A Framework for Analyzing the Impact of Data Integrity/Quality on Electricity Market Operations

Dae-Hyun Choi

(Advisor: Prof. Le Xie)

Department of Electrical and Computer Engineering
Texas A & M University

February 4, 2014

Smart Grid: A Cyber-Physical System

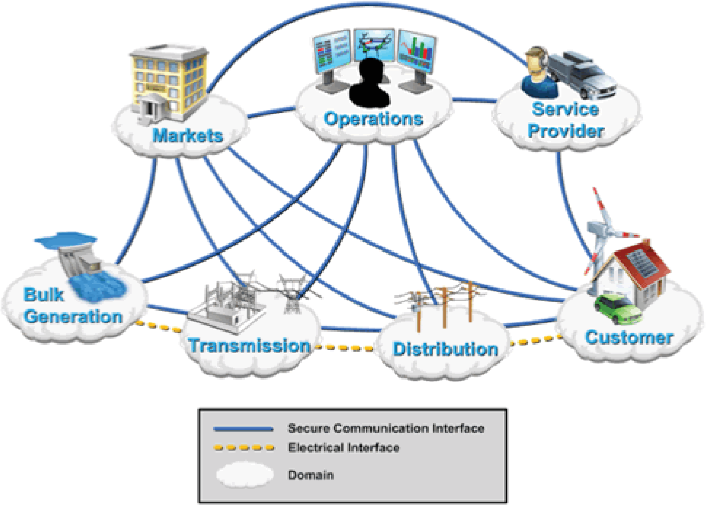


Image Source: http://www.sgiclearinghouse.org/sites/default/files/images/nist/Slide1-1.png

Advanced Grid Sensors Improve Smart Grid Operations

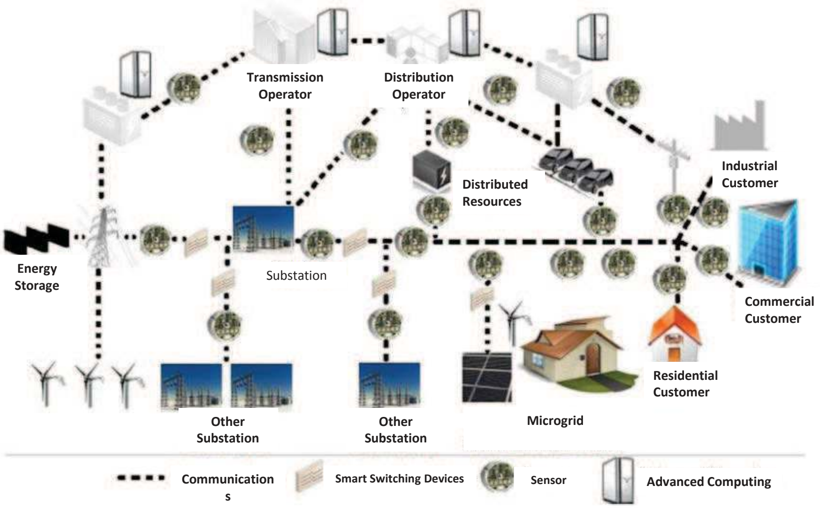


Image Source: EPRI

Advanced Grid Sensors Improve Smart Grid Operations

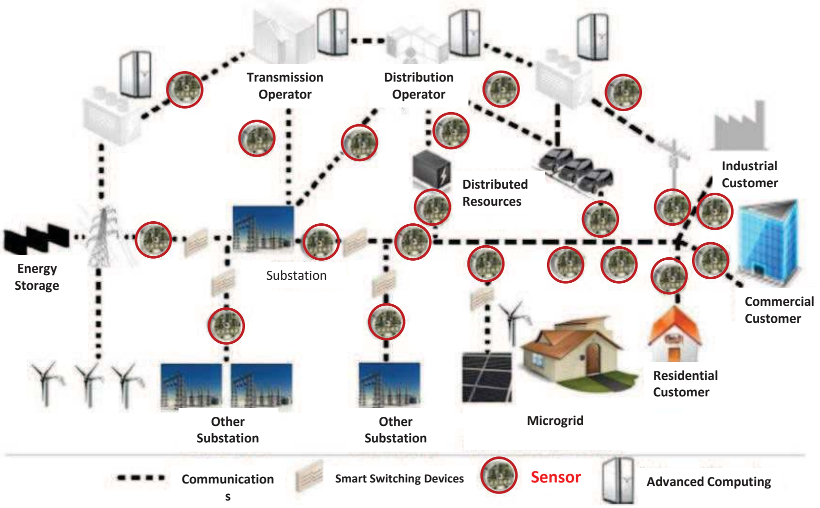


Image Source: EPRI

Data Quality in Future Grid

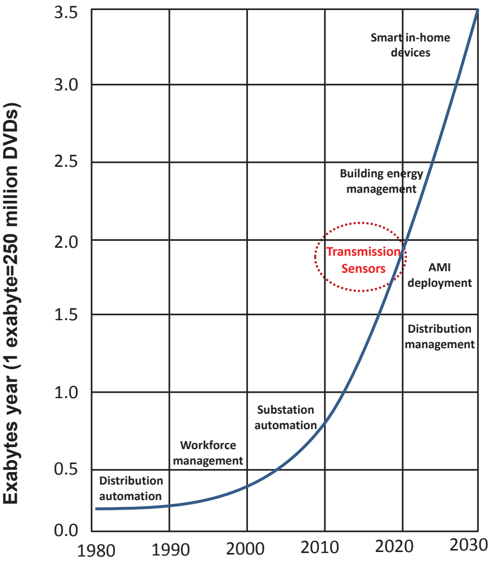


Image Source:EPRI

- Big data explosion!
  - Big data analytics
- **Data quality analysis**
  - Pre-processing for big data analytic
- **Transmission sensors**
  - **SCADA/PMU data**

Data Integrity in SCADA System

Stuxnet Worm, 2010

Nuclear power plant attacked via **SCADA** systems

Stuxnet: Malware more complex, targeted and dangerous than ever

Stuxnet: Computer worm opens new era of warfare

A Silent Attack, but Not a Subtle One

* SCADA (Supervisory Control And Data Acquisition)

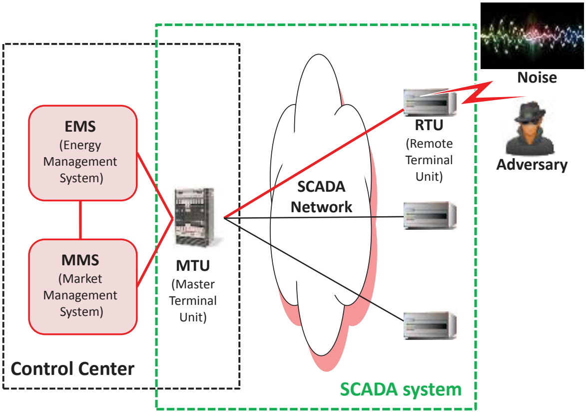
SmartGrid Update Report*

"Even *small* changes in the data could affect the stability of the grid and even jeopardize human safety"

SCADA Weak Cybersecurity+Data Integrity Violation $\Rightarrow$ Grid Malfunction

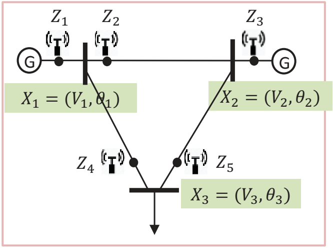
* http://www.smartgridupdate.com/dataforutilities/pdf/DataManagementWhitePaper.pdf

Motivation

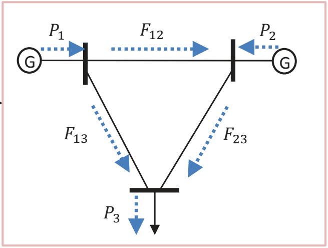


- **Data quality/integrity** violation ⇒ **blackouts** & **financial losses**

Three-bus system



Base case power flow solution



Measurement model

$$\boldsymbol{Z = h(X) + e}$$

$$Z = (Z_1, Z_2, Z_3, Z_4, Z_5)$$
$$e = (e_1, e_2, e_3, e_4, e_5)$$
$$X = (\theta_1, \theta_2, \theta_3, V_1, V_2, V_3)$$

Weighted least squares method

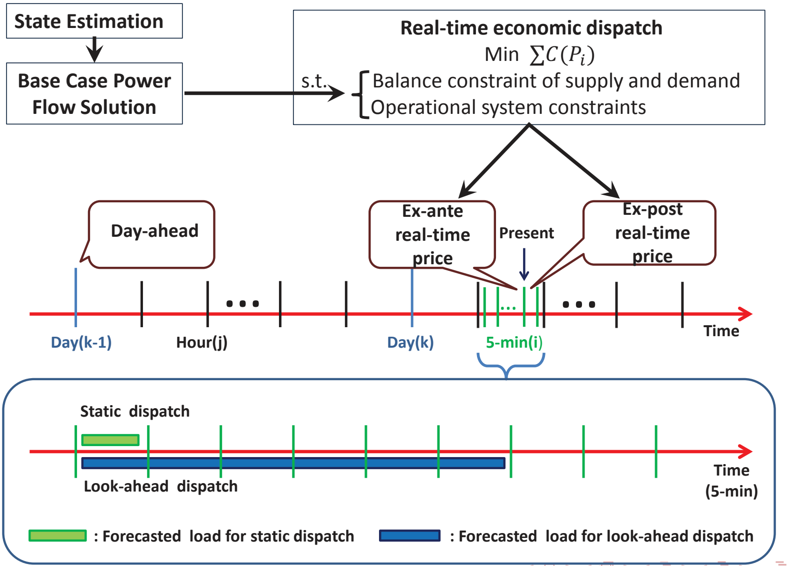
minimize $\boldsymbol{J(X) = r^T R^{-1} r}$

s.t $\boldsymbol{r = Z - h(X)}$

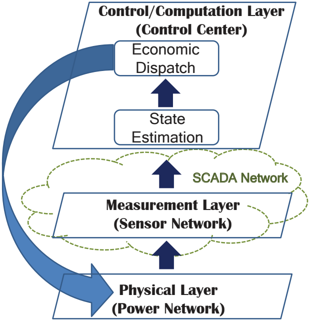
State estimation solution

$$\boldsymbol{\hat{X} = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{V}_1, \hat{V}_2, \hat{V}_3)}$$

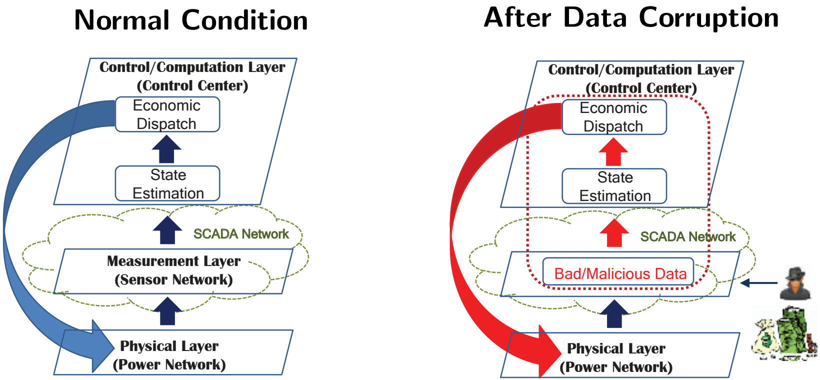
Background: Electricity Market Operations



Normal Condition



Problem Statement



1. What are the impacts of data integrity/quality on real-time market prices, namely locational marginal price (**LMP**), via state estimation?
2. What are analytical tools for quantifying such impacts?

Previous and Missing Work

I. Data Integrity Attack on Physical and Economical Grid Operations

- **Attack modeling & analysis based on continuous data manipulation:** [1, Liu et al., 2009], [2, Kosut et al., 2010], [3, Kim et al., 2011]
- **Attack modeling & analysis based on discrete data manipulation:** [4, Kim et al., 2013]
- **Data attack on static economic dispatch:** [5, Xie et al., 2011]
- **Data attack on look-ahead economic dispatch: ?**

II. LMP Sensitivity Analysis Subject to Power System Condition

- **Impact of physical system conditions (e.g., load variations) on LMP sensitivity:** [6, Conejo et al., 2005], [7, Li et al., 2007]
- **Impact of sensor data quality on LMP sensitivity: ?**

Research Goals

- **A Market Participant's Perspective**

Part I: Data Integrity Attack on Look-Ahead Economic Dispatch

- Ramp-induced data (RID) attack **[Choi, Xie, TSG2013]**
- *Undetectable* and *profitable* RID attack strategy
- Economic impact of RID attack

- **A System Operator's Perspective**

Part II: Sensitivity Analysis of LMP to Data Corruption

- Impact of *continuous* data quality on real-time LMP **[Choi, Xie, TPS2014]**
- Impact of *discrete* data quality on real-time LMP **[Choi, Xie, SmartGridComm2013]**

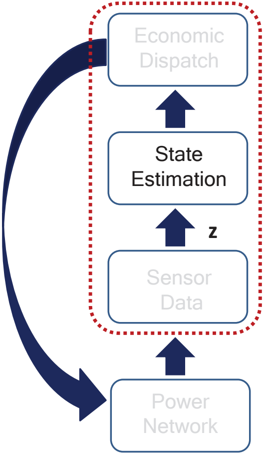
- **A Market Participant's Perspective**

RID Attack on Look-Ahead Dispatch in Real-Time Market

1. **Attack Modeling**
   - Generation capacity withholding
   - Covert change of generators' *inter-temporal* ramp constraints

2. **Performance Evaluation**
   - Undetectability
   - Profitability

State Estimation Model



- **Measurement Model** $\Rightarrow \mathbf{z} = \mathbf{S}\mathbf{x} + \mathbf{e}$
  - $\mathbf{z}$: measurements vector, $\mathbf{e} \sim \mathcal{N}(0, \mathbf{R})$
  - $\mathbf{S} = \begin{bmatrix} \mathbf{I} \\ \mathbf{H_d} \end{bmatrix}$: system factor matrix
  - $\mathbf{x}$: (nodal power injection) states vector
- **Weighted Least Squares Estimate**

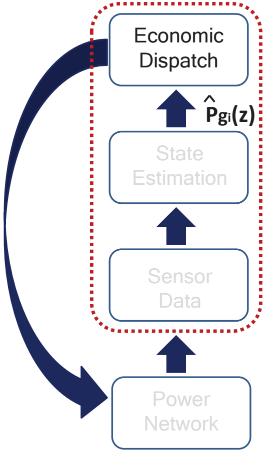
$$\hat{\mathbf{x}}(\mathbf{z}) = (\mathbf{S}^T\mathbf{R}^{-1}\mathbf{S})^{-1}\mathbf{S}^T\mathbf{R}^{-1}\mathbf{z} = \mathbf{B}\mathbf{z}$$

- **Bad Data Detection (Chi-squares test)**

$$J(\hat{\mathbf{x}}(\mathbf{z})) = \mathbf{r}^T\mathbf{R}^{-1}\mathbf{r} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_\chi$$

where $\mathbf{r} = \mathbf{z} - \mathbf{S}\hat{\mathbf{x}}(\mathbf{z})$

Economic Dispatch Model



▶ **Look-Ahead Dispatch Model**

$$\min_{P_{g_i}[k]} \sum_{k=1}^{K} \sum_{i \in G} C_i(P_{g_i}[k])$$

s.t.

$$\sum_{i \in G} P_{g_i}[k] = \sum_{n=1}^{N} D_n[k] \quad \forall k = 1, \ldots, K$$

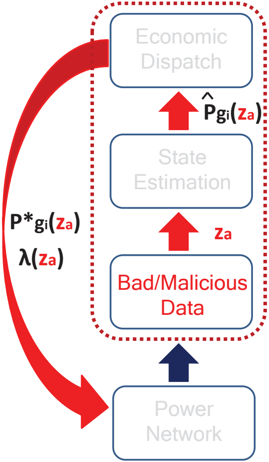
$$|P_{g_i}[k] - P_{g_i}[k-1]| \leq R_i \Delta T \quad \forall k = 1, \ldots, K$$

$$P_{g_i}^{\min} \leq P_{g_i}[k] \leq P_{g_i}^{\max} \quad \forall k = 1, \ldots, K$$

$$F_l^{\min} \leq F_l[k] \leq F_l^{\max} \quad \forall k = 1, \ldots, K$$

$$\forall l = 1, \ldots, L$$

Attack Target: $P_{g_i}[0]$ is updated with $\hat{P}_{g_i}(\mathbf{z})$ at every dispatch interval!



- **Attack Measurement Model**

  $\Rightarrow z_a = Sx + e + a$

  - z_a: corrupted measurement vector
  - a: injected attack vector

- **A Domino Effect of Data Attack**
  - $z_a \Rightarrow \hat{P}_{g_i}(z_a) \Rightarrow \lambda(z_a)$

Two Main Features of RID Attack: (1) Undetectability

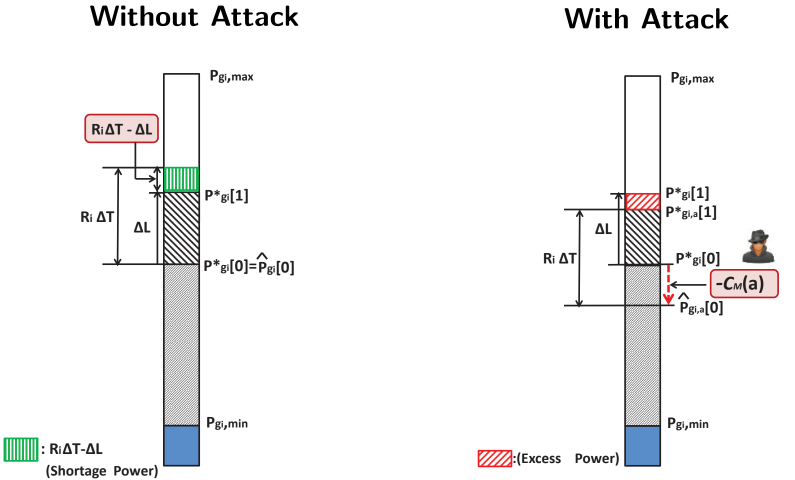
- After data attack, we have
  - New estimator: $\hat{\mathbf{x}}(\mathbf{z_a}) = \mathbf{B}\mathbf{z_a} = \hat{\mathbf{x}}(\mathbf{z}) + \mathbf{B}\mathbf{a}$
  - New residual:

$$||\mathbf{r}'||_2 = ||\mathbf{r} + (\mathbf{I} - \mathbf{SB})\mathbf{a}||_2 \leq \underbrace{||\mathbf{r}||_2}_{\text{Without attack}} + \underbrace{||(\mathbf{I} - \mathbf{SB})\mathbf{a}||_2}_{\text{With attack}}$$

- For **undetectability**, the attacker's goal is to
  - Construct $\mathbf{a}$ such that the contribution of $||(\mathbf{I} - \mathbf{SB})\mathbf{a}||_2$ still makes the following *healthy* detection condition hold true:

$$||\mathbf{r}'||_2 < \eta$$

Two Main Features of RID Attack: (2) Profitability



- Capacity withholding condition: $-\mathcal{C}_M(\mathbf{a}) > R_i \Delta T - \Delta L$
  - increasing LMP and profit

Attack Strategy: Compute the Attack Vector $\mathbf{a}$

$$\max_{\mathbf{a} \in \mathrm{span}(\mathcal{A})} \delta$$

s.t.

$$||(\mathbf{I} - \mathbf{SB})\mathbf{a}||_2 \leq \epsilon \Rightarrow \textbf{Undetectable Condition}$$

$$\alpha \mathcal{C}_M(\mathbf{a}) + \beta \mathcal{C}_B(\mathbf{a}) \leq \Delta L - R_i \Delta T - \delta \Rightarrow \textbf{Profitable Condition}$$

$$\delta > 0$$

where

$$\mathcal{C}_M(\mathbf{a}) = E[\hat{P}_{g_i,\mathbf{a}}[0] - P^*_{g_i}[0]] = \mathbf{B}_i\mathbf{a}$$

$$\mathcal{C}_B(\mathbf{a}) = \sum_{j \in \underline{G}^c_M} E[\hat{P}_{g_j,\mathbf{a}}[0] + R_j \Delta T - P^{\max}_{g_j}[0]] = \sum_{j \in \underline{G}^c_M} [\mathbf{B}_j\mathbf{a} + R_j \Delta T]$$

- $\alpha = 1$, $\beta = 0$: Marginal unit attack (Case I)
- $\alpha = 0$, $\beta = 1$: Binding unit attack (Case II)
- $\alpha = 1$, $\beta = 1$: Coordinated attack (Case III)

Simulation Setup

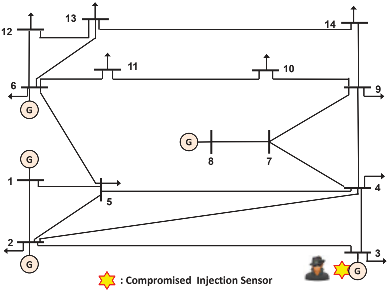


Figure : IEEE 14-Bus System.

Unit Type	P_{min}	P_{max}	Ramp Rate	Marginal Cost
Coal(1)	0MW	200MW	10MW/5min	30\$/MWh
Wind(2)	0MW	300MW	150MW/5min	20\$/MWh
Nuclear(3)	0MW	300MW	8MW/5min	40\$/MWh
Coal(6)	50MW	250MW	15 MW/5min	55\$/MWh
Oil(8)	60MW	150MW	60 MW/5min	60\$/MWh

Table : Generator Parameters.

Attack Performance

Case	Static (PE(3)%)	Look-ahead (PE(3)%)	J (η_χ=37.6)
I	131.9	148.9	28.2
II	101.2	102.6	35.5
III	108.9	113.8	31.5

- Case I: P_3 injection sensor compromised
- Case II: P_1 injection sensor compromised
- Case III: P_1, P_3 injection sensors compromised

Observation 1

- Attack profitability (PE(3) $>$ 100%)
- Attack undetectability ($J < \eta_\chi$=37.6)

Ramp-Induced Data Attack Increases LMPs

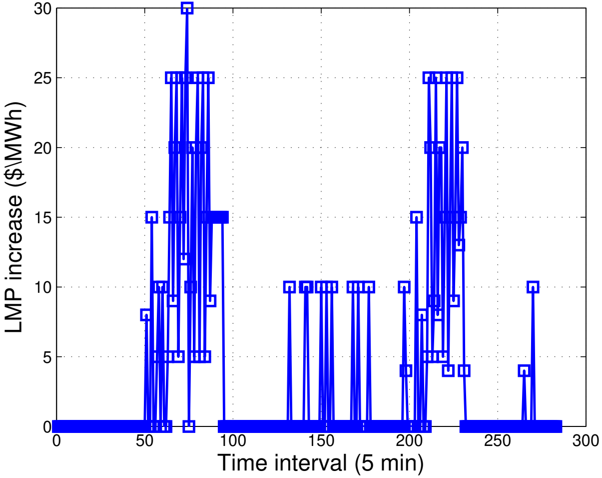


Figure : LMP Increase of Look-ahead Dispatch with Case 1 Attack.

Attack Relative Magnitude vs Attack Performance

	Attack Relative Magnitude (ARM) $\left(\frac{\|\|\mathbf{a}\|\|_\infty}{\|\|\mathbf{z}\|\|_\infty}\%\right)$			
	0.25	0.5	0.75	1
Static (PE(3))	111.8	120.8	126.4	126.9
Look-ahead (PE(3))	112.2	125.8	127.6	137.7
J	21.1	25.4	29.2	33.1

Observation 2

- Increasing ARM $\Rightarrow$ increasing attack profit at the expense of increasing J

Ramp Rate & Data Accuracy vs Attack Profit

	Ramp Rate (MW/5min)				Variance (σ^2)			
	8	10	12	14	0.0005	0.005	0.05	0.5
Static (PE(3))	131.9	119.7	106.4	100.5	123.2	129.1	130.3	136.9
Look-ahead (PE(3))	148.9	123.5	108.5	103.1	143.5	144.75	146.1	152.8

Observation 3

- A **slower** ramp rate unit targeted $\Rightarrow$ increasing attack profit

Observation 4

- A **less accurate** sensor compromised $\Rightarrow$ increasing attack profit

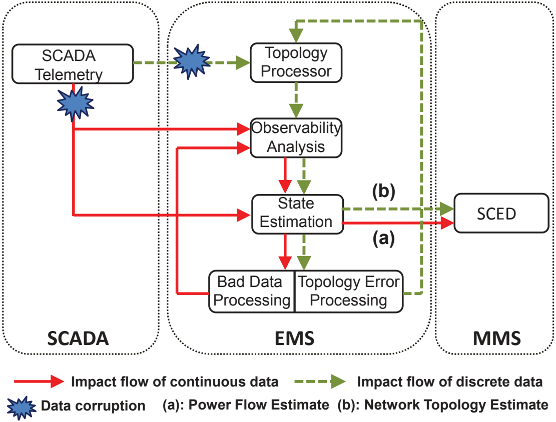
Part I: Remarks

Main Contributions

1. Problem formulation of a novel ramp-induced data attack
   - covert generation capacity withholding
2. An optimization-based undetectable/profitable attack strategy
3. Economic impacts on real-time electricity market operations

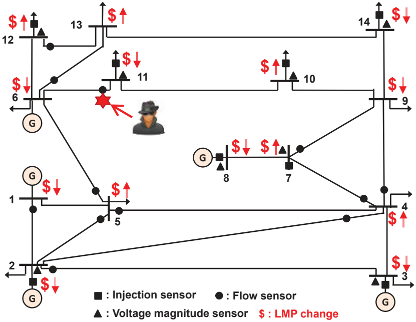
Part II: Sensitivity Analysis of LMP to Data Corruption

- **A System Operator's Perspective**

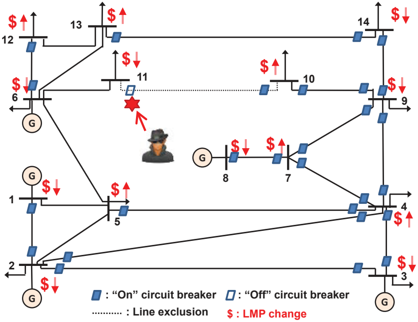


- **Part II-A:** impact of undetectable error in **(a)** on LMP
- **Part II-B:** impact of undetectable error in **(b)** on LMP

- Develop **analysis tools** to study the impact of data quality on LMP

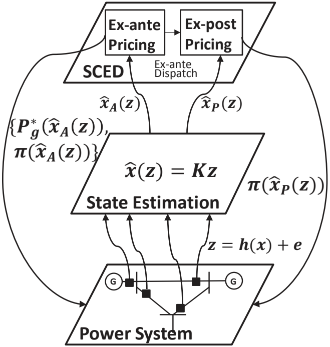


(a) Continuous data corruption (Part II-A)



(b) Discrete data corruption (Part II-B)

Part II-A: LMP Sensitivity to Continuous Data Corruption



- Composite function of the Ex-ante and Ex-post LMP vectors:

$$\mathbf{LMP} = \pi(\hat{\mathbf{x}}(\mathbf{z}))$$

- Proposed sensitivity matrix:

$$\mathbf{\Lambda} = \frac{\partial \pi}{\partial \mathbf{z}} = \underbrace{\frac{\partial \pi}{\partial \hat{\mathbf{x}}}}_{\mathbf{\Lambda}_A} \underbrace{\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{z}}}_{\mathbf{\Lambda}_B}$$

- $\mathbf{\Lambda}_A$: Sensitivity matrix of LMPs to state estimates (**Economic Impact**)
- $\mathbf{\Lambda}_B$: Sensitivity matrix of state estimates to sensor data (**Cyber Impact**)

$\Rightarrow$ A unified *closed-form* LMP sensitivity matrix $\mathbf{\Lambda}$

Continuous Data Corruption Manipulates LMP

$$\min_{P_{g_i}} \sum_{i=1}^{N_b} C_i(P_{g_i})$$

s.t.

$$\lambda(\mathbf{z_a}) : \sum_{i=1}^{N_b} P_{g_i} = \sum_{i=1}^{N_b} L_{d_i}$$

$$\boldsymbol{\tau}(\mathbf{z_a}) : \hat{P}_{g_i}^{\min}(\mathbf{z_a}) \leq P_{g_i} \leq \hat{P}_{g_i}^{\max}(\mathbf{z_a}) \qquad \forall i = 1, \dots, N_b$$

$$\boldsymbol{\mu}(\mathbf{z_a}) : F_l^{\min} \leq \sum_{i=1}^{N_b} S_{li}(P_{g_i} - L_{d_i}) \leq F_l^{\max} \qquad \forall l = 1, \dots, N_l$$

- **Domino effect:** $\mathbf{z_a} \Rightarrow \left\{ \hat{P}_{g_i}^{\min}(\mathbf{z_a}), \hat{P}_{g_i}^{\max}(\mathbf{z_a}) \right\} \Rightarrow \{\lambda(\mathbf{z_a}), \boldsymbol{\mu}(\mathbf{z_a})\} \Rightarrow \mathbf{LMP}(\mathbf{z_a})$

$$\mathbf{LMP}(\mathbf{z_a}) = \lambda(\mathbf{z_a})\mathbf{1}_{N_b} - \mathbf{S}^T \left[\boldsymbol{\mu}_{\max}(\mathbf{z_a}) - \boldsymbol{\mu}_{\min}(\mathbf{z_a})\right]$$

Derivation of Λ_A: KKT Condition Perturbation Approach

- **KKT equations**

$$\text{(i)}\quad \frac{\partial C_i(P_{g_i})}{\partial P_{g_i}} - \lambda + \sum_{j=1}^{B_g} \tau_j A_{ji} + \sum_{l=1}^{B_f} \mu_l S_{li} = 0 \qquad \forall i = 1, \ldots, N_b$$

$$\text{(ii)}\quad \sum_{i=1}^{N_b} P_{g_i} = \sum_{i=1}^{N_b} L_{d_i}$$

$$\text{(iii)}\quad \sum_{i=1}^{N_b} A_{ji} P_{g_i} = \hat{C}_j \qquad \forall j = 1, \ldots, B_g$$

$$\text{(iv)}\quad \sum_{i=1}^{N_b} S_{li}[P_{g_i} - L_{d_i}] = D_l \qquad \forall l = 1, \ldots, B_f.$$

- **Perturbed KKT equations**

$$\text{(i)}\quad \underbrace{\frac{\partial}{\partial P_{g_i}} \left(\frac{\partial C_i(P_{g_i})}{\partial P_{g_i}} \right)}_{M_i} dP_{g_i} - d\lambda + \sum_{j=1}^{B_g} A_{ji} d\tau_j + \sum_{l=1}^{B_f} S_{li} d\mu_l = 0 \quad \forall i = 1, \ldots, N_b$$

$$\text{(ii)}\quad \sum_{i=1}^{N_b} dP_{g_i} = \sum_{i=1}^{N_b} dL_{d_i}$$

$$\text{(iii)}\quad \sum_{i=1}^{N_b} A_{ji} dP_{g_i} = d\hat{C}_j \qquad \forall j = 1, \ldots, B_g$$

$$\text{(iv)}\quad \sum_{i=1}^{N_b} S_{li} dP_{g_i} = \sum_{i=1}^{N_b} S_{li} dL_{d_i} \qquad \forall l = 1, \ldots, B_f.$$

For example,

$$\text{(ii)}\ \sum_{i=1}^{N_b} P_{g_i} = \sum_{i=1}^{N_b} L_{d_i} \implies \sum_{i=1}^{N_b} dP_{g_i} = \sum_{i=1}^{N_b} dL_{d_i}$$

Derivation of $\boldsymbol{\Lambda}_A$: KKT Condition Perturbation Approach

- Perturbed KKT equations in matrix form

$$\underbrace{\begin{bmatrix} \mathbf{M} & -\mathbf{1}_{N_b} & \Upsilon \\ \mathbf{1}_{N_b}^T & 0 & \mathbf{0} \\ \Upsilon^T & \mathbf{0} & \mathbf{0} \end{bmatrix}}_{\boldsymbol{\Xi}} \begin{bmatrix} d\mathbf{P}_g \\ d\lambda \\ d\boldsymbol{\tau}_s \\ d\boldsymbol{\mu}_s \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{U_1}^T & \mathbf{U_2}^T \end{bmatrix}}_{\boldsymbol{\Phi}} \begin{bmatrix} d\mathbf{L}_d \\ d\hat{\mathbf{C}}_s \end{bmatrix}$$

- Sensitivity of lagrangian multipliers to estimated capacity limit

$$\begin{bmatrix} d\mathbf{P}_g \\ d\lambda \\ d\boldsymbol{\tau}_s \\ d\boldsymbol{\mu}_s \end{bmatrix} = \underbrace{\boldsymbol{\Xi}^{-1}\boldsymbol{\Phi}}_{\boldsymbol{\Lambda}_p} \begin{bmatrix} d\mathbf{L}_d \\ d\hat{\mathbf{C}}_s \end{bmatrix} \Longrightarrow \boldsymbol{\Lambda}_p = \begin{bmatrix} \boldsymbol{\Lambda}_{\mathbf{L}_d} & \Big| & \boldsymbol{\Lambda}_{\hat{\mathbf{C}}_s} \end{bmatrix} = \left[\begin{array}{c|c} \frac{\partial \mathbf{P}_g}{\partial \mathbf{L}_d} & \frac{\partial \mathbf{P}_g}{\partial \hat{\mathbf{C}}_s} \\ \frac{\partial \lambda}{\partial \mathbf{L}_d} & \frac{\partial \lambda}{\partial \hat{\mathbf{C}}_s} \\ \frac{\partial \boldsymbol{\tau}_s}{\partial \mathbf{L}_d} & \frac{\partial \boldsymbol{\tau}_s}{\partial \hat{\mathbf{C}}_s} \\ \frac{\partial \boldsymbol{\mu}_s}{\partial \mathbf{L}_d} & \frac{\partial \boldsymbol{\mu}_s}{\partial \hat{\mathbf{C}}_s} \end{array}\right]$$

Finally, $\boldsymbol{\Lambda}_A$ is constructed with $\frac{\partial \lambda}{\partial \hat{\mathbf{C}}_s}$ and $\frac{\partial \boldsymbol{\mu}_s}{\partial \hat{\mathbf{C}}_s}$

Derivation of $\mathbf{\Lambda}_B$: Iterative State Estimation Equation

- Gauss-Newton iterative equation for state estimation

$$d\hat{\mathbf{x}}^{k+1} = \underbrace{[\mathbf{G}(\hat{\mathbf{x}}^k)]^{-1}\mathbf{H}^T(\hat{\mathbf{x}}^k)\mathbf{R}^{-1}}_{\mathbf{\Psi}(\hat{\mathbf{x}}^k)} d\mathbf{z}^k$$

$$\Updownarrow$$

$$\left[\frac{d\hat{\boldsymbol{\theta}}^{k+1}}{d\hat{\mathbf{V}}^{k+1}}\right] = \left[\frac{\mathbf{\Psi}_{\hat{\theta}}(\hat{\mathbf{x}}^k)}{\mathbf{\Psi}_{\hat{\mathbf{V}}}(\hat{\mathbf{x}}^k)}\right] d\mathbf{z}^k$$

$$\Downarrow$$

- Sensitivity of linearized real power estimates to sensor data

$$d\hat{\mathbf{z}}_r = \begin{bmatrix} \mathbf{B}^S_{P\theta} \\ \mathbf{B}_{P\theta} \\ \mathbf{B}_{F\theta} \end{bmatrix} d\hat{\boldsymbol{\theta}} = \begin{bmatrix} \mathbf{B}^S_{P\theta} \\ \mathbf{B}_{P\theta} \\ \mathbf{B}_{F\theta} \end{bmatrix} \mathbf{\Psi}_{\hat{\theta}} d\mathbf{z}$$

- Desired sensitivity matrix

$$\mathbf{\Lambda}_B = \begin{bmatrix} \mathbf{B}^S_{P\theta} \\ \mathbf{B}_{P\theta} \\ \mathbf{B}_{F\theta} \end{bmatrix} \mathbf{\Psi}_{\hat{\theta}}$$

Simulation Setup

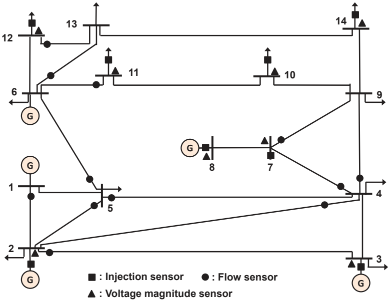
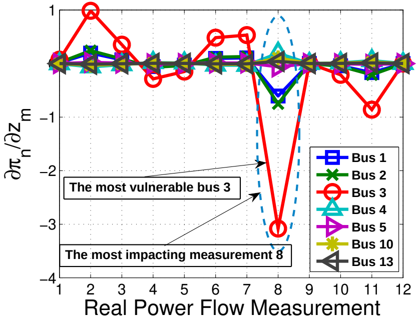


Figure : IEEE 14-bus system with a given measurement configuration.

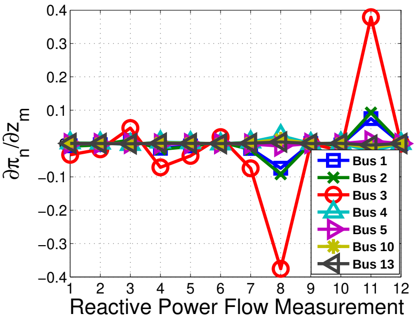
Table : Generator parameters in the IEEE 14-bus system.

Bus	$P_{g_i}^{\min}$(MW)	$P_{g_i}^{\max}$(MW)	a_i(\$/MWh)	b_i(\$/(MW)^{2}h)
1	0	332.4	20	0.043
2	0	140	20	0.25
3	0	100	40	0.01
6	0	100	40	0.01
8	0	100	40	0.01

Using a closed-form LMP sensitivity matrix $\mathbf{\Lambda} = \mathbf{\Lambda}_A \cdot \mathbf{\Lambda}_B$,



(a)

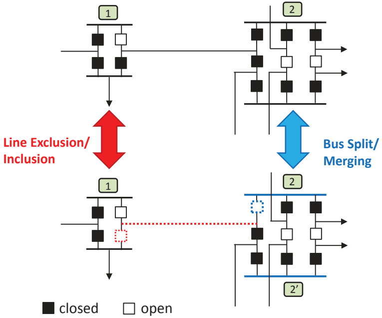


(b)

Key Observations

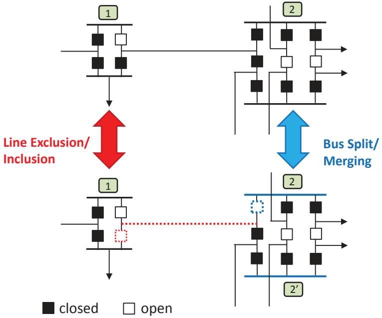
1. **Sensitivity grouping property**
   - *Identical* positive or negative sensitivity bus group to data corruption
2. **Economically sensitive physical and cyber assets**
   - *Buses* with LMP highly sensitive to data corruption.
   - Significantly influential *sensors* on LMP change.
3. **Impact of different types of sensor data on LMP**
   - A more significant impact of *real* power sensor data on LMP sensitivity

- Two types of topology error

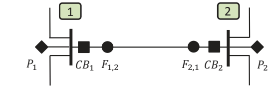


Part II-B: LMP Sensitivity to Network Topology Error

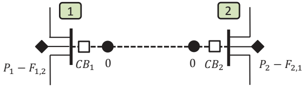
- Two types of topology error



- Attack scenario [4, Kim et al., 2013]



(a) Without attack



(b) With attack (line exclusion)

[Continuous data sensor]
- ◆ Injection sensor
- ● Flow sensor

[Discrete data sensor]
- ■ Circuit breaker sensor ("Closed=0")
- □ Circuit breaker sensor ("Open=1")

Topology Data Attack Manipulates LMP

$$\min_{p_i} \sum_{i \in G} C_i \cdot p_i$$

s.t.

$$\lambda(\mathbf{z_a}) : \sum_{n=1}^{N_b} P_{g_n} = \sum_{n=1}^{N_b} L_{d_n}$$

$$\boldsymbol{\tau}(\mathbf{z_a}) : p_i^{\min} \leq p_i \leq p_i^{\max} \qquad \forall i \in G$$

$$\boldsymbol{\mu}(\mathbf{z_a}) : F_l^{\min} \leq \sum_{n=1}^{N_b} \widehat{H_{l,n}}(\mathbf{z_a})(P_{g_n} - L_{d_n}) \leq F_l^{\max} \qquad \forall l = 1, \ldots, N_l$$

- **Domino effect:** $\mathbf{z_a} \Rightarrow \widehat{H_{l,n}}(\mathbf{z_a}) \Rightarrow \{\lambda(\mathbf{z_a}), \boldsymbol{\mu}(\mathbf{z_a})\} \Rightarrow \mathbf{LMP}(\mathbf{z_a})$

$$\mathbf{LMP}(\mathbf{z_a}) = \lambda(\mathbf{z_a})\mathbf{1}_{N_b} - \widehat{\mathbf{H}}(\mathbf{z_a})^T [\boldsymbol{\mu}_{\max}(\mathbf{z_a}) - \boldsymbol{\mu}_{\min}(\mathbf{z_a})]$$

LMP Sensitivity Analysis to Topology Error

Proposition 1 (A Closed-Form Shadow Price Expression)

The shadow price μ_l for the congested transmission line l:

$$\mu_l = \frac{\Delta C(j,i)}{\Delta H_l(i,j)}$$

where

$$\Delta C(j,i) = C_j - C_i \quad \text{: Marginal Unit Energy Costs Difference}$$
$$\Delta H_l(i,j) = H_{l,i} - H_{l,j} \text{ : Distribution Factors Difference}$$

LMP Sensitivity Analysis to Topology Error (cont'd)

Corollary 2 (A Closed-Form LMP Sensitivity Index to Topology Error)

LMP sensitivity index with respect to the line k status error ($k \neq l$):

$$\Delta \mathbf{LMP}_l^k = \Delta C(j,i)\mathbf{v}_l^k$$

where

$$\Delta \mathbf{LMP}_l^k = \left[\Delta LMP_{l,1}^k, \ldots, \Delta LMP_{l,N_b}^k\right]^T$$

$$\mathbf{v}_l^k = [v_{l,1}^k, \ldots, v_{l,N_b}^k]^T, \quad v_{l,n}^k = \frac{\tilde{H}_{l,n}^k}{\Delta \tilde{H}_l^k(i,j)} - \frac{H_{l,n}}{\Delta H_l(i,j)}$$

- **Benefit:** less computational time than exhaustive numerical simulations

LMP Sensitivity Analysis to Topology Error (cont'd)

Corollary 3

(a) $v^k_{l,n} > 0 \Rightarrow$ decreasing LMP at bus n with topology error
- A quick prediction of post-LMP direction by topology error

(b) $|v^k_{l,n}| > |v^k_{l,m}| \Rightarrow$ LMP sensitivity at bus n is higher than at bus m
- A quick comparison of LMP sensitivity magnitude

(c) Increasing $\Delta C(j, i) \Rightarrow$ increasing LMP sensitivity at any bus
- Guidelines for a bidding strategy of generation company

Simulation Setup

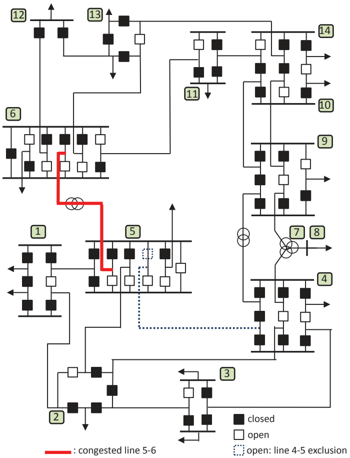


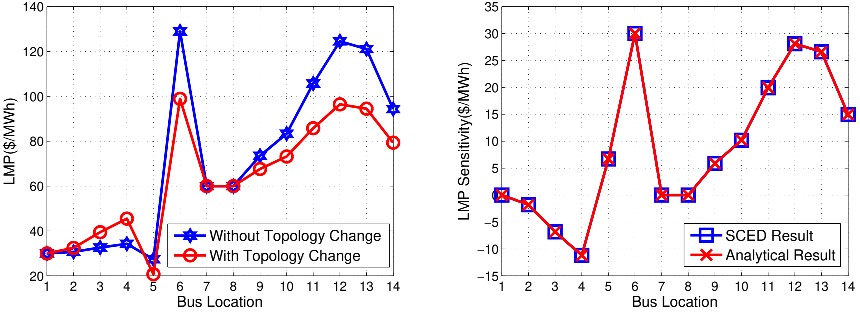
Figure : IEEE 14-bus system including bus-breaker model.

Table : Generator parameters of the IEEE 14-bus system.

Bus	$P_{\min}$	$P_{\max}$	Marginal Cost
1	0MW	330MW	30\$/MWh
2	0MW	140MW	20\$/MWh
3	0MW	100MW	40\$/MWh
6	0MW	100MW	55\$/MWh
8	0MW	100MW	60\$/MWh

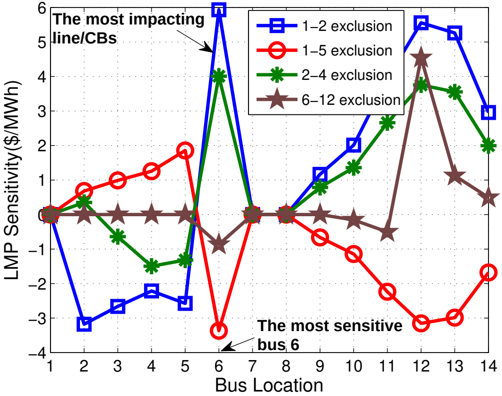
- Line 5-6 is congested
- Line 4-5 is excluded due to data corruption

Topology Errors Significantly Change LMPs



(a) (b)

LMP Sensitivities Under the Same Line 5-6 Congestion



- The highest sensitivity at **bus 6** to line 1-2, 1-5 and 2-4 exclusions
- **Line 1-2 exclusion** (blue plot) changes sensitivities the most

Part II: Remarks

Main Contributions

1. New analytical frameworks to study real-time LMP sensitivity with respect to data corruption
2. Derivation of closed-form LMP sensitivity analysis tools
   - economically sensitive buses to data corruption
   - influential sensors and transmission lines on LMP change
3. Easily integrated with the existing EMS/MMS

Conclusions

- **Impact of Data Integrity/Quality on Economic Dispatch**

Part I

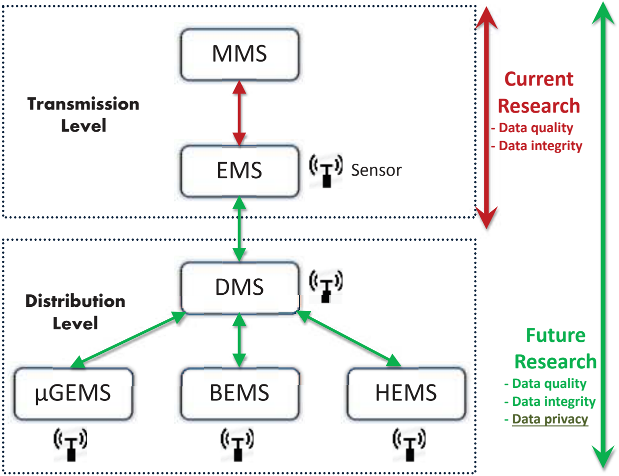
- Data Attack on Look-Ahead Dispatch
  - A market participant's perspective
  - Feasible ramp-induced data (RID) attack strategy for:
    - Undetectability
    - Profitability

Part II

- LMP Sensitivity to Data Corruption
  - A system operator's perspective
  - Analytical tools for LMP sensitivity quantification with respect to:
    - Continuous Data Corruption
    - Discrete Data Corruption

The Bigger Picture

Data Quality, Integrity, Privacy-Aware Multi-Scale Decision Making Tool



DMS: Distribution Management System, μGEMS: Microgrid Energy Management System
BEMS: Building Energy Management System, HEMS: Home Energy Management System

Multidisciplinary Approach to Future Work

- **A Unified System-Wide Monitoring Tool for Multi-Scale Spatial Data Quality Analysis**


Smart Meter


Solar Power


Electric Vehicle


Energy Storage

  - Design of interface between EMS and DMS
  - Performance index for the impact analysis of distribution **data quality**
  - **Power system engineering**, operations research/optimization

- **Smart Grid Cyber Security and Privacy**
  - **Data integrity** attack modeling and countermeasures
  - Smart meter **data privacy**-preserving algorithm
  - **Power system engineering**, computer networking, cyber security, statistical signal processing

References

[1] Y. Liu, M. K. Reiter, and P. Ning, “False data injection attacks against state estimation in electric power grids,” *Proc. 16th ACM Conf. Comput. Commun. Security*, 2009.

[2] O. Kosut, L. Jia, R. Thomas, and L. Tong, “Malicious data attacks on smart grid state estimation: Attack strategies and countermeasures,” *2010 First International Conference on Smart Grid Communications*, October 2010.

[3] T. T. Kim, and H. V. .Poor, “Strategic Protection Against Data Injection Attacks on Power Grids,” *IEEE Trans. Smart Grid*, vol. 2, no. 2, pp. 326–333, May 2011.

[4] J. Kim and L. Tong, “On Topology Attack of a Smart Grid: Undetectable Attacks and Counter Measures,” *IEEE J. Selected Areas in Communications*, July 2013.

[5] L. Xie, Y. Mo, and B. Sinopoli, “Integrity Data Attacks in power market operations,” *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 659–666, December 2011.

[6] A. J. Conejo, E. Castillo, R. Minguez, and F. Milano, “Locational marginal price sensitivities,” *IEEE Trans. Power Syst*, vol. 20, no. 4, pp. 2026–2033, November 2005.

[7] F. Li and R. Bo, “DCOPF-based LMP simulation: Algorithm, comparison with ACOPF, and sensitivity,” *IEEE Trans. Power Syst*, vol. 22, no. 4, pp. 1475–1485, November 2007.

Publications (During Ph.D. Study)

- Journal

[1] **D.-H Choi** and L. Xie, "Sensitivity Analysis of Real-Time LocationalMarginal Price to SCADA Sensor Data Corruption," *IEEE Trans. Power Syst* (accepted).

[2] **D.-H Choi** and L. Xie, "Ramp-Induced Data Attacks on Look-ahead Dispatch in Real-time Power Markets," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1235–1243, September 2013.

[3] L. Xie, **D.-H. Choi**, S. Kar and H. Vincent Poor "Fully Distributed State Estimation for Wide-Area Monitoring Systems," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1154–1169, September 2012.

[4] S. Wang, L. Cui, J. Que, **D.-H. Choi**, X. Jiang, S. Cheng and L. Xie "A Randomized Response Model for Privacy Preserving Smart Metering," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1154–1169, September 2012.

- Conference Proceedings

[5] **D.-H Choi** and L. Xie, "Impact Analysis of Locational Marginal Price Subject to Power System Topology Errors," *2013 Fourth International Conference on Smart Grid Communications*, October 2013.

[6] **D.-H Choi** and L. Xie, "Malicious Ramp-Induced Temporal Data Attack in Power Market with Look-Ahead Dispatch" *2012 Third International Conference on Smart Grid Communications*, November 2012 (**The Best Paper Award**).

[7] **D.-H Choi** and L. Xie, "Fully Distributed Bad Data Processing for Wide Area State Estimation," *2011 Second International Conference on Smart Grid Communications*, October 2011.

[8] L. Xie, **D.-H. Choi** and S. Kar, "Cooperative distributed state estimation: Local observability relaxed," *Proc. IEEE Power and Energy Society General Meeting*, Detroit, 2011.

- Book Chapter

[9] L. Xie, **D.-H. Choi**, S. Kar and H. Vincent Poor "Bad/malicious data detection in distributed power system state estimation," *in E. Hossain, Z. Han, and H. V. Poor, editors, Smart Grid Communications and Networking*, Cambridge University Press, 2012 (to appear).

Thank You!