construct the attacks when the network information is unknown. In the following section, we give a firm answer to this question, which is yes.

**IV. Construct Data Integrity Attacks without Network Information**

In this section, we consider the strategy to construct data integrity attacks without network information, in other words, when matrix \( H, S \) and \( K \) are unknown to the attacker. The measurement \( z \), the cost coefficient \( C^G, C^L \), and flow limit \( f_{max} \) are known to the attacker. By examining P1, we can see that the attacker need to specifically deal with \( \| (I - K) a \|_2^2 \leq \varepsilon \) in the constraint, and \( \lambda_i(a, z, S) \) and \( K_i a \) in the objective.

**A. Constraint \( \| (I - K) a \|_2^2 \leq \varepsilon \)**

The constraint \( \| (I - K) a \|_2^2 \leq \varepsilon \) determines whether or not the attack can be detected by the bad data detector. Since \( K \) is unknown, to safely launch an attack, the attacker would need to consider the extreme case of constraint \( \| (I - K) a \|_2^2 \leq \varepsilon \), which is when \( \varepsilon = 0 \). In other words, if we could construct a vector \( a \) which has \( \| (I - K) a \|_2^2 = 0 \), then such an \( a \) would always satisfy the constraint and the bad data detector can never detect it. Note from (4) that when \( a = H c, \forall v \in R^n \), we always have \( \| (I - K) a \|_2^2 = 0 \). This is the so called unobservable attack [4]. Therefore, to satisfy the constraint, we just need to construct a vector \( a \), which always lies in \( \mathbb{R}(H) \), the column space of matrix \( H \). The last question would be how to determine \( \mathbb{R}(H) \) when \( H \) is unknown and even dynamic?

Inspired by [13], we can directly estimate and track the subspace \( \mathbb{R}(H) \) using the measurement \( z \). Let \( z_t \) denote the measurement vector at each time \( t \), from (1):

\[
z_t = H x_t + e_t
\]

To estimate \( \mathbb{R}(H) \), at each time \( t \), we aim at minimizing the following loss function,

\[
J_t = \arg \min_{J \in R^{m \times n}} \sum_{i=1}^{t} \rho^{t-i} u_i(J),
\]

where forgetting factor \( 0 \ll \rho \leq 1 \) controls the memory and tracking ability, \( J \) is the estimated subspace with rank \( n \) since \( H \) is always full rank, and

\[
u_i(J) = \min_x \| (z_i - J x) \|_2^2, \quad i = 1, ..., t
\]

To solve (18), we alternate between the coefficient estimation and subspace update at each time \( t \). Specifically, the coefficient vector is estimated by:

\[
x_t = \arg \min_x \| (z_t - J_t x) \|_2^2
\]

\[
= (J_{t-1}^T J_{t-1})^{-1} J_{t-1}^T z_t
\]
where \( J_0 \) is a random initialization. Then \( J_t \) is solved from:

\[
J_t = \arg \min_J \sum_{i=1}^{t} \rho^{t-i} \| (z_i - J x_i) \|_2^2
\]

(21)

where \( x_i, i = 1, ..., t \), are estimated from (20).

Note for all row \( h = 1, 2, ..., m \), the objective function in (21) can be rowwise decomposed [14] as \( J_t = [J_1^T, J_2^T, ..., J_m]^T \):

\[
J_t^h = \arg \min_{\alpha} \sum_{i=1}^{t} \rho^{t-i} (z_i(h) - x_i^T J_h)^2
\]

(22)

\[
= J_t^{h-1} + (z_t(h) - x_t^T J_t^{h-1})(W^t)^{T} x_t
\]

where \( W^t = \rho W^{t-1} + x_t x_t^T \) and \( \dagger \) means pseudoinverse. Equation (22) is the classical formulation of Recursive Least Square (RLS) estimation with forgetting [15]. Based on RLS updating formula, we further have:

\[
(W^t)^{\dagger} = \rho^{-1}(W^{t-1})^{\dagger} - (\rho^{-1} - \alpha_t^{(t)})^{\dagger}
\]

(23)

\[
\beta_t = 1 + \rho^{-1} x_t(W^{t-1})^{\dagger} x_t, \quad \alpha_t = \rho^{-1}(W^{t-1})^{\dagger} x_t
\]

(24)

We summarize the subspace estimation and tracking process for \( \mathbb{R}(H) \) in Algorithm 1. At each time \( t \), the attacker acquires the current estimated subspace \( J_t \) from Algorithm 1. Then the most conservative approach to replace the constraint \( \| (I - K) a \|_2^2 \leq \varepsilon \) is:

\[
a = J_t \cdot \eta, \quad \forall \eta \in \mathbb{R}^n
\]

(25)

Note that this constraint can be further relaxed in section D.

### B. Objective \( \lambda(a, z, S) \)

To calculate \( \lambda(a, z, S) \), the attacker should figure out the unknown matrix \( S \). Assume \( z \) contains all the branch flow measurements and power injection measurements at all buses. Based on (3), let \( l \) denote the number of branches, for a particular branch flow measurement \( z_f(j) \), \( j = 1, 2, ..., l \) in \( z_f \), at each time \( t \), we have:

\[
z_f^t(j) = (z_i^m)^T S_j + e_t
\]

(26)

where \( S_j^t \) is the \( j \)th row of matrix \( S \) at time \( t \). Therefore, we can also estimate \( S_j^t \) through RLS with forgetting:

\[
S_j^t = \arg \min_{S_j^t} \sum_{i=1}^{t} \rho^{t-i}((z_i^m)^T S_j^t)^2
\]

(27)

\[
= S_j^{t-1} + ((z_i^m)^T S_j^{t-1})(W^t)^{T} z_i^m
\]

where \( W^t = \rho W^{t-1} + z_i^m (z_i^m)^T \). Similar routines as in (23) (24) can be applied to find \( S_j^t \). The process of estimating \( S \) is summarized in Algorithm 2. \( S_0 \) is initialized as a random \( l \times n \) matrix.

Therefore, to calculate \( \lambda(a, z, S) \), at each time \( t \), the attacker first get \( S_t \) from Algorithm 2, then invoke simulator \( \lambda(a, z, S_t) \) as in Figure 1.

### Algorithm 2 Estimation for Shift Factor Matrix \( S \)

1: **Input:** A sequence of real-time measurements \( z_t, t = 1, 2, ..., \)
2: **Initialize:** An \( n \times n \) row vector \( J_0 \), and a diagonal matrix \( (W^0)^{\dagger} = \delta I, \delta \gg 0 \)
3: for \( t=1,2,... \) do
4: \( \beta_t = 1 + \rho^{-1} x_t(W^{t-1})^{\dagger} x_t, \quad \alpha_t = \rho^{-1}(W^{t-1})^{\dagger} x_t \)
5: for \( h=1,2,...,m \), in parallel do
6: \( J_h^t = J_h^{t-1} + (z_t(h) - x_t^T J_h^{t-1})(W^t)^{T} x_t \)
7: end for
8: Form \( S_t = [S_1^t, S_2^t, ..., S_l^t]^T \);
9: end for

C. Objective \( K_{\alpha} \)

From (4), when \( H \) is unknown, \( K \) is unknown, so we cannot calculate \( K_{\alpha} \) directly. However, the following Lemma 1 sheds some light on the method to tackle this problem.

**Lemma 1.** \( K \) in (4) is an orthogonal projector onto \( \mathbb{R}(H) \).

**Proof:** Suppose \( b \in \mathbb{R}^n \) and let \( b = H x \) be the orthogonal projection of \( b \) onto \( \mathbb{R}(H) \). Then the residual \( r = b - H x \) is orthogonal to \( \mathbb{R}(H) \), hence, it is orthogonal to each of the columns of \( H \). As a result, we have:

\[
H^T (b - H x) = 0 \implies H^T H x = H^T b \implies x = (H^T H)^{-1} H^T b \implies \hat{x} = H (H^T H)^{-1} H^T b
\]

(28)

\[
\implies b = H (H^T H)^{-1} H^T b = K b
\]

Therefore, \( K \) is an orthogonal projector onto \( \mathbb{R}(H) \).

**Lemma 1.** \( K \) in (4) is an orthogonal projector onto \( \mathbb{R}(H) \).

**Proof:** Suppose \( b \in \mathbb{R}^n \) and let \( b = H x \) be the orthogonal projection of \( b \) onto \( \mathbb{R}(H) \). Then the residual \( r = b - H x \) is orthogonal to \( \mathbb{R}(H) \), hence, it is orthogonal to each of the columns of \( H \). As a result, we have:

\[
H^T (b - H x) = 0 \implies H^T H x = H^T b \implies x = (H^T H)^{-1} H^T b \implies b = H (H^T H)^{-1} H^T b = K b
\]

Therefore, \( K \) is an orthogonal projector onto \( \mathbb{R}(H) \).

**Based on Lemma 1, we present the following theorem to calculate \( K_{\alpha} \) when \( K \) is unknown.**

**Theorem 1.** Let matrix \( K = (H^T H)^{-1} H^T b \), where \( H \in \mathbb{R}^{m \times n} \) with full rank \( n \). Suppose there is another matrix \( J \in \mathbb{R}^{m \times n} \) also with full rank \( n \), and \( \mathbb{R}(J) = \mathbb{R}(H) \). Define matrix \( K' \) as:

\[
K' = J (J^T J)^{-1} J^T
\]

(29)

then \( K = K' \).
Proof: Since $\mathbb{R}(J) = \mathbb{R}(H)$, then:
\[
\forall x \in \mathbb{R}^n, \exists y \in \mathbb{R}^n, \quad \text{s.t.} \quad Hx = Jy. \tag{30}
\]
Based on Lemma 1, matrix $K'$ is also an orthogonal projector onto $\mathbb{R}(H)$. Therefore, for any $u \in \mathbb{R}^n$, we can have:
\[
u = Ku + r, \quad u = K' u + r' \tag{31}
\]
where $Ku, K'u \in \mathbb{R}(H)$, and residual $r, r'$ are orthogonal to $\mathbb{R}(H)$. So,
\[
(r - r')^T(Ku - K'u) = 0 \tag{32}
\]
Since $r = u - Ku$, $r' = u - K'u$, we further have:
\[
(K'u - Ku)^T(Ku - K'u) = 0 \tag{33}
\]
which means for any $u \in \mathbb{R}^n$, we have $Ku = K'u$. Then the orthogonal projector must be unique and $K = K'$.

Therefore, at each time $t$, the attacker can construct $K' = J_t(J_t^T J_t)^{-1}J_t^T$ using $J_t$ generated from Algorithm 1, then use $K' a$ to replace $K a$ in the objective.

D. Summary

Based on Theorem 1, instead of using equation (25), we can further replace constraint $\| (I - K)a \|_2^2 \leq \varepsilon$ with $\| I - K' \|_2^2 \leq \varepsilon$. Therefore, the problem of attack construction without network information is formulated as P2:

\[
\begin{align*}
\text{max} \quad & \mathcal{V}(a) = \sum_{i \in \mathbb{B}} \lambda_i(a, z, S)K'_i a_i \\
\text{s.t.} \quad & a = \Psi c, \quad \forall c \in \mathbb{R}^m \\
& \| I - K' \|_2^2 \leq \varepsilon
\end{align*} \tag{34, 35, 36}
\]

The attack construction process at each time $t$ is summarized in Algorithm 3, where $\nu$ is a small constant.

**Algorithm 3 Construct Data Integrity Attacks Against Real-Time Electrical Market without Network Information**

1. **Input:** A sequence of real-time measurements $z_t, t = 1, 2, \ldots$.
2. **Initialize:** Launch Algorithm 1 and Algorithm 2;
3. for $t=1,2,\ldots$ do
4. Get $J_t$ from Algorithm 1 and $S_t$ from Algorithm 2;
5. if $\| J_t - J_{t-1} \|_F \leq \nu$ and $\| S_t - S_{t-1} \|_F \leq \nu$ then
6. Construct $K' = J_t(J_t^T J_t)^{-1}J_t^T$, update P2 objective;
7. Update the constraint $\| I - K' \|_2^2 \leq \varepsilon$ in P2;
8. Solving P2 using derivative-free optimization method. (In search process, evaluations of objective function invoke simulator routine $\lambda(a, z_t, S_t)$);
9. Based on solved vector $a$, modify the measurements of corresponding meters in $\zeta_A$;
10. end if
11. end for

\[ \frac{\text{LMP at buses in different congestion patterns}}{\text{(a)}} \quad \text{Maximum revenues with different numbers of compromised meters} \quad \text{(b)} \]

**Fig. 2. Data Integrity Attacks in IEEE 14 Bus ystem**

V. EVALUATION

In this section, we evaluate our proposed attacking strategies through IEEE bus benchmark system [16]. All the numerical simulations are conducted in Matlab platform with software packages including @MATPOWER and patternsearch solver in Global Optimization Toolbox.

A. Network Information is Known

In this part, we examine the performance of P1 through IEEE 14 bus system (14 buses and 20 branches), in which the network information is known. All power injection measurements and power flow measurements (in both directions for each line) are employed, such that $m = 54$, $n = 14$. We first examine the functionality of LMP simulator. Since the LMP at all buses totally depend on the congestion pattern, we directly plot the LMP under different congestion patterns in Figure 2a. The congestion pattern includes the ID of branches whose power flows exceed the security limits. One interesting fact is that different congestion patterns could result in the same LMP at a particular bus, e.g., the LMP at bus 4 are the same in the last two congestion patterns, both of which have branch 7, 9 congested, and are incident with bus 4.

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>OPTIMAL ATTACK VECTOR $a$ AGAINST IEEE14 WITH DIFFERENT SIZES OF $\zeta_A$</th>
</tr>
</thead>
<tbody>
<tr>
<td>size of $\zeta_A$</td>
<td>optimal attack vector $a$</td>
</tr>
<tr>
<td>2</td>
<td>(0.63,0.2,2.34)</td>
</tr>
<tr>
<td>4</td>
<td>(0.65,1.2,3.2,0.1,14.48,53,34.64)</td>
</tr>
<tr>
<td>6</td>
<td>(0.69,3.1,2.32,0.1,13.48,0,15.32,0,34.64)</td>
</tr>
<tr>
<td>8</td>
<td>(0.79,3.2,2.32,0.1,13.48,0,14.48,0,15.32,0,34.64)</td>
</tr>
<tr>
<td>10</td>
<td>(0.1,0.4,3.2,2.32,0.1,13.48,0,14.48,0,15.32,0,34.64)</td>
</tr>
</tbody>
</table>

Then we demonstrate the optimal attack vectors in P1 when different number of meters are compromised. Table I lists the optimal attack vector $a$ against IEEE14 system under different size of $\zeta_A$. The notation $(p, q)$ denotes the nonzero entries of vector $a$, and $p$ is the index and $q$ is the value. The corresponding maximum revenues under optimal attack vectors are given in Figure 2b. We can see that the number of compromised meters has a significant impact on the revenues.

B. Network Information is Unknown

In this part, we examine the performance of P2, in which the network information is unknown. Both IEEE14 and IEEE118
In this paper, we present the construction strategies of data integrity attacks against real-time electrical market, with or without network information. Simulation-based global optimization problems are formulated and our results show that the attacker could generate a fair amount of revenues through data integrity attacks in both cases. In future work, countermeasure based on the sparsity feature of attack vector will be investigated to mitigate the financial risks.

**REFERENCES**


