Smart-Grid Data Integrity Attacks: Characterizations and Countermeasures

Kameshwar Poolla
UC Berkeley

May 24, 2011

Co-conspirators

- Annarita Giani [Berkeley]
- Eilyan Bitar [Berkeley]
- Miles McQueen [Idaho National Labs]
- Pramod Khargonekar [Florida]

... and thanks to many useful discussions with:
Anjan Bose, Joe Eto, Pravin Varaiya

Outline

Cybersecurity of Electricity Grids

SCADA/EMS

Assumptions

- Simply connected power system
- n buses
  - injection buses [loads or generators]
  - null buses [zero external real power injected]
- Standard DC power flow assumptions [for now]
  - Quasi-steady state, balanced operation
  - Lossless lines
  - Reactive power support at all nodes
  - Small power angles

Power Meters

- Real power flow meters
  - Every injection bus
  - Some lines [need only one per line because lossless]
- The reality
  - All injection buses are metered [for settlement]
  - Only large > 50 MW are connected to SCADA network
  - Only large lines are metered [$\sim 10\%$]
  - Data transmitted to EMS at system operator every 2-10 seconds

Models

x	bus voltage angles
y_1	injection meters
y_2	line meters
y	$\begin{bmatrix} 0 & y_1 & y_2 \end{bmatrix}^T \in \mathbb{R}^{m+n}$
h	$\mathbb{R}^n \rightarrow \mathbb{R}^{m+n}$
H	bus susceptance matrix $\in \mathbb{R}^{(m+n)\times n}$
m	number of line meters
n	number of buses [including zero injections]

- general nonlinear model $\mathcal{M}: \quad y = h(x)$
- DC power flow $\mathcal{L}: \quad y = Hx$

Attacks

Definition

Attack $\mathcal{A} = (\mathbb{S}, a)$ consists of

$$\begin{aligned} \mathbb{S} \quad & \text{set of compromised meters} \\ a \neq 0 \quad & \text{attack vector } a = [0\ a_1\ a_2]^T \in \mathbb{R}^{m+n} \end{aligned}$$

Sparsity of $\mathcal{A} = |\mathbb{S}|$

- Meter readings y are changed to $y + a$
- Nonzero components of $a \leftrightarrow$ compromised meters $\mathbb{S}$

Unobservable Attacks

General model $y = h(x)$
Current operating point x^o
Current measurements $y^o = h(x^o)$

Definition

$\mathcal{A}$ is *unobservable* at operating pt x^o wrt the model $\mathcal{M}$ if

$$\exists \;\; x^a : \quad y^o + a = h(x^o + x^a)$$

i.e. there exists some system state consistent with the compromised observations and model
x^a is the *perceived state* perturbation associated with $\mathcal{A}$

Unobservable Attacks ...

For linear DC power flow model $\mathcal{L}$

- $\mathcal{A}$ is unobservable $\iff$ $a = Hx^a$ is solvable
- unobservability is independent of current operating point

Theorem

Consider DC power flow model $y = Hx$ and attack $\mathcal{A} = (\mathbb{S}, a)$.
K delete rows corresp to $\mathbb{S}$ from H

(a) $\mathcal{A}$ is unobservable if and only if rank $(K) < n$

(b) Attack vector a must lie in the subspace:

$$\mathcal{T} = \{a \in \mathbb{R}^{m+n} : a = Hx,\ Kx = 0\}$$

Coordination

- Unobservable attacks require coordination
- Attack vector must be arranged carefully across spatially separated meters
- Requires that attacker knows the model [and operating point]
- Consequence: low sparsity attacks are more probable
- Low sparsity unobservable attacks
  - Liu *et al* [2009]
  - 3-sparse attacks commonly exist in power system examples
- Our research focus
  - What do sparse attacks look like?
  - What countermeasures can we take?

Observable Islands

Definition

$\mathcal{A} = (\mathbb{S}, a)$: unobservable attack.

x^a: perceived state perturbation.

Partition the set of buses $\mathbb{V}$ into disjoint union

$$\mathbb{V} = \mathbb{V}_1 \cup \cdots \cup \mathbb{V}_s, \quad \mathbb{V}_i \cap \mathbb{V}_j = \phi \text{ for } i \neq j$$

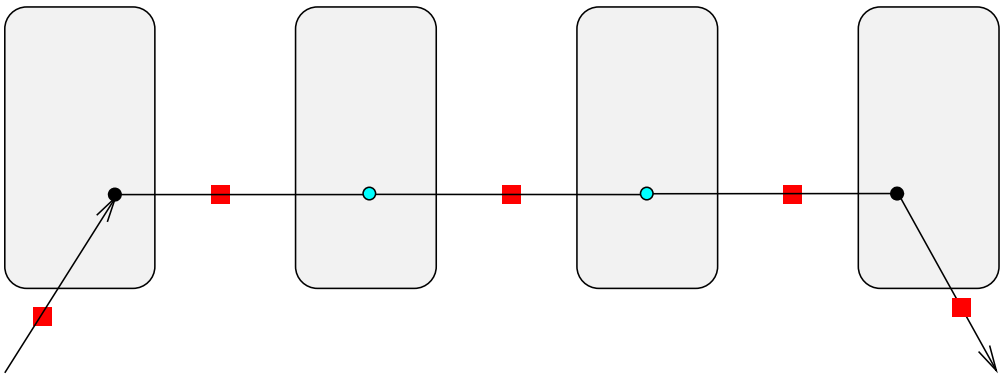
defined by the equivalence classes

$$v_1, v_2 \in \mathbb{V}_i \iff x^a(v_1) = x^a(v_2)$$

The sets $\{\mathbb{V}_i, \cdots, \mathbb{V}_s\}$ are the *observable islands* of attack $\mathcal{A}$.

Observable Islands Interpretation

- Unobservable attack $\mathcal{A}$
- Corresponds to some apparent power flow perturbation
- Apparent power flow must be on lines connecting islands
- No power flow between buses within same island



Related Work

- Liu et al (2009)
  Showed existence of sparse unobservable attacks for common power system examples
- Bobba et al (2010)
  Countermeasures using PMUs
- Kosut et al (2010)
  Sufficient characterizations of unobservable attacks in terms of graph properties
- Teixeira et al (2010)
  Measure of cyberattack resilience of a power system
  Studies of vulnerability with nonlinear models
- Pasqualetti et al (2011)
  Dynamic detection of attacks

Main Results

1. Irreducible Attacks
2. Attacks Involving 2 Power Injection Meters
3. Characterization of Sparse Attacks
4. Countermeasures using Known-secure PMUs
5. Countermeasures based on State-estimation
6. Beyond the DC Power Flow Model

A. Irreducible Attacks

Definition

$\mathcal{A} = (\mathbb{S}, a)$ is *irreducible* $\iff$ there is no unobservable attack $\mathcal{A}' = (\mathbb{S}', a')$ with $\mathbb{S}' \subsetneq \mathbb{S}$

Theorem

Consider DC power flow model $y = Hx$ *and attack* $\mathcal{A} = (\mathbb{S}, a)$.

K *delete rows corresp to* $\mathbb{S}$ *from* H

L *keep row corresp to* $\mathbb{S}$ *in* H

(a) *If* $\mathcal{A}$ *is irreducible then rank* $(K) = n - 1$

(b) *Suppose rank* $(K) = n - 1$. *Let* $0 \neq x \in \mathcal{N}(K)$. *If all the entries of the vector* Lx *are nonzero, then* $\mathcal{A}$ *is irreducible.*

(c) *Attack vector* a *must lie in the one dimensional subspace:*

$$\mathcal{T} = \{a \in \mathbb{R}^{m+n} : a = Hx,\ Kx = 0\}$$

B. Attacks Involving 2 Power Injection Meters

- Finding all irreducible attacks is NP hard
  - $\equiv$ finding minimal sets of rows of H whose deletion reduces rank by 1
  - $O\left(mn!/k!(n-k)!\right)$ flops if there are k compromised injection meters
  - Computationally intractable problem even for small power networks
- Fast algorithm in special case of exactly 2 power injection meters
  - $O(n^2m)$ flops
  - 1 minute on a 3Ghz PC for CAISO 4000 bus system

Algorithm

0 | define $G = \begin{bmatrix} H_0^T & H_1^T \end{bmatrix}^T$
$X = H_2 G^{-1}$

1 | select any two injection nodes i and j

2 | define $M^{i,j} = X[e_i e_j] \in \mathbb{R}^{m \times 2}$
where $e_i, e_j \in \mathbb{R}^n$ are unit vectors

3 | use elementary column operations to factorize

$$M^{i,j}\ T = \begin{bmatrix} 0_{r\times 1} & 0_{r\times 1} \\ 1 & 0 \\ 0 & 1 \\ x_1 & x_2 \end{bmatrix}$$

4 | $I_1 =$ indices of nonzero rows of x_1
$I_2 =$ indices of nonzero rows of x_2

5 | $\mathcal{A}_1 = ([i, j, r+2, r+I_1], a)$
$\mathcal{A}_2 = ([i, j, r+1, r+I_2], a)$

6 | goto 1

C. Characterization of Sparse Attacks

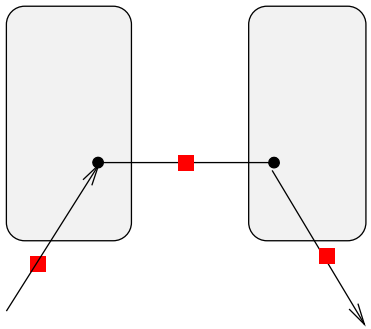
- Assume all lines are metered
- Can characterize 3,4,5 sparse attacks graphically

Theorem

$\mathcal{A} = (\mathcal{S}, a)$ is a 3-sparse irreducible attack $\iff$

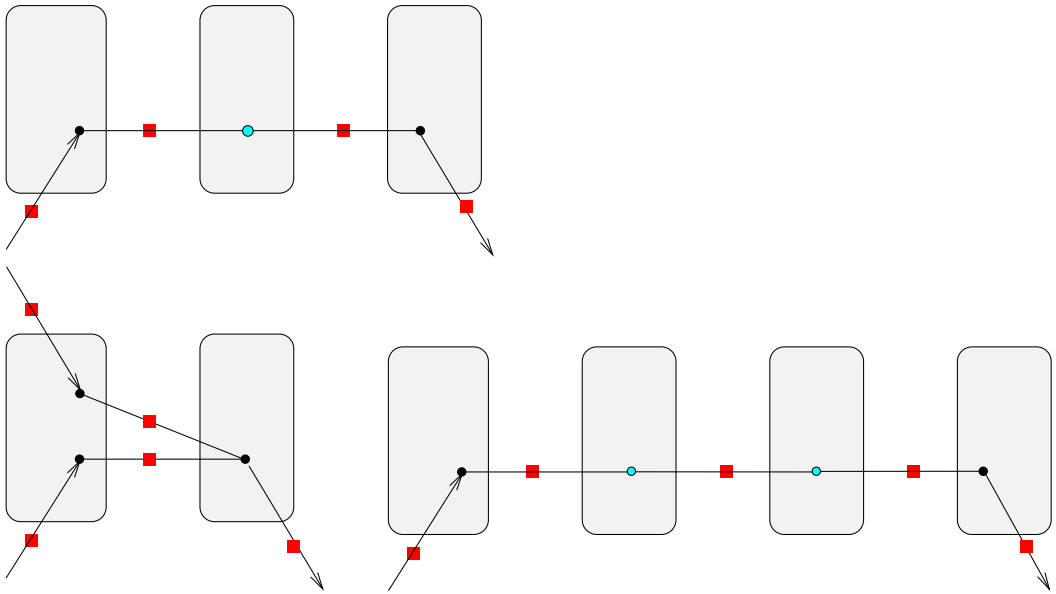
(a) $\mathbb{S}$ consists of power meters at two adjacent injection buses and the power meter on the line ℓ connecting these buses

(b) ℓ is a cutset of the power system graph



Characterization of Sparse Attacks ...

- Canonical forms for 4 sparse attacks
- Canonical forms for 5 sparse attacks
- Algorithm to find all k-sparse attacks, $k = 3, 4, 5$
  - find all instances of canonical forms in power system graph
  - depth-first search $O(n^2)$



D. Countermeasures using Known-secure PMUs

- PMUs are on the new NASPInet architecture
- Designed for security? Encrypted?
- Main idea
  - Disabling an attack with secure PMUs
  - Consider an unobservable attack $\mathcal{A}$
  - Associated observable islands $\mathbb{V}_1, \cdots, \mathbb{V}_s$
  - If $\mathcal{A}$ occurred, voltage phases between any pair of islands is $\neq 0$
  - Disable $\mathcal{A} \iff$ 2 PMUs are placed in any 2 distinct islands

Countermeasures using Known-secure PMUs ...

Theorem

Arbitrary collection of unobservable attacks $\mathbb{A} = \{\mathcal{A}_1, \cdots \mathcal{A}_p\}$
Observable islands:

$$\begin{array}{lccccc} \textit{attack } \mathcal{A}_1 & \mathbb{V}_1^1 & \mathbb{V}_2^1 & \cdots & \mathbb{V}_{s_1}^1 & \\ \textit{attack } \mathcal{A}_2 & \mathbb{V}_1^2 & \mathbb{V}_2^2 & \mathbb{V}_3^2 & \cdots & \cdots \quad \mathbb{V}_{s_2}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ \textit{attack } \mathcal{A}_p & \mathbb{V}_1^p & \cdots & \mathbb{V}_{s_p}^p & & \end{array}$$

$\mathbb{A}$ *is made observable by PMUs placed at buses* $\mathbb{B}$

$\iff \quad \forall k, \exists i_1 \neq i_2 : \ \mathbb{V}_{i_1}^k \cap \mathbb{B} \neq \phi, \mathbb{V}_{i2}^k \cap \mathbb{B} \neq \phi$

i.e. every attack has two distinct islands which contain PMUs

Countermeasures using Known-secure PMUs ...

- minimal PMU placement $\equiv$ vertex touching problem
- NP hard
- 2X approximations are P

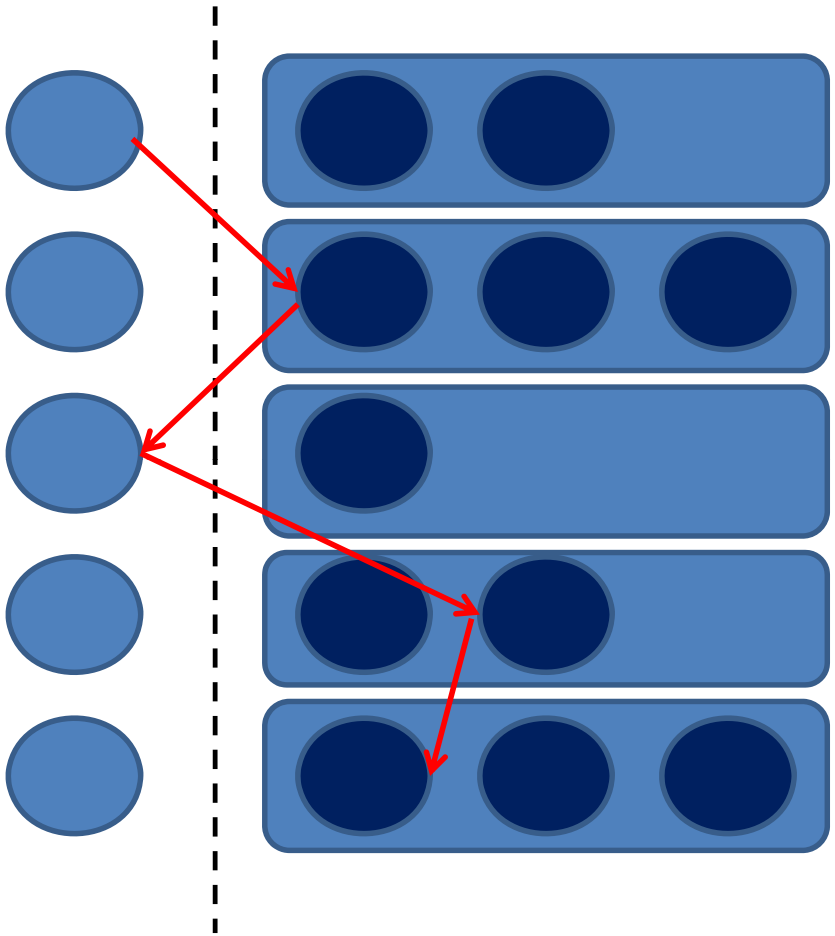
Theorem

Consider any *collection* $\mathbb{A} = \{\mathcal{A}_1, \cdots \mathcal{A}_p\}$ *of* p *unobservable attacks. It is sufficient to place* $p+1$ *PMUs at specific buses to make every attack in the collection* $\mathbb{A}$ *observable.*

- algorithm complexity: $O(n^2 p)$ flops
- sketch of proof [next slide]

Countermeasures using Known-secure PMUs ...

Sketch of proof:



E. Countermeasures using State Estimation

- Method
  - No additional sensors
  - Modify BDD algorithm
- Idea:
  - Suppose we have an unobservable attack $\mathcal{A}$ at time t^o
  - State estimator will predict phase angle translations on observable islands
  - Very unlikely if we have large islands

$$\gamma(\mathcal{A}) = \text{size of second largest island}$$

  - γ is a measure of detectability of $\mathcal{A}$
- Attacks with large γ: sparse [few islands] – easily detected
- Attacks with small γ: involve many meters – improbable

F. Beyond the DC Power Flow Model

- Topological Observability: Korres *et al* [2003], Monticelli [2000]
- Attack $\mathcal{A}(\mathcal{S}, a)$ unobservable wrt linear model $\mathcal{L}$
  $\iff$ $\mathcal{A}(\mathcal{S}, \hat{a})$ unobservable wrt nonlinear model $\mathcal{N}$
- attack vector $\hat{a}$ for nonlinear model lives on a manifold
  Sub-tangent space of $h(x)$ at current oper point
- Vulnerable sensor set $\mathcal{S}$ is independent of
  oper point, bus admittance values
  depends only on topology

Future Directions

- Computational aspects of attack detection and classification
- Stale data problems
- Noise issues – graded observability notions
- Low-grade financially motivated cyberattacks
- Grid operations and cybersecurity