USC

USC Viterbi
School of Engineering

On the Conditional Mutual Information in Gaussian-Markov Structured Grids

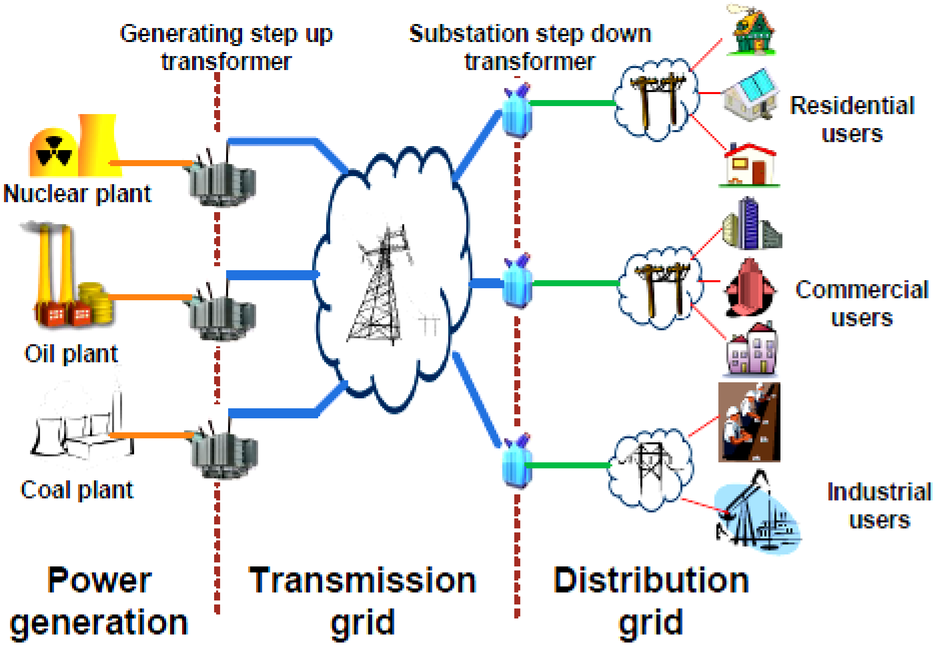
Hanie Sedghi &
Edmond Jonckheere

Agenda

- “Smart” Grid
- Phasor Measurement Units
- False Data Injection Attacks
- Gaussian Markov Random Field
- DC power flow
- Conditional Covariance Test
- Stealthy Deception False Data Injection Attack
- Attack Detection
- Conclusion and Future Works

Traditional Grid

- Traditional Grid
  - *electricity generation, electricity transmission, electricity distribution, and voltage/frequency stability control*

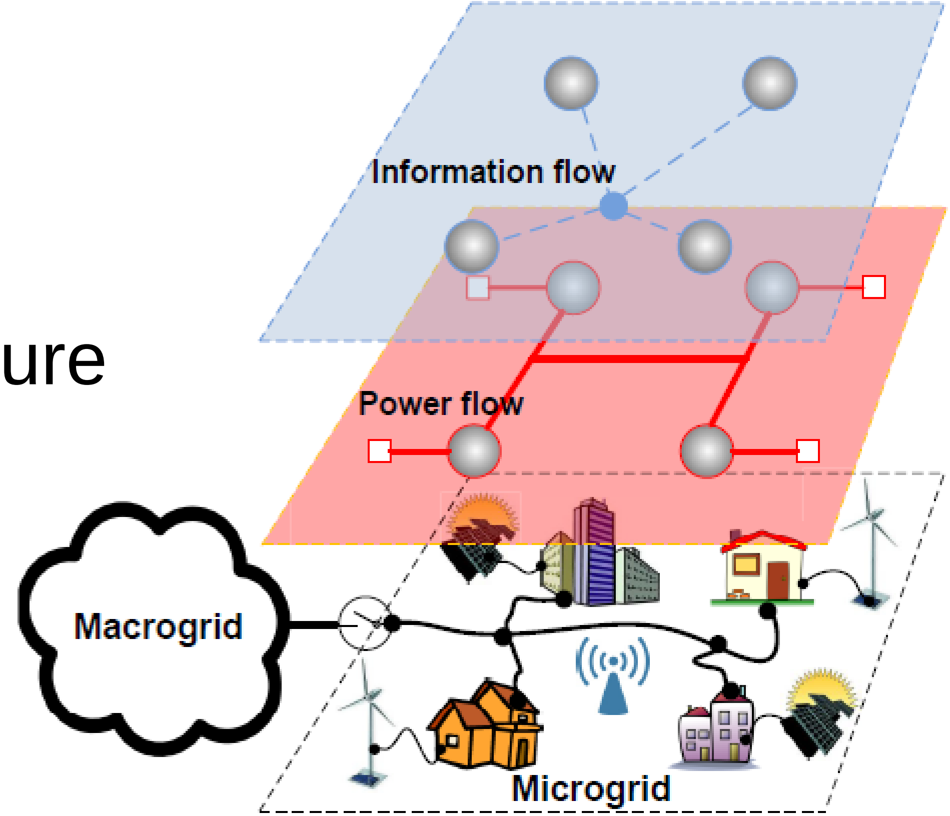


  - *Colocation of generation and distribution*

“Smart” Grid

- New Grid
  - *a large-scale generation-transmission-distribution NETWORK*
  - *Management and Control*

Overall architecture



- *Large scale power flow across the grid to allow consumers to purchase electricity at cheaper prices*

Fault Detection

- Today's power systems
  - *not adequately equipped with fault diagnosis mechanisms against various attacks*
- Fast and accurate uncovering of possibly malicious events
  - *preventing faults that may lead to blackouts*
  - *routine monitoring and control tasks of the smart grid, including state estimation and optimal power flow*
- Fault localization in nation’s grid
  - *challenging*
    - *due to the massive scale and inherent complexity*

Phasor Measurement Units (PMU’s)

- Synchronous PMU's with GPS time stamp
  - *being massively deployed across the grid*
  - *considered the most reliable sensing information to monitor the state of “health" of the grid*
- Recently Suggested Applications:
  - *Voltage security to avoid voltage collapse by using synchronized PMU measurements and decision tree*
  - *Fault detection through apparent changes in the bus susceptance parameters using PMU phase angles and generalized likelihood ratio*
  - *Detecting line outages using PMU angle measurements and Lasso, to avoid cascading events*
  - *And so many more!*

False Data Injection Attack

- False Data Injection attack refers to PMU data being manipulated before reaching the aggregator.
- All of the suggested applications fail in case of False Data Injection attack.
- PMU's are being massively deployed for "Smart" Grid control and monitoring.

It is crucial to have a mechanism to guarantee reliability of PMU data.

- We will consider the most recent false data injection attack that is capable of deluding the state estimator. Prior to us, no remedy was suggested for it.

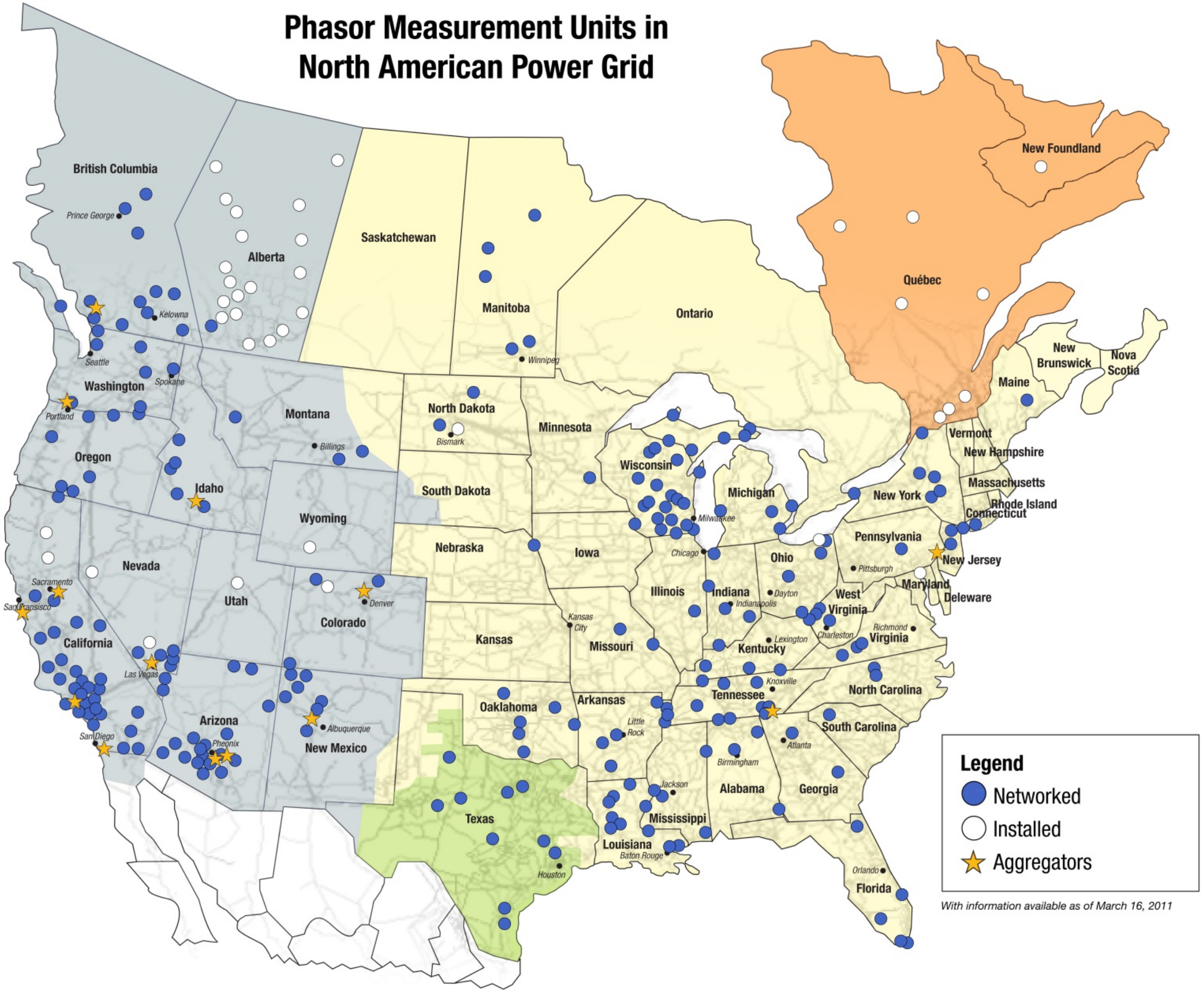
False Data Injection Attack

- False Data Injection attack refers to PMU data being manipulated before reaching the aggregator.
- All of the suggested applications fail in case of False Data Injection attack.
- PMU's are being massively deployed for Smart Grid control and monitoring.

It is crucial to have a mechanism to guarantee reliability of PMU data.

- We will consider the most recent false data injection attack that is capable of deluding the state estimator. Prior to us, no remedy was suggested for it.

PMU Network



Phasor Measurement Units in North American Power Grid

Legend
- ● Networked
- ○ Installed
- ☆ Aggregators

With information available as of March 16, 2011

Gaussian Markov Random Field

- A Gaussian Markov Random Field (GMRF) is a family of jointly Gaussian random variables with distribution that factors in accordance with a given graph.
- Given a graph $G = (V, E)$ with $V = \{1, ..., p\}$ consider a vector of Gaussian random variables $\vec{X} = [X_1, X_2, ..., X_p]^T$ where each node $i \in V$ is associated with a scalar Gaussian random variable X_i.
- A GMRF on *G* has a probability density function

$$f_{\mathbf{X}}(\mathbf{x}) \propto \exp\left[-\frac{1}{2}\mathbf{x}^T \mathbf{J}_G \mathbf{x} + \mathbf{h}^T \mathbf{x}\right]$$

where $\mathbf{J}_G$ is a positive-definite symmetric matrix whose sparsity pattern corresponds to that of the graph

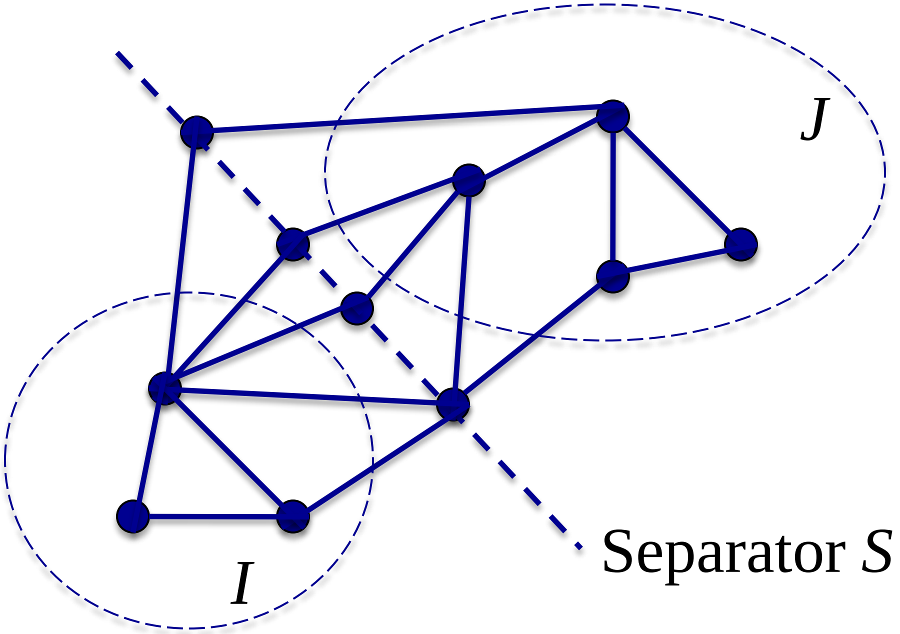
$$J_G(i, j) = 0 \iff (i, j) \notin G.$$

The matrix $J_G = \Sigma^{-1}$ is known as the potential or information matrix.

- For a Gaussian Markov Random Field, local Markov property states that

$$X_i | \mathbf{X}_{-i} = X_i | X_{N(i)}$$

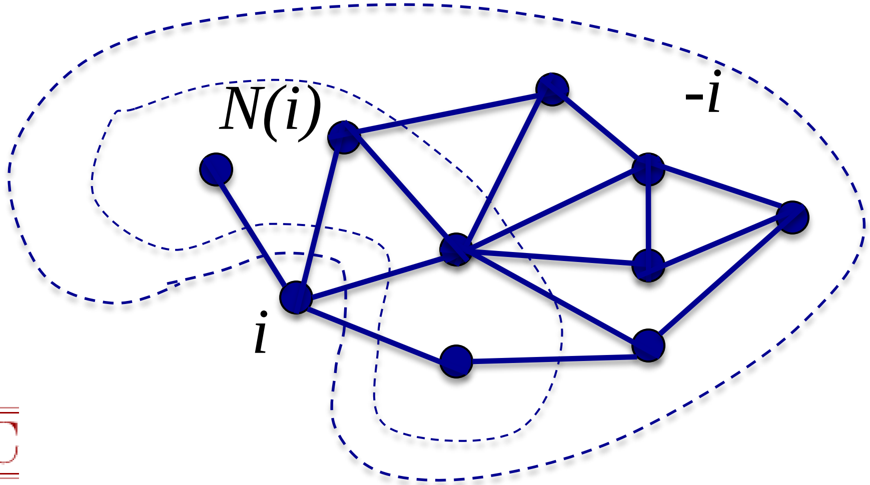
Gaussian Markov Random Field: Separator



$$J_G = \begin{pmatrix} J_{II} & J_{IS} & 0 \\ J_{SI} & J_{SS} & J_{SJ} \\ 0 & J_{JS} & J_{JJ} \end{pmatrix}$$

$$f_X(x) \propto e^{-\frac{1}{2}\left((x_I - r_{IS}x_S)^2 + (x_J - r_{JS}x_S)^2\right)}$$

$$x_I \perp x_J \mid x_S$$



$$E\left(X_i \mid X_{N(i)}\right) = E\left(X_i \mid X_{-i}\right)$$

DC Power Flow Equations

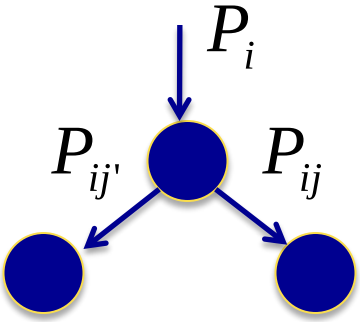
- Often used for analysis of power systems in normal steady-state operations
- Voltages are 1 p.u. and angle differences are small $\sin(\theta_i - \theta_j) \sim \theta_i - \theta_j$
- The power flow on the transmission line connecting bus *i* to bus *j* is given by

$$P_{ij} = b_{ij}(X_i - X_j)$$

X_i and X_j denote the phasor angles at bus *i* and *j* .
b_{ij} denotes the inverse of the line inductive reactance.

- The probabilistic landscape is given by the power injected at the buses:

$$P_i = \sum_{j \in N(i)} P_{ij}$$



- So,

$$X_i = \sum_{j \neq i} c_{ij} X_j + \frac{1}{\sum_{j \neq i} b_{ij}} P_i$$

- where

$$c_{ij} = \frac{b_{ij}}{\sum_{i \neq j} b_{ij}}$$

PMU angle measurements as GMRF

- Aggregated power (generation>0 & load<0) injection at buses are modeled as Gaussian random variables.
- DC power flow is linear; hence

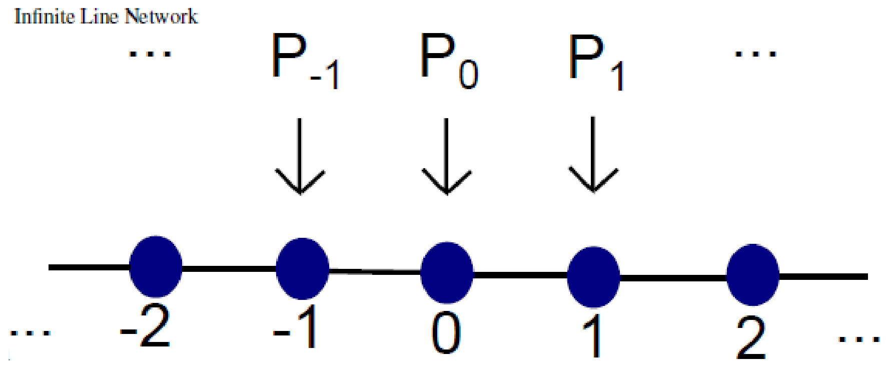
PMU angle measurements can be considered as Gaussian random variables.

- DC power flow shows the GMRF property of PMU angle measurements:

$$X_i = \sum_{j \neq i} c_{ij} X_j + \frac{1}{\sum_{j \neq i} b_{ij}} P_i$$

  - *The first term shows that grid graph neighbors are probabilistic neighbors too.*
  - *What about the second term?*
  - *What is the correct set of neighbors?*

Infinite Chain-Structured Network



$$\begin{bmatrix} \vdots \\ P_{-1} \\ P_0 \\ P_1 \\ \vdots \end{bmatrix} = \begin{bmatrix} \ddots & \ddots & & & \\ \ddots & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & \ddots \\ & & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ X_{-1} \\ X_0 \\ X_1 \\ \vdots \end{bmatrix}$$

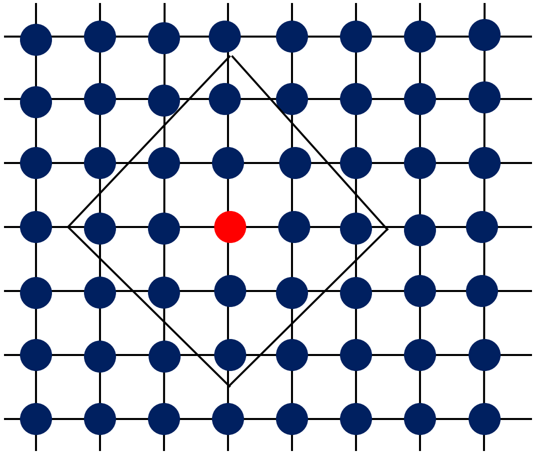
$$\widehat{P}(e^{j\alpha}) = \sum_{k=-\infty}^{+\infty} P_k e^{jk\alpha}, \quad \widehat{X}(e^{j\alpha}) = \sum_{k=-\infty}^{+\infty} X_k e^{jk\alpha}$$

$$\widehat{P}\left(e^{j\alpha}\right) = \widehat{B}(e^{j\alpha})\widehat{X}\,(e^{j\alpha})$$

$$f_X(P) \sim e^{-\frac{1}{2}P^T \Sigma_d^{-1} P} = e^{-\frac{1}{2}X^T B^T \Sigma_d^{-1} BX}$$

$$f_X(x) \propto \exp\left(-\frac{1}{2}X^T \begin{bmatrix} \ddots & \ddots & \ddots & \ddots & \ddots & & & & \\ & & 1 & -4 & 6 & -4 & 1 & & \\ & & & & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix} X \right)$$

Euclidean Lattice structured network



$$\widehat{P}(e^{j\alpha},e^{j\beta})=\widehat{B}(e^{j\alpha},e^{j\beta})\widehat{X}(e^{j\alpha},e^{j\beta})$$

$$\widehat{X}(e^{j\alpha},e^{j\beta})=\sum_{k,l\in\mathbb{Z}}X_{k,l}e^{jk\alpha}e^{jl\beta} \qquad \widehat{P}(e^{j\alpha},e^{j\beta})=\sum_{k,l\in\mathbb{Z}}P_{k,l}e^{jk\alpha}e^{jl\beta}$$

$$f_X(P)\propto e^{-\frac{1}{2}\sum_{k,l\in\mathbb{Z}}P_{kl}^2}=e^{-\frac{1}{2}\frac{1}{2\pi}\oint\oint|\widehat{B}|^2|\widehat{X}|^2 d\alpha d\beta}$$

$\sum_{k,l\in\mathbb{Z}}P_{k,l}^2$ is quadratic in the $X_{k,l}$ variables, but those variables that are multiplied have their indexes within at most a 2-neighbor relationship in the lattice structure.

Model Selection

- **We use Conditional Covariance Test (CCT)[1]:**
  - *Two nodes are connected in the Markov graph iff the Conditional Mutual Information between those measurements is greater than a threshold.*
  - *For Gaussian variables, testing Conditional Mutual Information is equivalent to Conditional Covariance Test.*

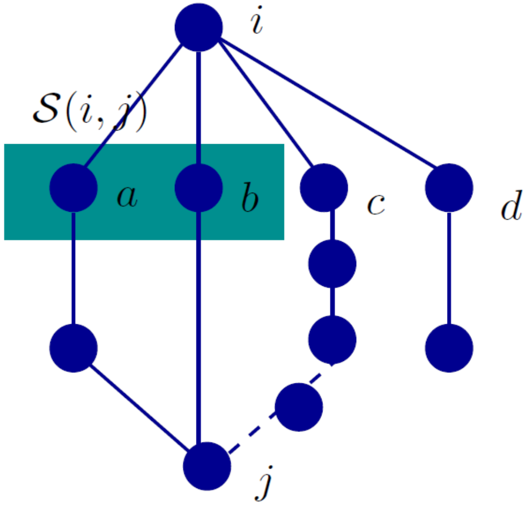
- **In order to have structural consistency, the model should satisfy two important properties: walk-summability and local separation property.**

walk-summability:

$$||\bar{\mathrm{R}}|| \leq \alpha < 1$$

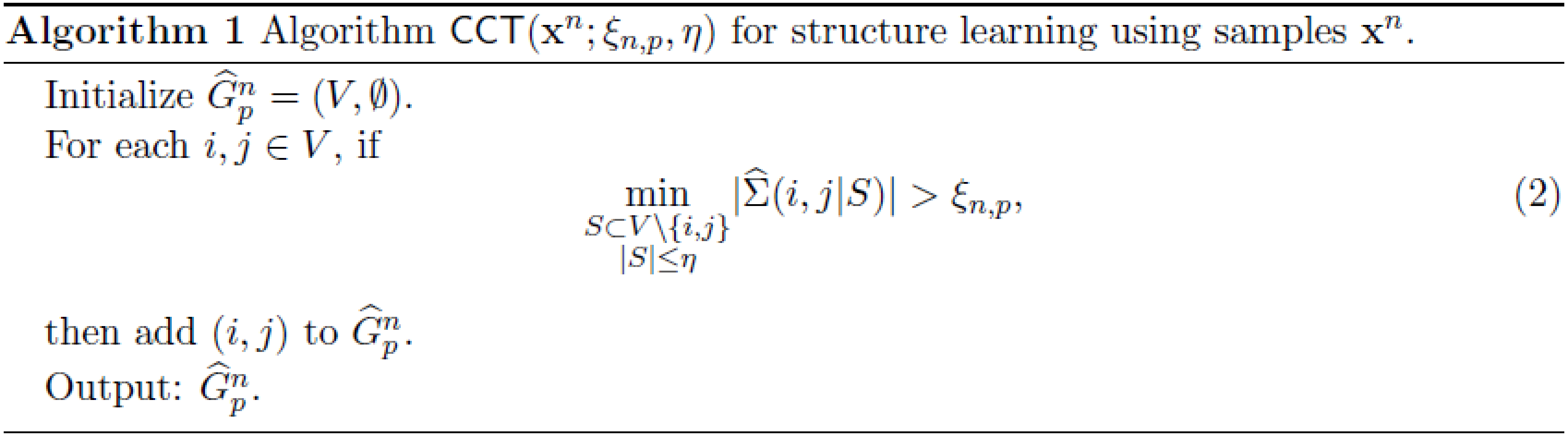
$$r_{ij} = \frac{\Sigma(i,j)\big|\big(V \setminus \{i,j\}\big)}{\sqrt{\Sigma(i,i)\big|\big(V \setminus \{i,j\}\big)}\sqrt{\Sigma(j,j)\big|\big(V \setminus \{i,j\}\big)}} = -\frac{J_{ij}}{\sqrt{J_{ii}}\sqrt{J_{jj}}}$$

local separation property [1]:



[1] A. Anandkumar, V. Tan, F. Huang, and A.S. Willsky. High-dimensional Gaussian graphical model selection: walk summability and local separation criterion. Journal of Machine Learning, June 2012. accepted

Conditional Covariance Test (CCT)



Algorithm 1 Algorithm $\mathsf{CCT}(\mathbf{x}^n; \xi_{n,p}, \eta)$ for structure learning using samples $\mathbf{x}^n$.

```
Initialize $\widehat{G}_p^n = (V, \emptyset)$.
For each $i, j \in V$, if
    $\min_{\substack{S \subset V \setminus \{i,j\} \\ |S| \leq \eta}} |\widehat{\Sigma}(i, j|S)| > \xi_{n,p},$    (2)
then add $(i, j)$ to $\widehat{G}_p^n$.
Output: $\widehat{G}_p^n$.
```

Neighboring Relationship

- Grid structure is **walk-summable.** ($\Leftarrow$ It is of bounded degree.)
- Under walk-summability the effect of faraway nodes on covariance decays with the distance and the error in approximating the covariance by local neighboring relationship decays exponentially with the distance [1].

- By correct tuning of threshold and enough number of samples, we expect the output of CCT method to follow the grid structure.

Stealthy Deception False Data Injection Attack

- The most recent and most realistically scary false data injection attack on the power grid is the stealthy deception attack [2]:

$$z = h(x) + \varepsilon$$

  - *z : measurement vector, x:state vector, h:measurement function,$^{\varepsilon}$ measurement error*

- The goal of a stealthy deception attacker is to compromise the measurements available to the State Estimator (SE) as

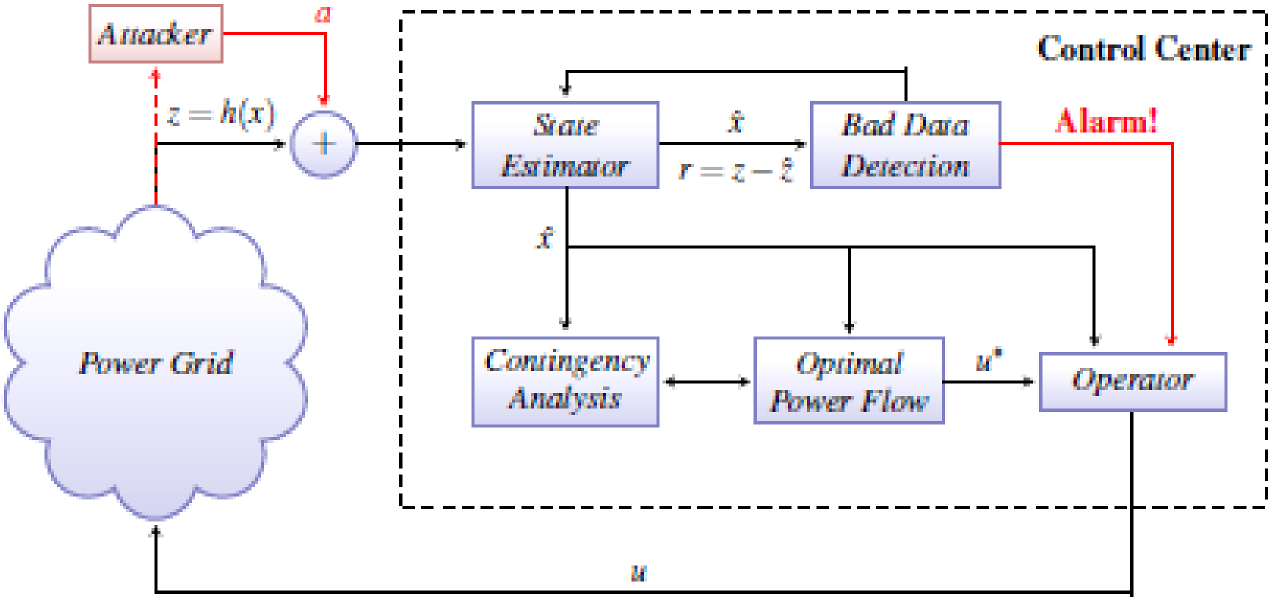
$$z^a = z + a$$

  - *a is the attack vector and is designed in a way that the difference between real measurement z and attacked measurement is the desired value*
  - *a is designed such that attack cannot be detected by Bad Data Detection in State Estimator*
  - *Such an a is proven to be achievable via* $a \in \mathrm{Im}(H)$ $H = \left(\partial f_i(x) / \partial x_j\right)_{i,j}$

[2] A. Teixeira, G. Dan, H. Sandberg, and K H. Johansson. A Cyber Security Study of a SCADA Energy Management System: Stealthy Deception Attacks on the State Estimator. In IFAC World Congress, September 2011.

False Data Injection Attack

- This attack is valid only if performed locally.
- Attack is performed under DC power flow assumption.



The state estimator under a cyber attack [2]

Attack Detection

- **DC power flow assumption**
  - $x=X$
  - $H = H_{P\theta}$
- $a \in \mathrm{Im}(H)$
  - $z^a = z + a = H(X+d)$ ⇒

$$Hd = z^a - HX = a$$

$$H_{ij} = -b_{ij} \text{ for } i \neq j$$

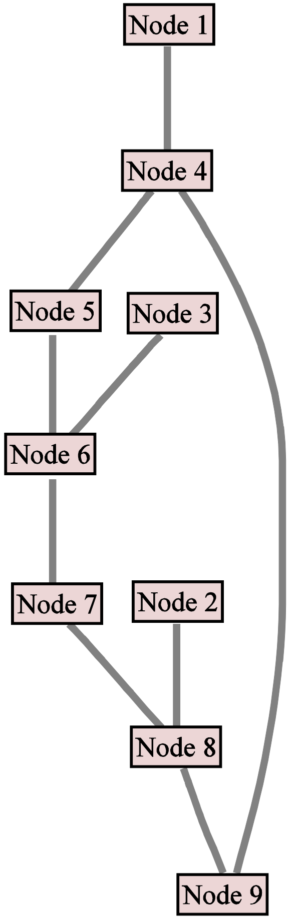
$$H_{ii} = \sum_{i \neq j} b_{ij}$$

- Numerical analysis on above equation shows that

  the Markov graph of an attacked system lacks at least one link from the grid graph.
- We use this to trigger the alarm.
- It should be emphasized that the attack assumes the knowledge of the system's bus-branch model. So the attacker is equipped with a wealth of information. Yet, we can detect such an attack by a *sophisticated* player with our method.

Simulation

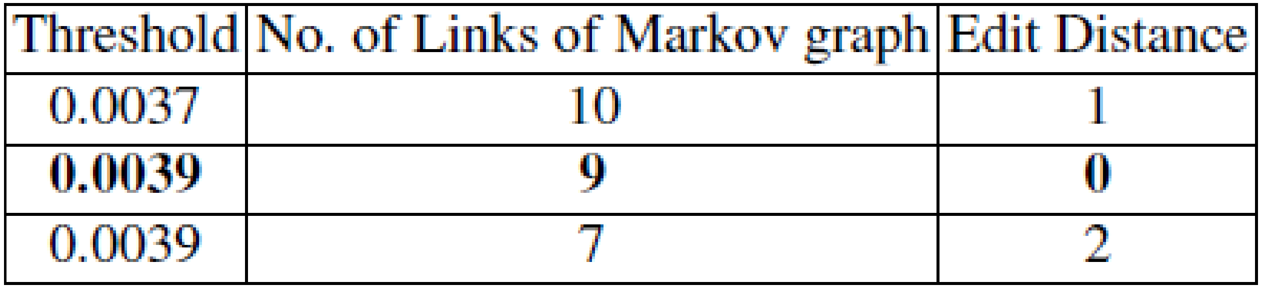
- **We considered a 9-node grid suggested by Zimmerman et al. [3]**



[3] C. E. Murillo-Snchez R. D. Zimmerman and R. J. Thomas. MATPOWER steady-state operations, planning and analysis tools for power systems research and education. Power Systems, IEEE Transactions on,26(1):12–19, Feb. 2011

Simulation (cont.)

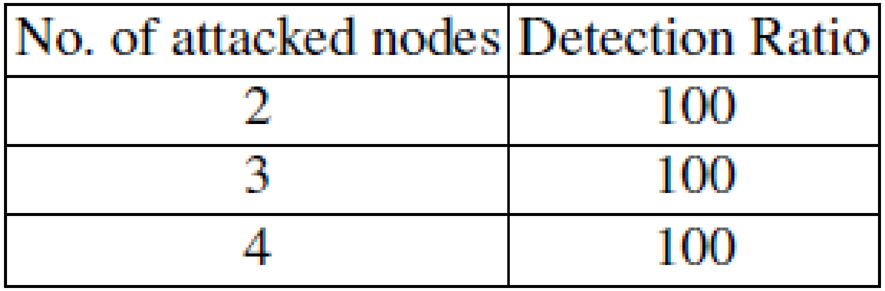
- First, we fed the system with Gaussian demand and simulated the power grid. We used MATPOWER for solving the DC power flow equations for various demand and used the resulting angle measurements as the input to CCT algorithm.
- We used YALMIP and SDPT3 to perform CCT.
- With the right choice of parameters and threshold, and enough *un-compromised* measurements, the Markov graph follows the grid structure.
- The edit distance between the Markov graph and the grid graph that is used to lead us to the correct threshold:



Threshold	No. of Links of Markov graph	Edit Distance
0.0037	10	1
0.0039	**9**	**0**
0.0039	7	2

Attack Simulation

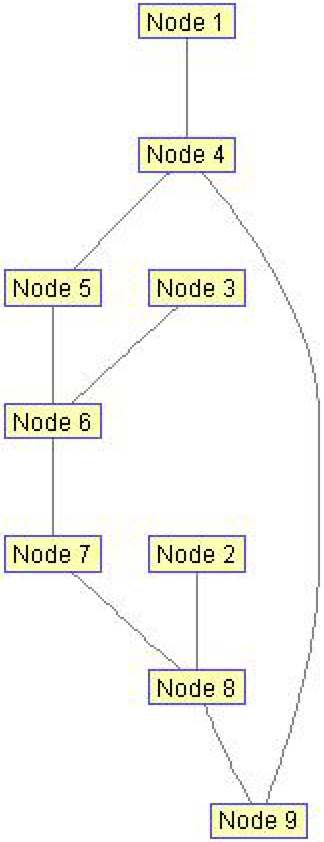
- We introduced the stealthy deception attack to the system.
- We investigated the cases where 2, 3 or 4 nodes were under attack.
- For each case, we simulated all possible attack combinations.
- In all attack scenarios, the Markov graph of tampered PMU measurements lacked at least one link that was present in grid graph, a discrepancy that triggered the alarm.

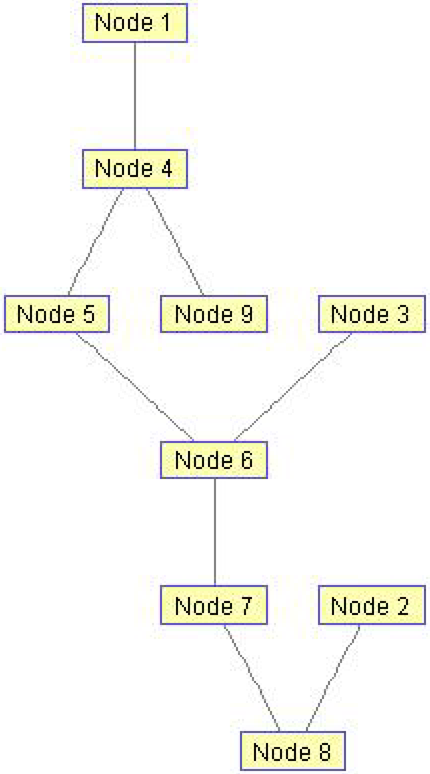


No. of attacked nodes	Detection Ratio
2	100
3	100
4	100

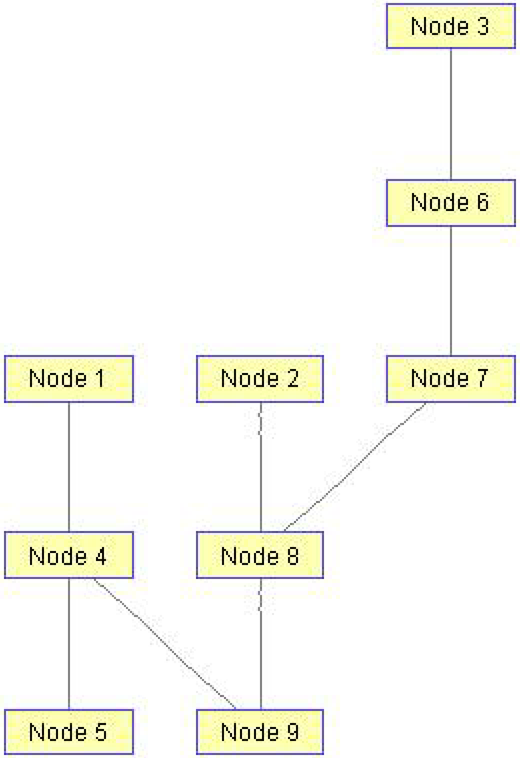
Attack Simulation

No attack





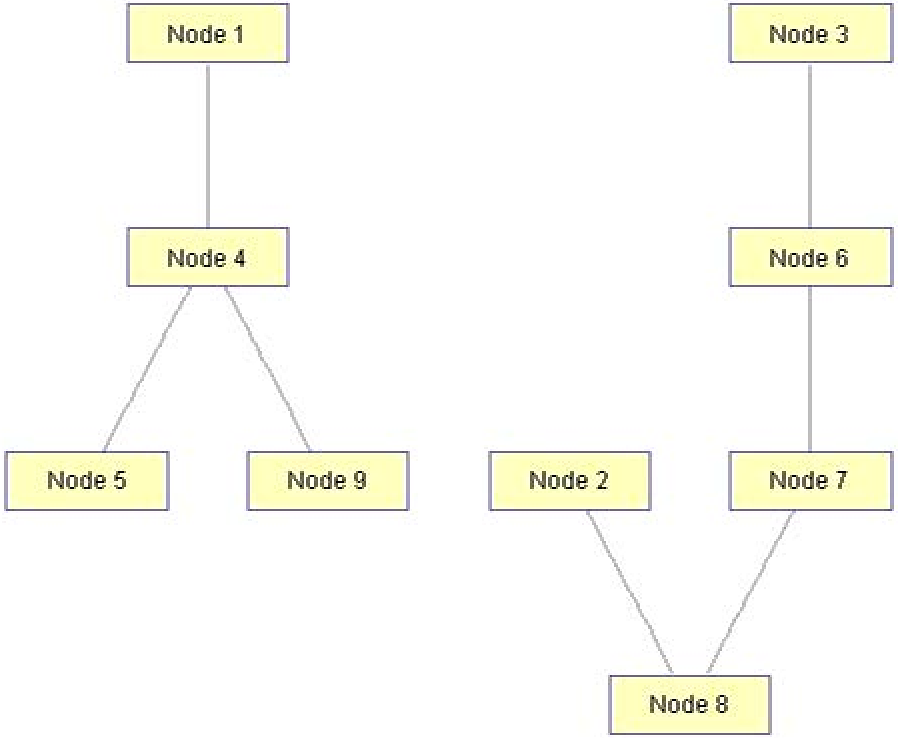
One case:
Nodes 3 and 9
under attack



One case:
Nodes 2 and 5
under attack

Attack Simulation

One case:
Nodes 6 and 9 under attack



Conclusion

- It is crucial to assure PMU data reliability
- Statistical structure learning of PMU angle measurements
- Markov graph of bus angle measurements follows grid topology.
- Discrepancy triggers the alarm that the system is under false data injection attack.
- This is the first remedy for the strong false data injection attack mentioned.
- We would like to extend this work to bigger grid networks.

References

1. A. Anandkumar, V. Tan, F. Huang, and A.S. Willsky. High-dimensional Gaussian graphical model selection: walk summability and local separation criterion. Journal of Machine Learning, June 2012. accepted.
2. A. Teixeira, G. Dan, H. Sandberg, and K H. Johansson. A Cyber Security Study of a SCADA Energy Management System: Stealthy Deception Attacks on the State Estimator. In IFAC World Congress, September 2011.
3. C. E. Murillo-Snchez R. D. Zimmerman and R. J. Thomas. MATPOWER steady-state operations, planning and analysis tools for power systems research and education. Power Systems, IEEE Transactions on,26(1):12–19, Feb. 2011