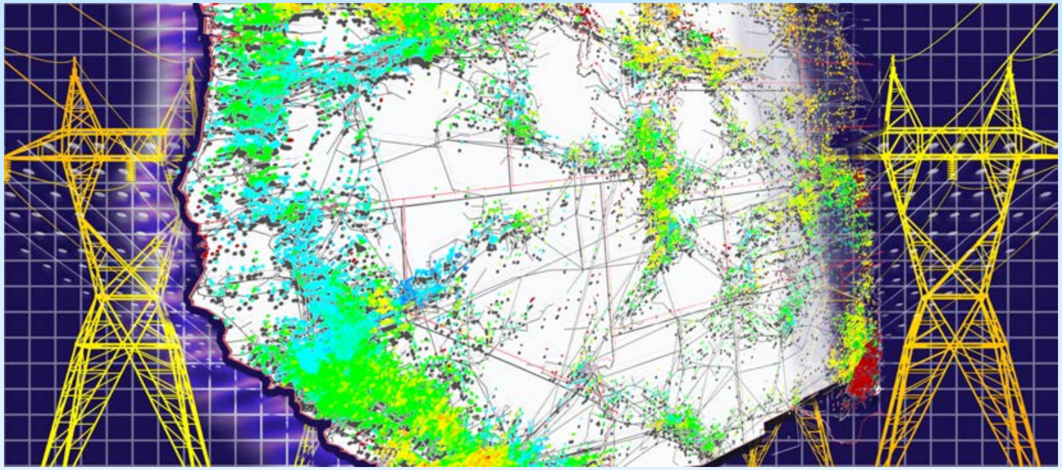
COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

wim.net

State Estimation following Cyber Attacks on the Power Grid



Saleh Soltan, Mihalis Yannakakis, Gil Zussman

Electrical Engineering and Computer Science

Columbia University, New York, NY

Security and Resilience

Physical Attacks/Failures

Cyber Attacks/Failures

Control

Data
Power

Power Grid
Physical Infrastructure

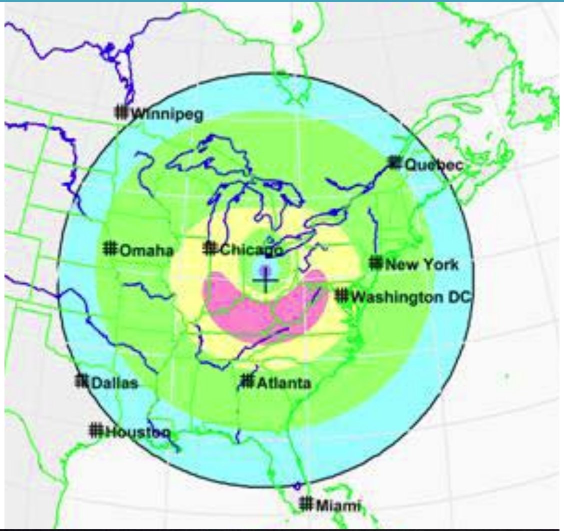
Communication networks

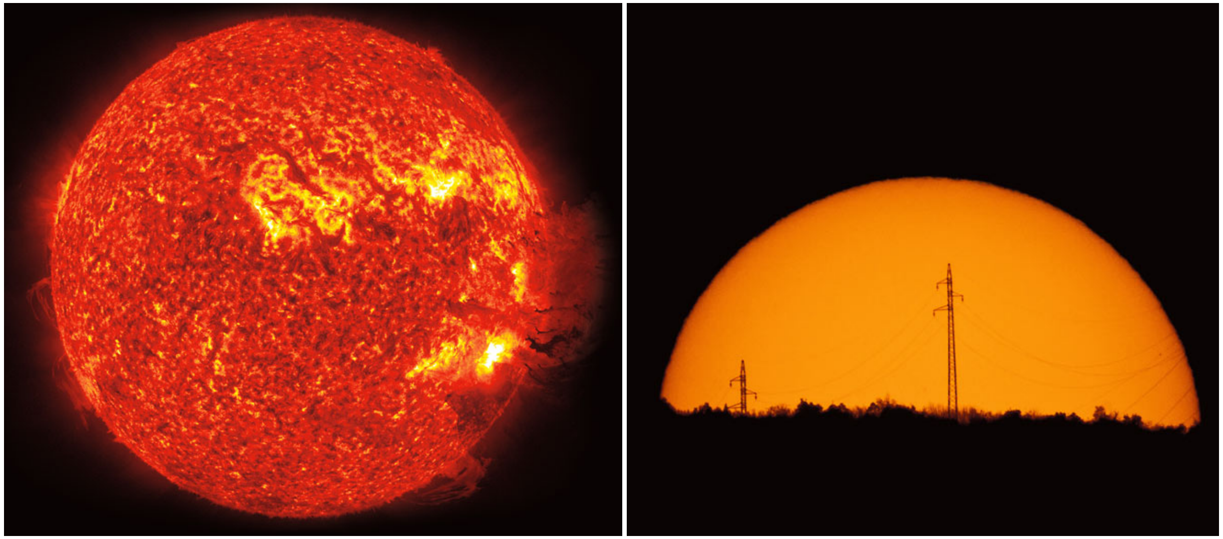
Supervisory Control and Data
Acquisition (SCADA) system

Large Scale Physical Attacks/Failures

- EMP (Electromagnetic Pulse) attack
- Solar Flares

DEFENSE THREAT REDUCTION AGENCY
UNITED STATES OF AMERICA

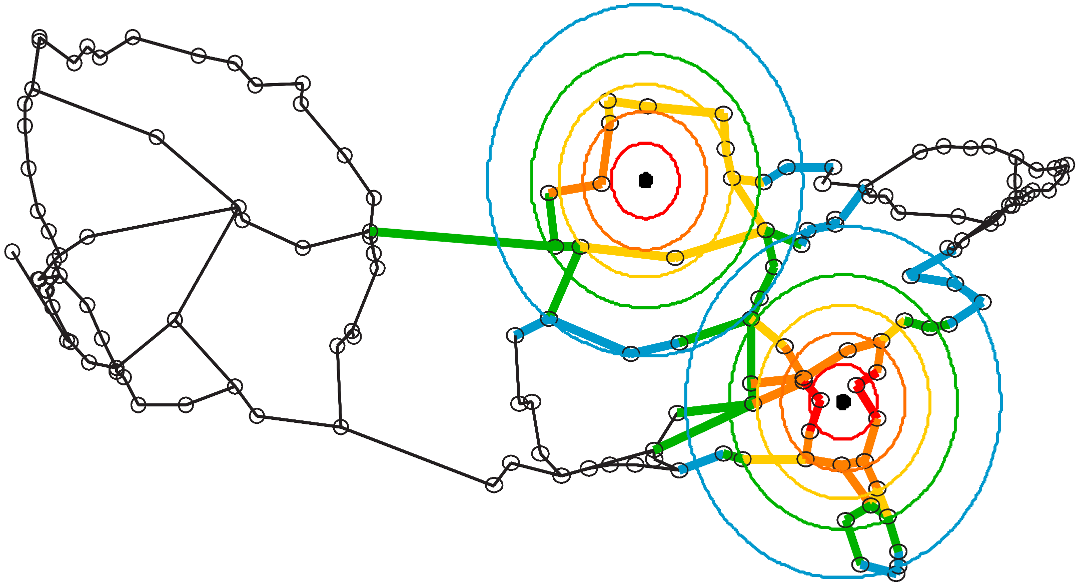



Photos from National Geographic

- Other natural disasters

Identifying Communications Network’s Most Vulnerable Parts

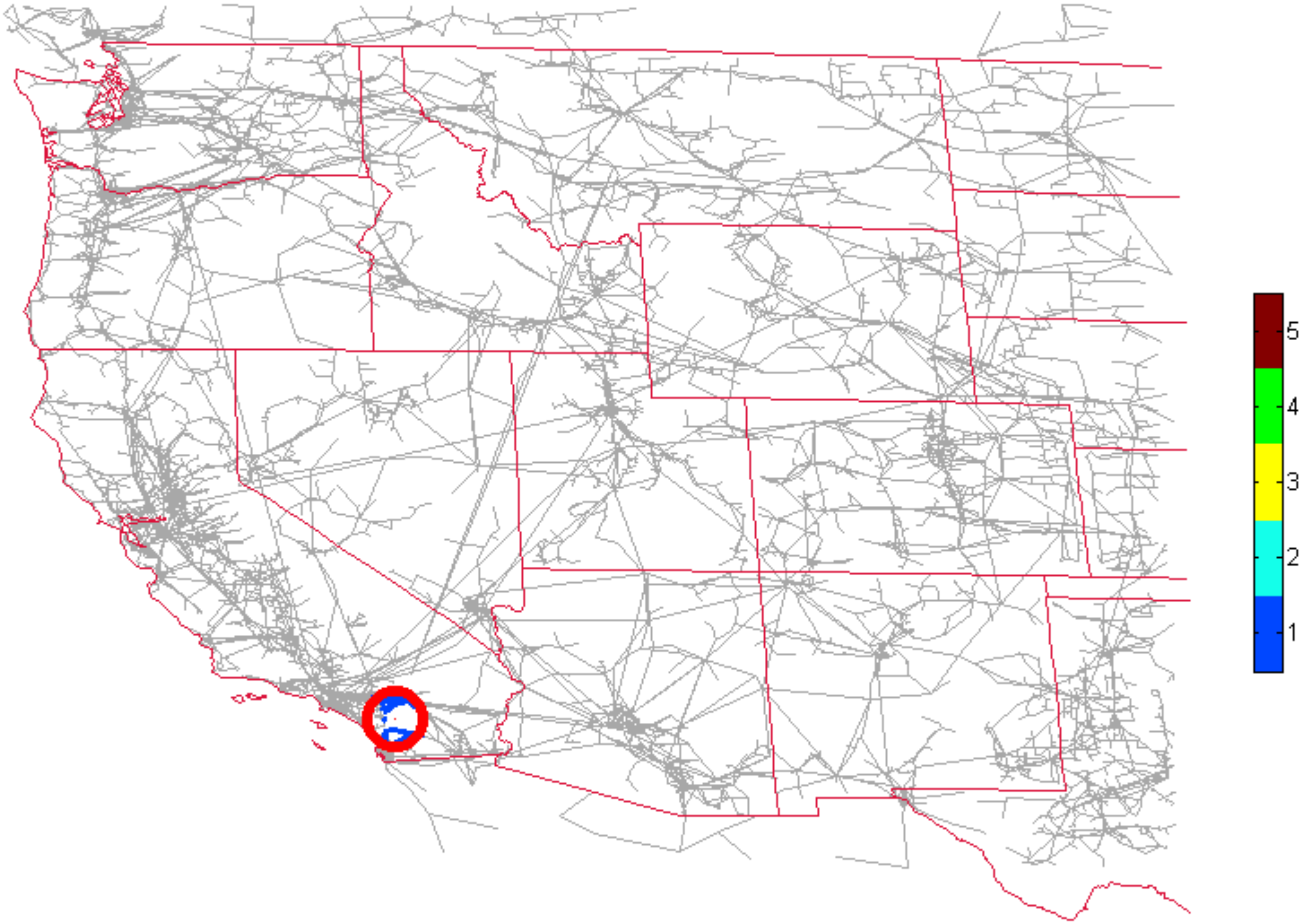
- Cuts with probabilistic properties
  - Distance from the attack’s epicenter
  - The topography of the surrounding area
  - The component’s specifications
- A number of simultaneous attacks
- Take into account protection and restoration



- Use computational geometric tools for complexity reduction

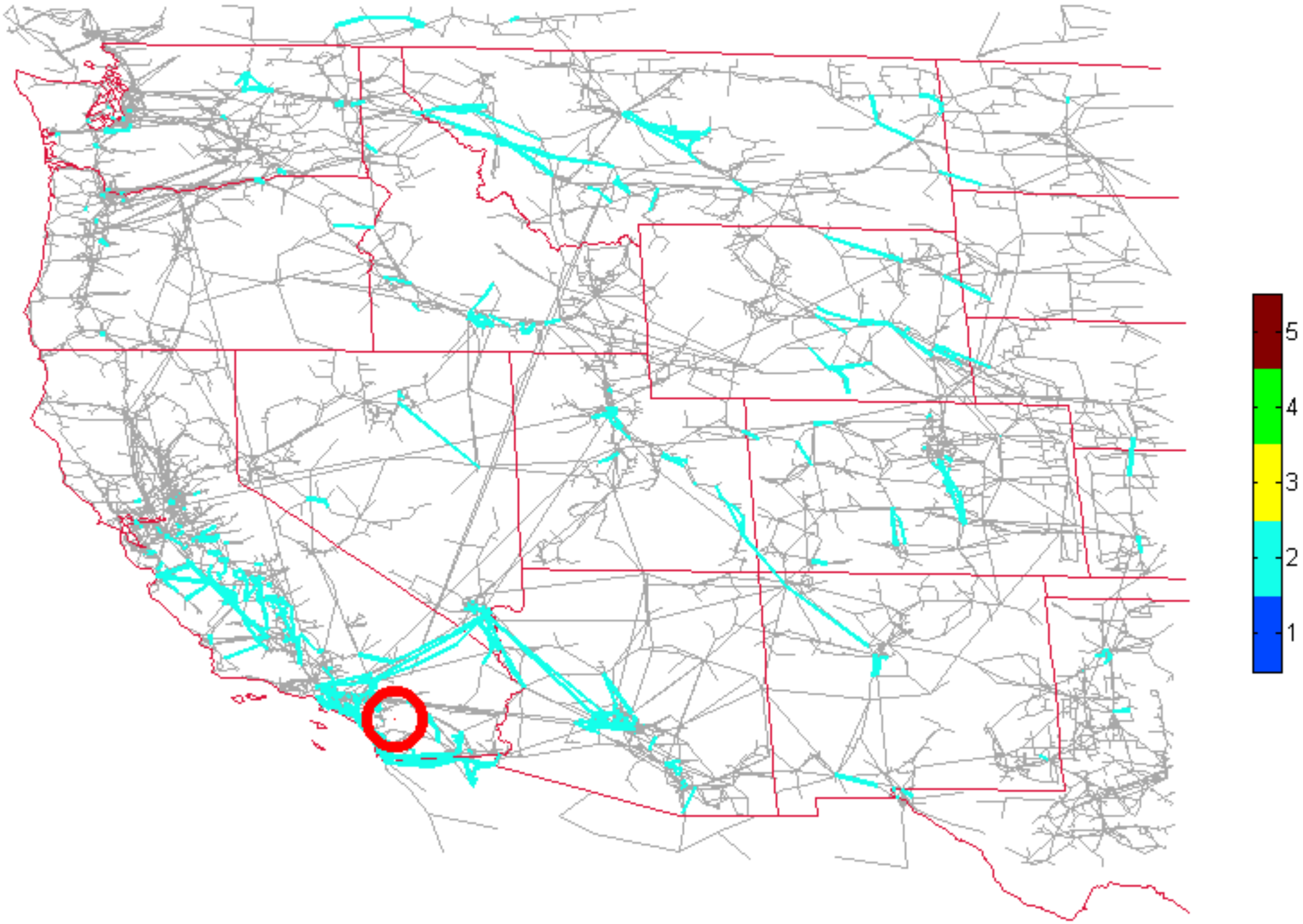
P. K. Agarwal, A. Efrat, S. K. Ganjugunte, D. Hay, S. Sankararaman, and G. Zussman, “The resilience of WDM networks to probabilistic geographical failures,” IEEE/ACM Transactions on Networking, vol. 21, no. 5, pp. 1525–1538, Oct. 2013.

Power Networks – Vulnerability and Cascade Analysis

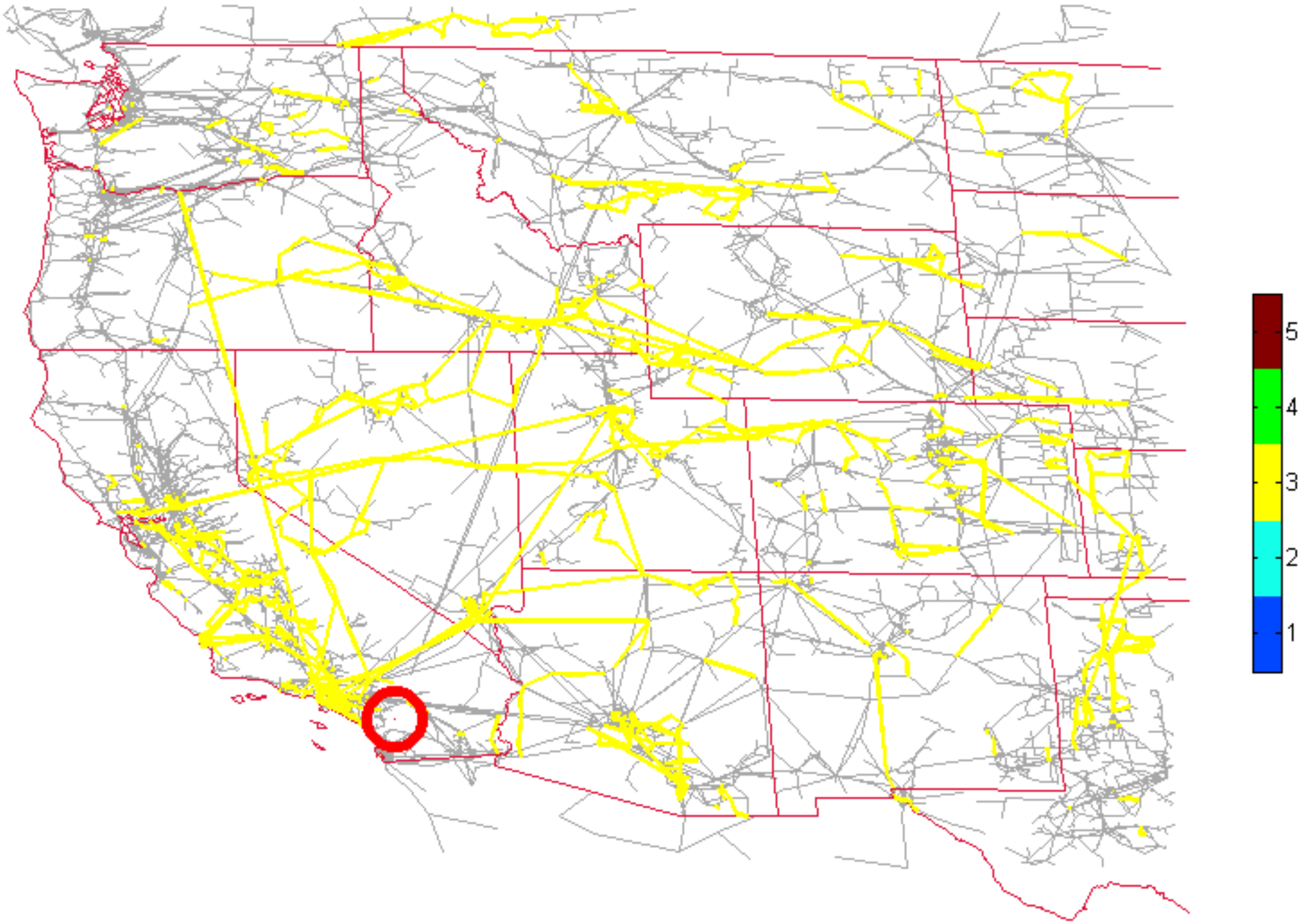


A. Bernstein, D. Bienstock, D. Hay, M. Uzunoglu, and G. Zussman, "Power grid vulnerability to geographically correlated failures - Analysis and control implications," in *Proc. IEEE INFOCOM'14*, 2014.

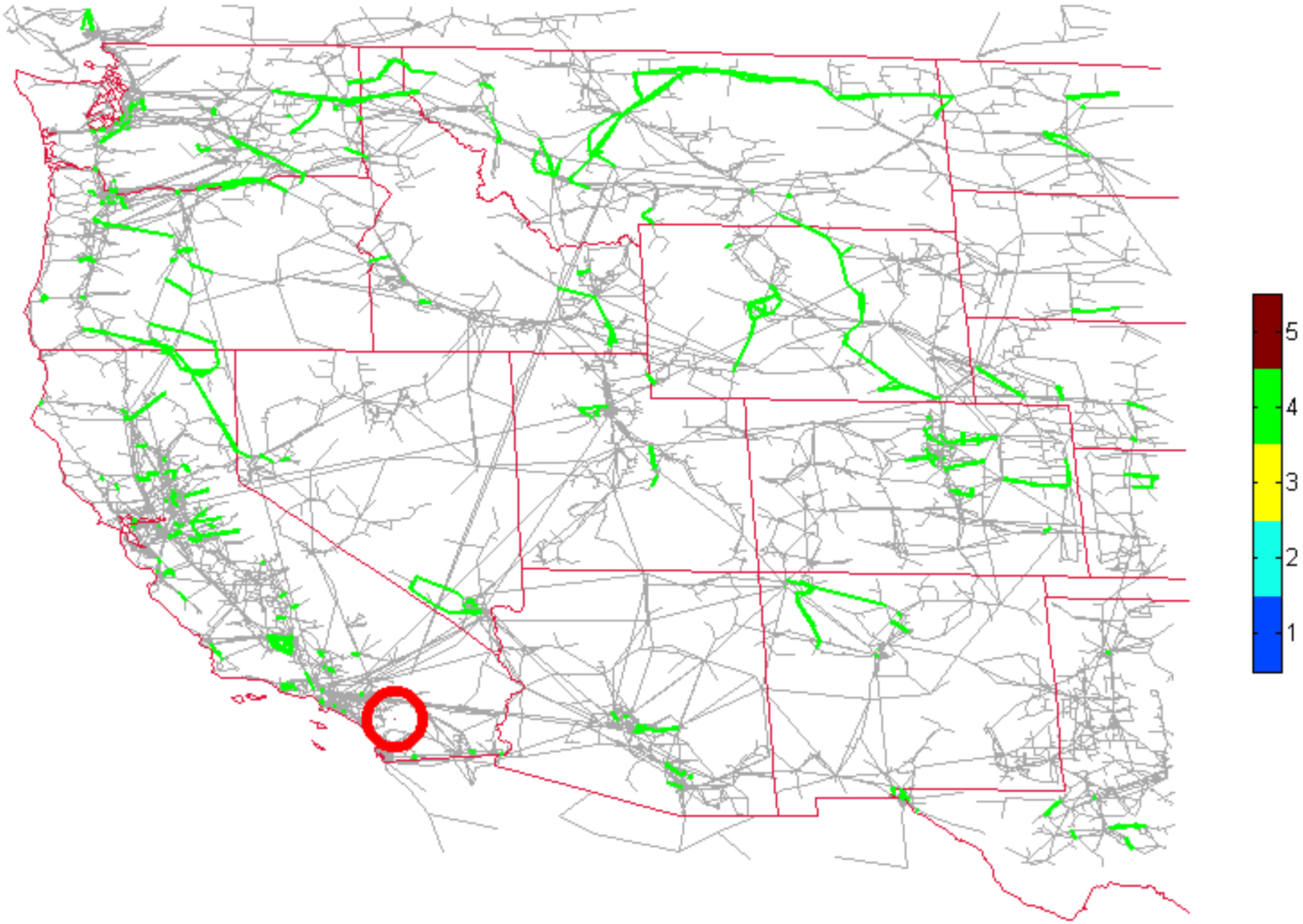
Power Networks – Vulnerability and Cascade Analysis



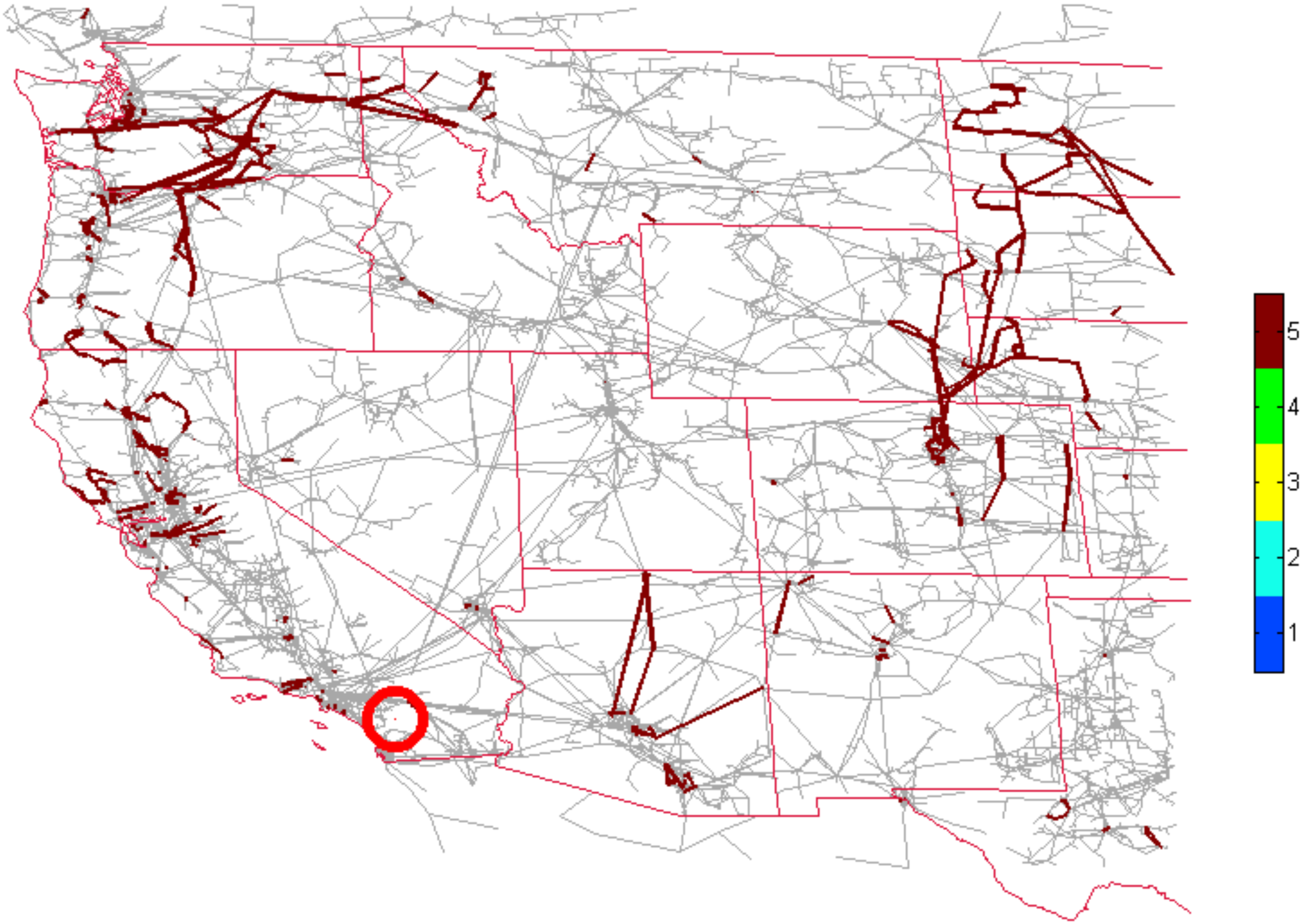
Power Networks – Vulnerability and Cascade Analysis



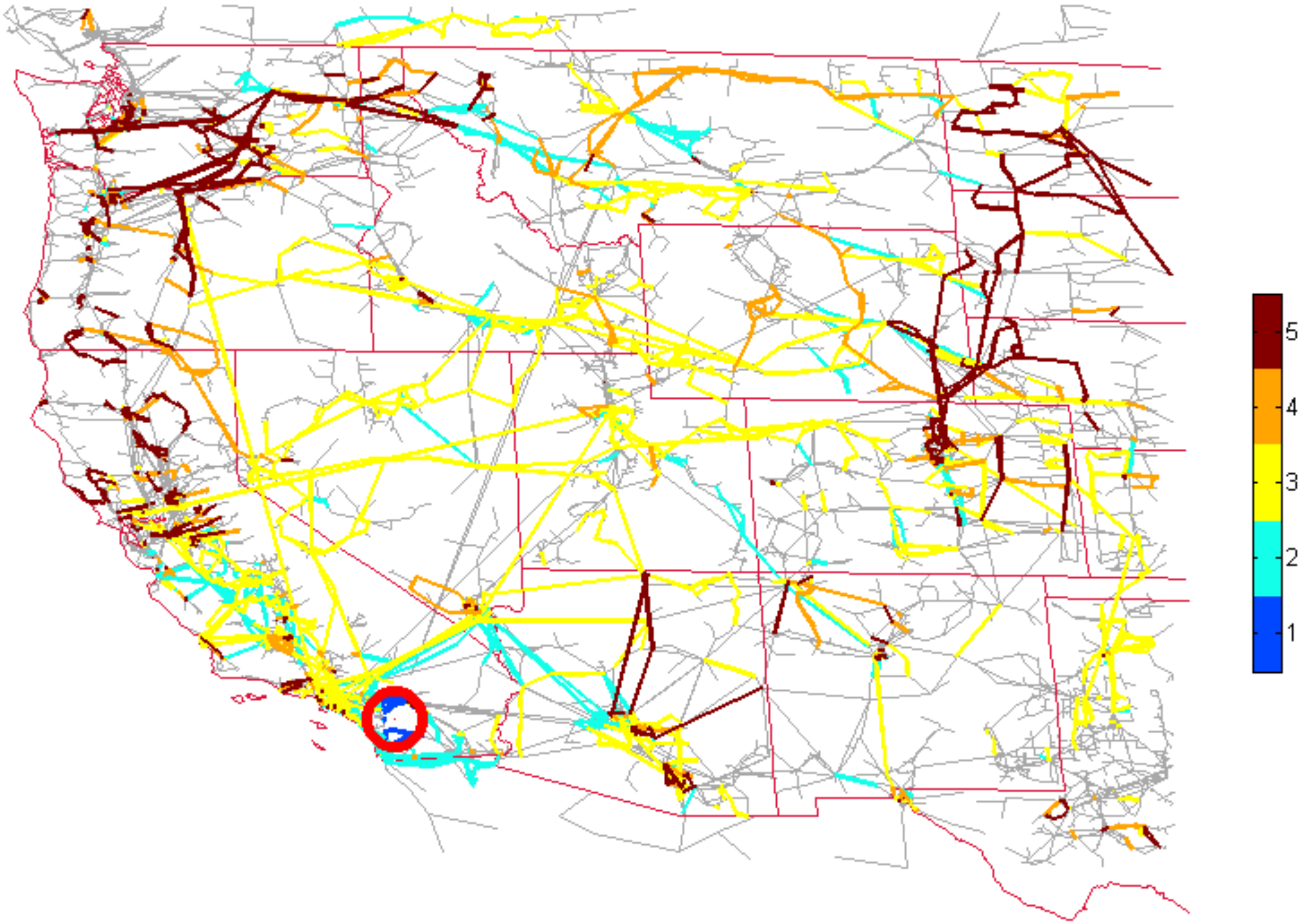
Power Networks – Vulnerability and Cascade Analysis



Power Networks – Vulnerability and Cascade Analysis



Power Networks – Vulnerability and Cascade Analysis



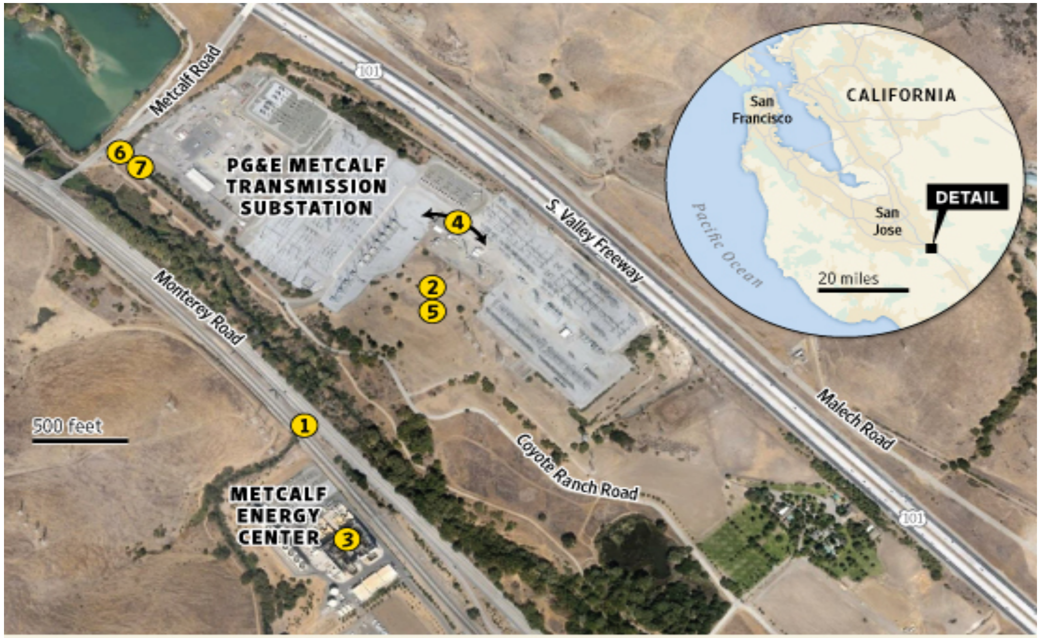
N-Resilient, Factor of Safety K = **1.2** → **Yield = 0.33**

For $(N\text{-}1)$-Resilient → **Yield = 0.35**   For K = **2** → **Yield = 0.7**

(Yield - the fraction of the demand which is satisfied at the end of the cascade)

Power Grid Attack in San Jose (Apr. 2014)

- “A sniper attack in April 2014 that knocked out an electrical substation near San Jose, Calif., has raised fears that the country's power grid is vulnerable to terrorism. ” –The Wall Street Journal



Shots in the Dark

A look at the April 16 attack on PG&E’s Metcalf Transmission Substation

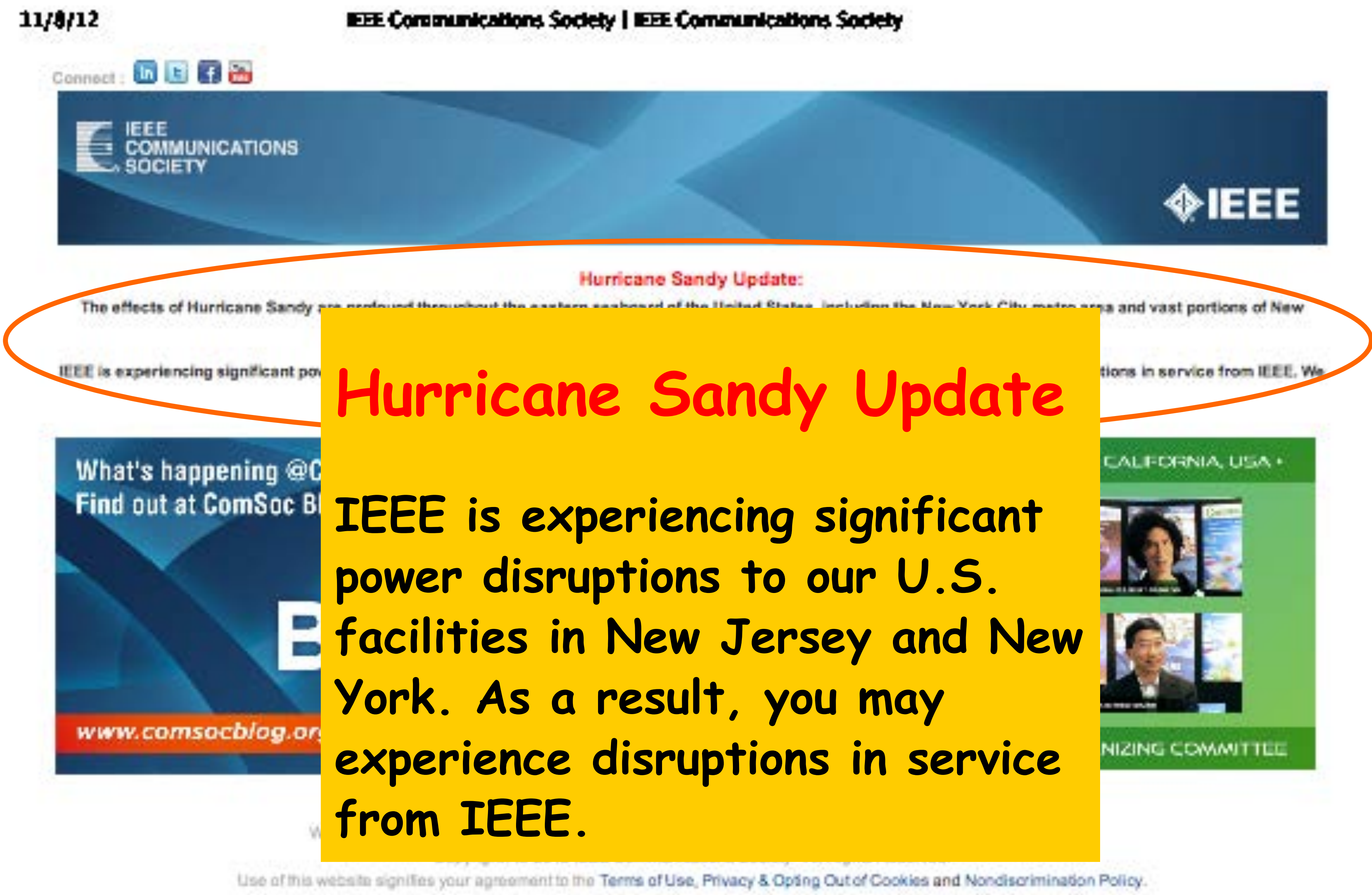
1	2	3	4	5	6	7
12:58 a.m., 1:07 a.m. Attackers cut telephone cables	**1:31 a.m.** Attackers open fire on substation	**1:41 a.m.** First 911 call from power plant operator	**1:45 a.m.** Transformers all over the substation start crashing	**1:50 a.m.** Attack ends and gunmen leave	**1:51 a.m.** Police arrive but can’t enter the locked substation	**3:15 a.m.** Utility electrician arrives

Sources: PG&E; Santa Clara County Sheriff’s Dept.; California Independent System Operator; California Public Utilities Commission; Google (image)
The Wall Street Journal

Cyber Attack in Ukraine (Dec. 2015)

- Unplugged 225,000 people from the Ukrainian electricity grid
  - Steal credentials for accessing the SCADA system, *before June 2015*
  - Explore of SCADA system and plan attack, *June-Dec. 2015*
  - Remotely operate circuit breakers, *day of attack*
  - Phone jamming attacks keeps operators unaware, *day of attack*

Interdependencies (Nov. 2012)



11/8/12

IEEE Communications Society | IEEE Communications Society

Connect :

Hurricane Sandy Update:

The effects of Hurricane Sandy are profound throughout the eastern seaboard of the United States, including the New York City metro area and vast portions of New

IEEE is experiencing significant pow … tions in service from IEEE. We

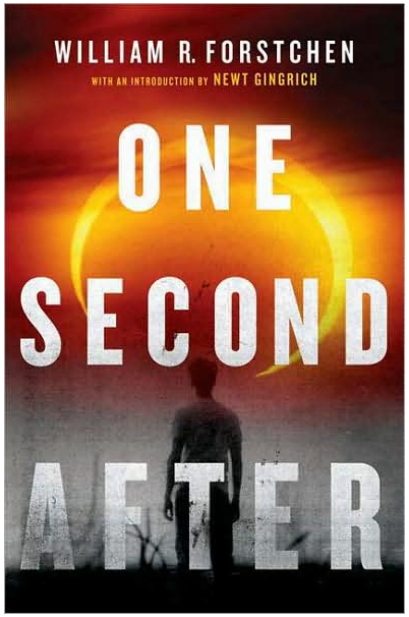
Use of this website signifies your agreement to the Terms of Use, Privacy & Opting Out of Cookies and Nondiscrimination Policy.

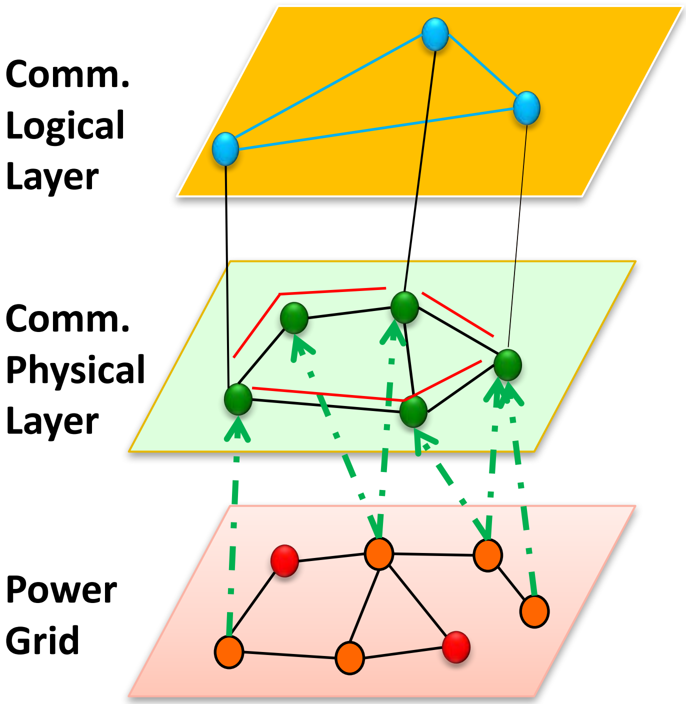
Hurricane Sandy Update

IEEE is experiencing significant power disruptions to our U.S. facilities in New Jersey and New York. As a result, you may experience disruptions in service from IEEE.

Interdependencies

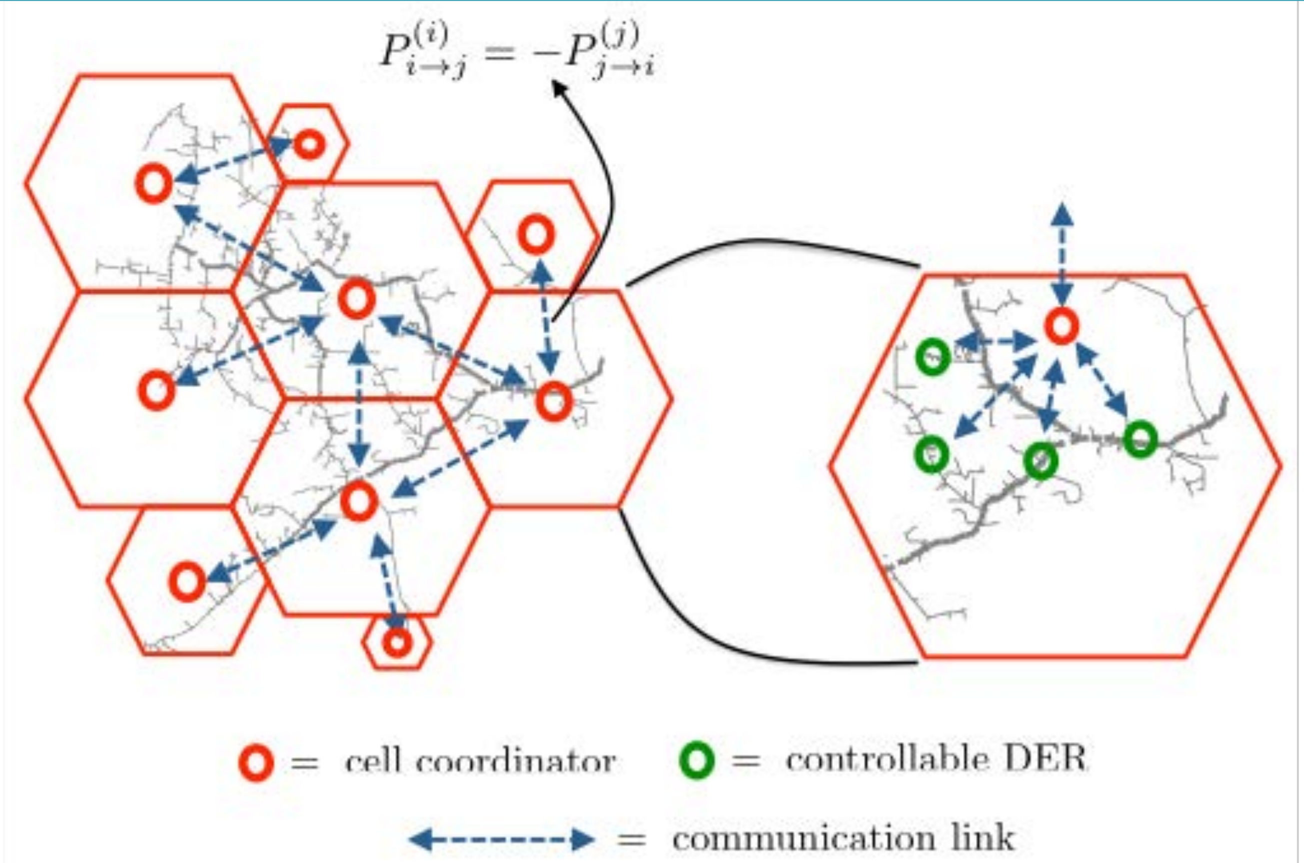
- FCC Workshop workshop on network resiliency (2013) https://edas.info/web/fcc-nr2013/program.html
- Report of the Commission to Assess the threat to the United States from Electromagnetic Pulse (EMP) Attack, 2008
- Modeling:
  - Simple cascades
  - Power control with limited communications/ imperfect information
  - Power loss (eventual) impact on communications





S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, S. Havlin, "Catastrophic cascade of failures in interdependent networks," *Nature* 464, 1025-1028 (15 April 2010) -> Google Scholar

Back to Autonomous Energy Grids


$P^{(i)}_{i\to j} = -P^{(j)}_{j\to i}$
○ = cell coordinator
○ = controllable DER
←-------→ = communication link

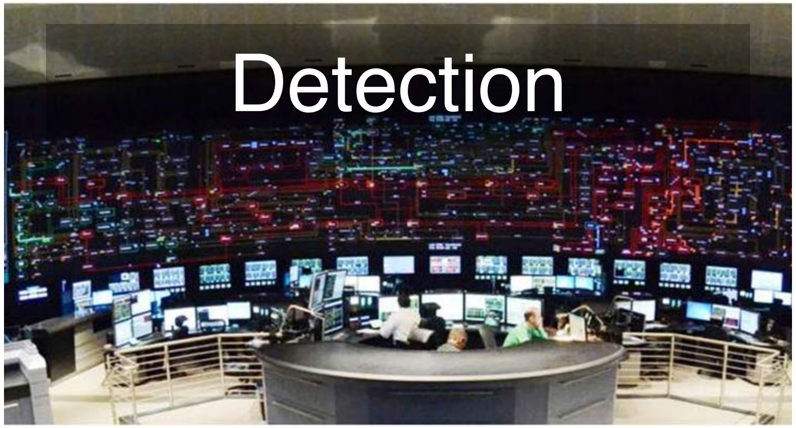
- Communication is crucial for coordinated distributed control
- Control without communication or with imperfect information – suboptimal?
- Are the cells resilient without communication?

B. Kroposki, E. Dall'Anese, A. Bernstein, Y. Zhang, B.-M. Hodge, "Autonomous Energy Grids ," Proc. HICSS, 2018

Ongoing Research

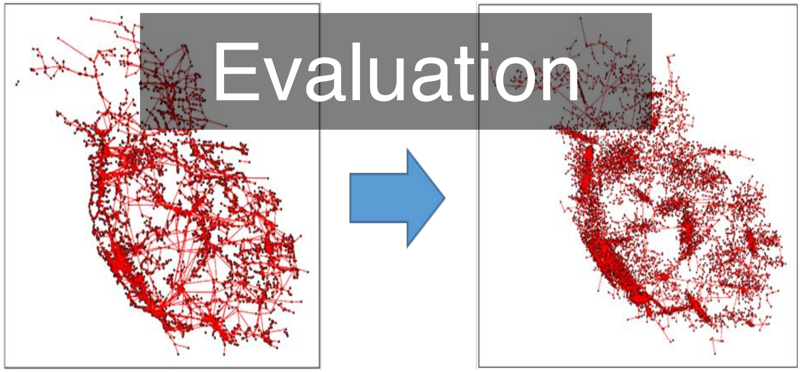
Prediction

Cascades and Interdependencies



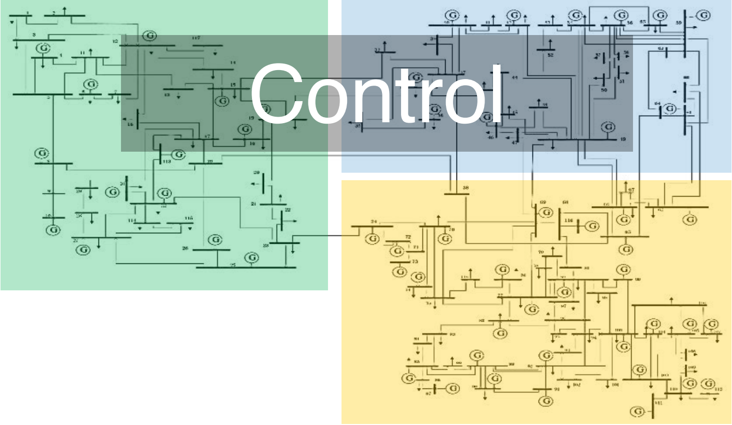
Detection

Cyber Attacks



Evaluation

Synthetic Power Grids

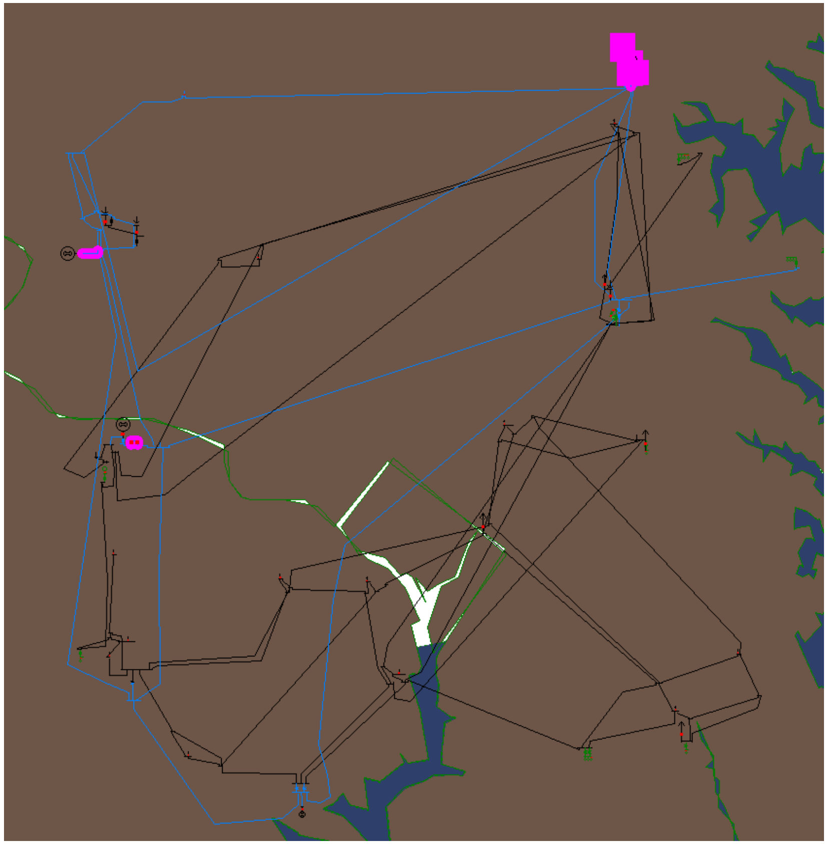


Control

Islanding, Limited Information

DARPA RADICS Program

- RADICS - Rapid Attack Detection, Isolation and Characterization Systems
- Respond to cyber-attacks on U.S. critical infrastructure
- TA1 - Use grid measurements to identify attacks
- Scenarios in “Exercise #1” (May 2016):
  - Disconnect line + false data injection
  - Disconnect line and load + false data injection
  - Large scale cascade



80-Bus Transmission system overlaid on the Washington DC area

State Estimation after a Cyber-Physical Attack - Outline

- State estimation under the DC model
- State estimation in the presence of measurement noise and uncertainty
- State estimation under the AC model
- Attack identification when the affected area is unknown

* focus on transmission

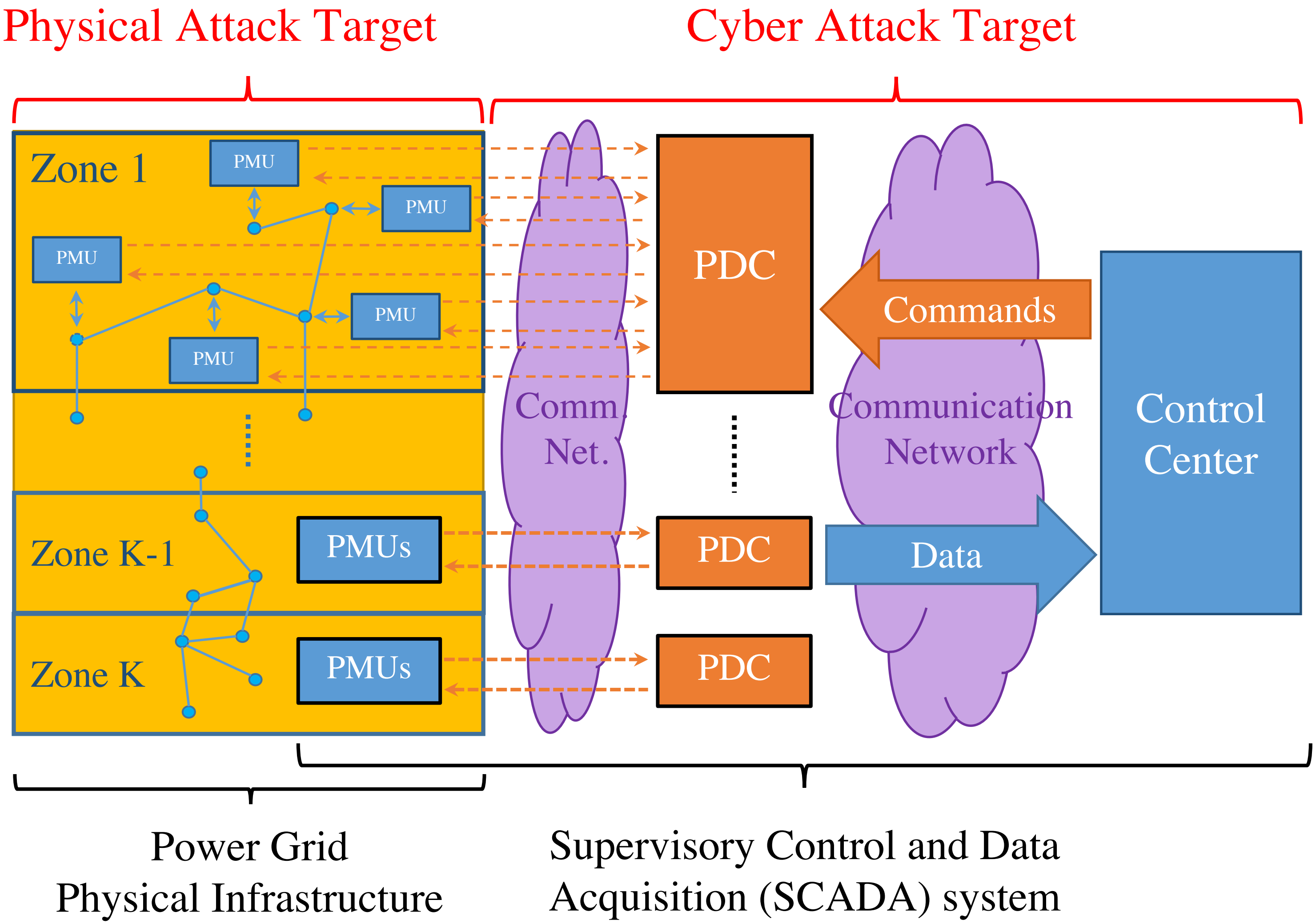
[1] Saleh Soltan, Mihalis Yannakakis, Gil Zussman, “Joint Cyber and Physical Attacks on Power Grids: Graph Theoretical Approaches for Information Recovery,” IEEE Transactions on Control of Network Systems (to appear), 2017.

[2] Saleh Soltan and Gil Zussman, “Power Grid State Estimation after a Cyber-Physical Attack under the AC Power Flow Model,” Proc. IEEE PES-GM’17, 2017.

[3] Saleh Soltan, Mihalis Yannakakis, Gil Zussman, “EXPOSE the Line Failures following a Cyber-Physical Attack on the Power Grid ,” in preparation.

[4] Saleh Soltan, Mihalis Yannakakis, Gil Zussman, “REACT to Cyber Attacks on Power Grids,” submitted.

Simplistic view of a Power Grids

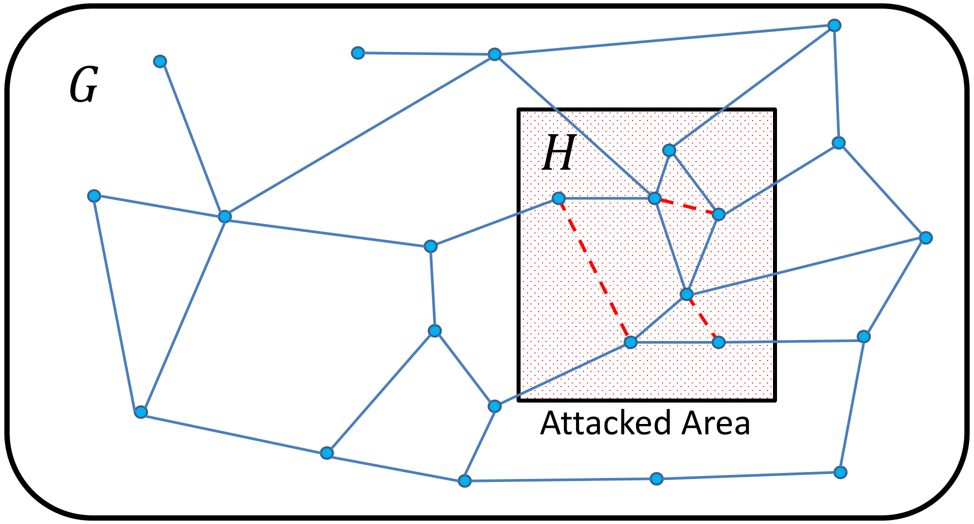


PMU: Phasor Measurement Unit

PDC: Phasor Data Concentrators

Simple Attack Model

- An adversary attacks an area by
  - Disconnecting some lines within the attacked area (physical attack)
  - Disallowing the information from the measurement devices within the area to reach the control center (cyber attack)
- Assume we know H
- A cyber only or physical only attack may result in a similar situation



Power Flow Equations - DC Approximation

- Represent the grid by a connected graph $G = (V, E)$
- The DC power flow is a solution $(\vec{f}, \vec{\theta})$ of:

$$A\vec{\theta} = \vec{p}$$
$$YD^T\,\vec{\theta} = \vec{f}$$

$D \in \{-1,0,1\}^{n\times m}$: the **incidence matrix** of the grid:

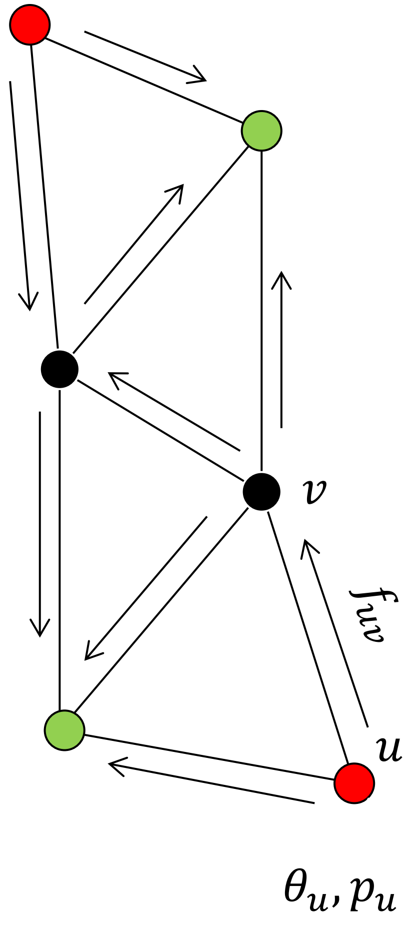
$$d_{ij} = \begin{cases} 0, & \text{if } e_j \text{ is not incident to node } i, \\ 1, & \text{if } e_j \text{ is coming out of node } i, \\ -1, & \text{if } e_j \text{ is going into of node } i, \end{cases}$$

$Y \in \mathbb{R}^{m\times m}$: the diagonal matrix of admittance values,

and $A = DYD^t$: the **admittance matrix** of the grid

θ_u: Phase Angle

x_{uv}: Reactance



🟢 Load ($p_u < 0$)

🔴 Generator ($p_u > 0$)

Objective

Objective: Use the information available outside of the attacked zone ($\vec{\theta}'_{\bar{H}}$) and the information before attack ($A, \vec{\theta}$)

- *Recover the phase angles ($\vec{\theta}'_{H}$)*
- *Detect the disconnected lines (F) or (A')*

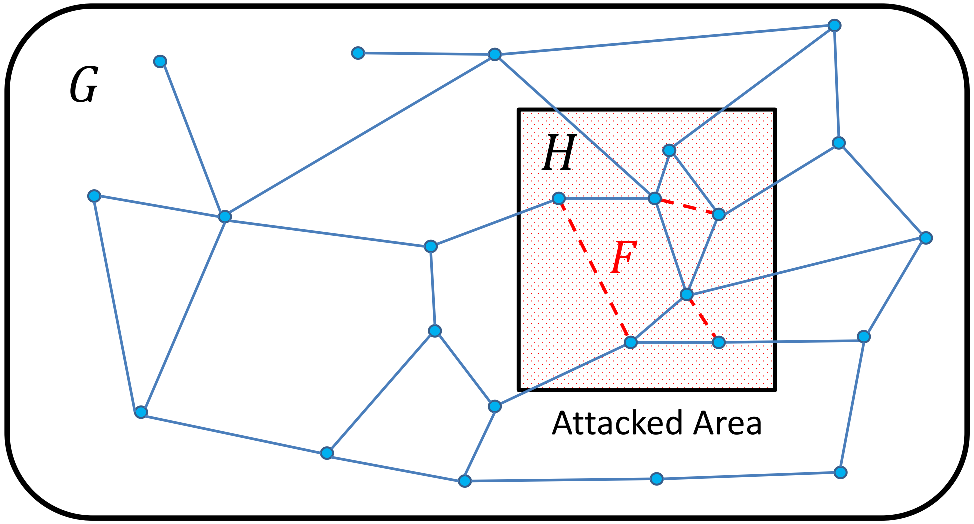
$$\vec{\theta}' = \begin{bmatrix} \vec{\theta}'_{H} \\ \vec{\theta}'_{\bar{H}} \end{bmatrix}$$

H : an induced subgraph of G that represents the attacked zone

$\bar{H} : G \backslash H$

F : Set of failed lines

○′ : The value of ○ after an attack

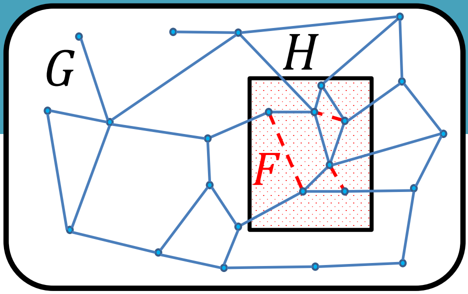


Related Work

- Line failure detection is a combinatorial problem
- Previous work on line failures detection using phase angle measurements:
  - Single or double line failures (Tate and Overbye, 2009)
  - Line failure identification in an internal system using the information from an external system (Zhu and Giannakis, 2012)
  - PMU location selection for line outage detection (Zhao, Goldsmith, Poor, 2012)
  - Recovery in transient state (Garcia et al., 2016)
  - Topology attacks (Kim and Tong, 2013)
  - …

- We use the algebraic properties of the DC power flow equations and the grid structure to efficiently (LP) detect line failures
- We show (empirically) that the approach operates well with measurement noise and under the AC power flows
- We extend to AC power flows and cases in which the attack area is unknown

Information Recovery



$\overline{H}: G\backslash H$

○′ : The value of ○ after an attack

- Assume that supply/demand values do not change or we know changes

$$\begin{cases} A\vec{\theta} = P \\ A'\vec{\theta}' = P \end{cases} \Rightarrow A(\vec{\theta} - \vec{\theta}') = (A' - A)\vec{\theta}'$$

$$\Rightarrow \begin{bmatrix} A_{H|G} \\ A_{\overline{H}|G} \end{bmatrix} (\vec{\theta} - \vec{\theta}') = \begin{bmatrix} A'_{H|G} - A_{H|G} \\ A'_{\overline{H}|G} - A_{\overline{H}|G} \end{bmatrix} \vec{\theta}'$$

Recover the phase angles

$$A_{\overline{H}|G}(\vec{\theta} - \vec{\theta}') = 0$$

Detect Line Failures

$$A_{H|G}(\vec{\theta} - \vec{\theta}') = D_H\vec{x}$$

Simultaneous
phase angles recovery
and failed lines detection

$$A_{\overline{H}|G}(\vec{\theta} - \vec{\theta}') = 0$$
$$A_{H|G}(\vec{\theta} - \vec{\theta}') = D_H\vec{x}$$

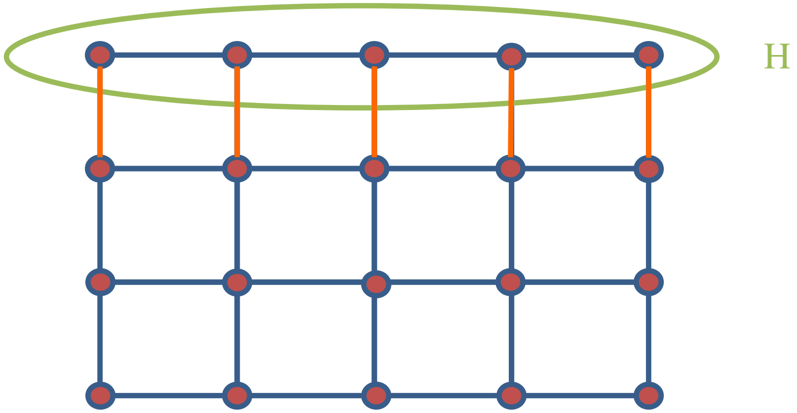
Does not depend on *the number of line failures*

Recovery of Phase Angles

$$\boxed{A_{\overline{H}|G}(\vec{\theta}-\vec{\theta}')=0 \Rightarrow A_{\overline{H}|H}\vec{\theta}'_H = A_{\overline{H}|G}\vec{\theta} - A_{\overline{H}|\overline{H}}\vec{\theta}'_{\overline{H}}}$$

$$A=\begin{bmatrix} A_{H|G} \\ A_{\overline{H}|G} \end{bmatrix} = \begin{bmatrix} A_{H|H} & A_{H|\overline{H}} \\ A_{\overline{H}|H} & A_{\overline{H}|\overline{H}} \end{bmatrix}, \vec{\theta}' = \begin{bmatrix} \vec{\theta}'_H \\ \vec{\theta}'_{\overline{H}} \end{bmatrix}$$

- $\vec{\theta}'_H$ can be recovered after any attack on H, if $A_{\overline{H}|H}$ has linearly independent columns
- $\vec{\theta}'_H$ can be recovered almost surely if there is a matching between the nodes inside and outside of H that covers V_H



Matching: A set of pairwise nonadjacent lines

Detecting Failed lines

Lemma. There exists a vector $\vec{x} \in \mathbb{R}^{|E_H|}$ such that

$$D_H\vec{x} = A_{H|G}(\vec{\theta} - \vec{\theta}')$$

and $supp(\vec{x})$ gives indices of the failed lines.

- $supp(\vec{x})$: Set of nonzero elements of vector $\vec{x}$

Lemma. The solution $\vec{x}$ is unique, if and only if H is acyclic.

- Failed lines can be detected, if H is acyclic
- *What if the set of line failures is sparse?*

$$\boxed{\min \| \vec{x} \|_1 \;\; s.t. \; D_H\vec{x} = A_{H|G}(\vec{\theta} - \vec{\theta}')} \quad (*)$$

Detecting Failed lines

$$\min \| \vec{x} \|_1 \ \ s.t. \ D_H\vec{x} = A_{H|G}(\vec{\theta} - \vec{\theta}') \qquad (*)$$

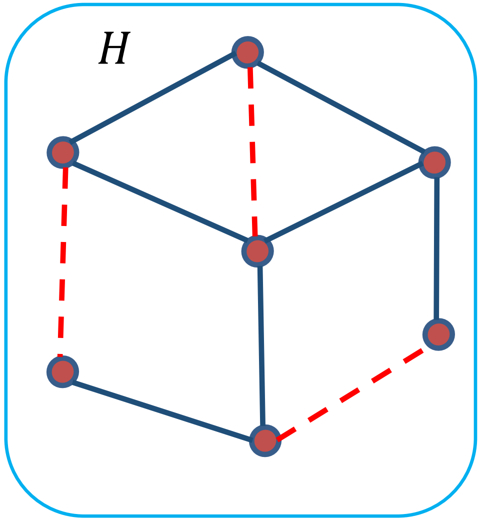
Lemma. If H is a cycle and less than half of the lines are failed, then the solution $\vec{x}$ to the optimization $(*)$ is unique and $supp(\vec{x})$ gives indices of the failed lines.

Theorem. In a planar graph H, *the solution to* $(*)$

$$\min \| \vec{x} \|_1 \ \ s.t. \ D_H\vec{x} = A_{H|G}(\vec{\theta} - \vec{\theta}')$$

is unique and $supp(\vec{x})$ gives indices of the failed lines, if the following conditions hold:
(i) for any cycle, less than half of its lines are failed,
(ii) F^* can be covered by edge-disjoint cycles in H^*

Example



Simultaneous Phase Angles Recovery and Failed lines Detection

- Find vectors $\vec{x} \in \mathbb{R}^{|E_H|}$ and $\vec{\theta'}_H \in \mathbb{R}^{|V_H|}$ such that

$$D_H\vec{x} = A_{H|G}(\vec{\theta} - \vec{\theta}')$$
$$A_{\overline{H}|G}\left(\vec{\theta} - \vec{\theta}'\right) = 0$$

- Unique solution if, and only if,
  1. H is acyclic
  2. There is a matching between the nodes in H and $\overline{H}$
- $supp(\vec{x})$ gives the indices of the failed lines

- Assuming the set of line failures is sparse:

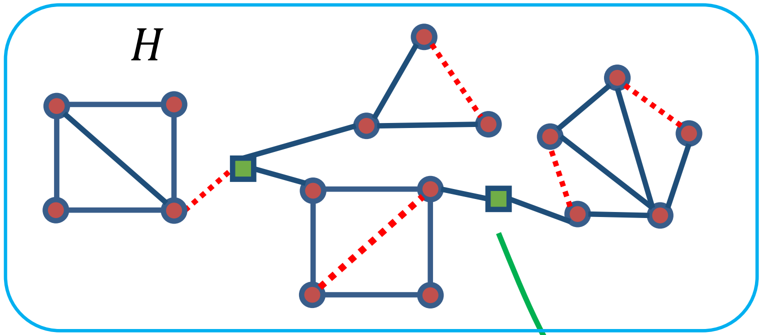
$$\boxed{\begin{aligned} &\min \| \vec{x} \|_1 \;\; s.t. \\ &D_H\vec{x} = A_{H|G}(\vec{\theta} - \vec{\theta}') \\ &A_{\overline{H}|G}\left(\vec{\theta} - \vec{\theta}'\right) = 0 \end{aligned}} \qquad (**)$$

Simultaneous Phase Angles Recovery and Failed Lines Detection

Theorem. Under some conditions on F and H the solution $\vec{x}, \vec{\theta}'_H$ to $(**)$ is unique and can recover the missing information.

$$\begin{aligned} &\min \|\vec{x}\|_1 \quad s.t. \\ &D_H\vec{x} = A_{H|G}(\vec{\theta} - \vec{\theta}') \qquad (**) \\ &A_{\overline{H}|G}(\vec{\theta} - \vec{\theta}') = 0 \end{aligned}$$

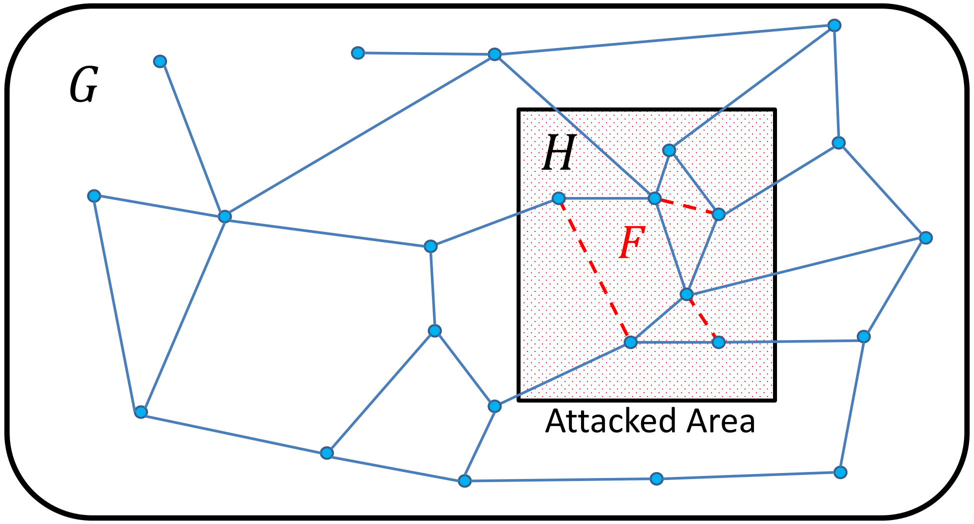
- *Example.*



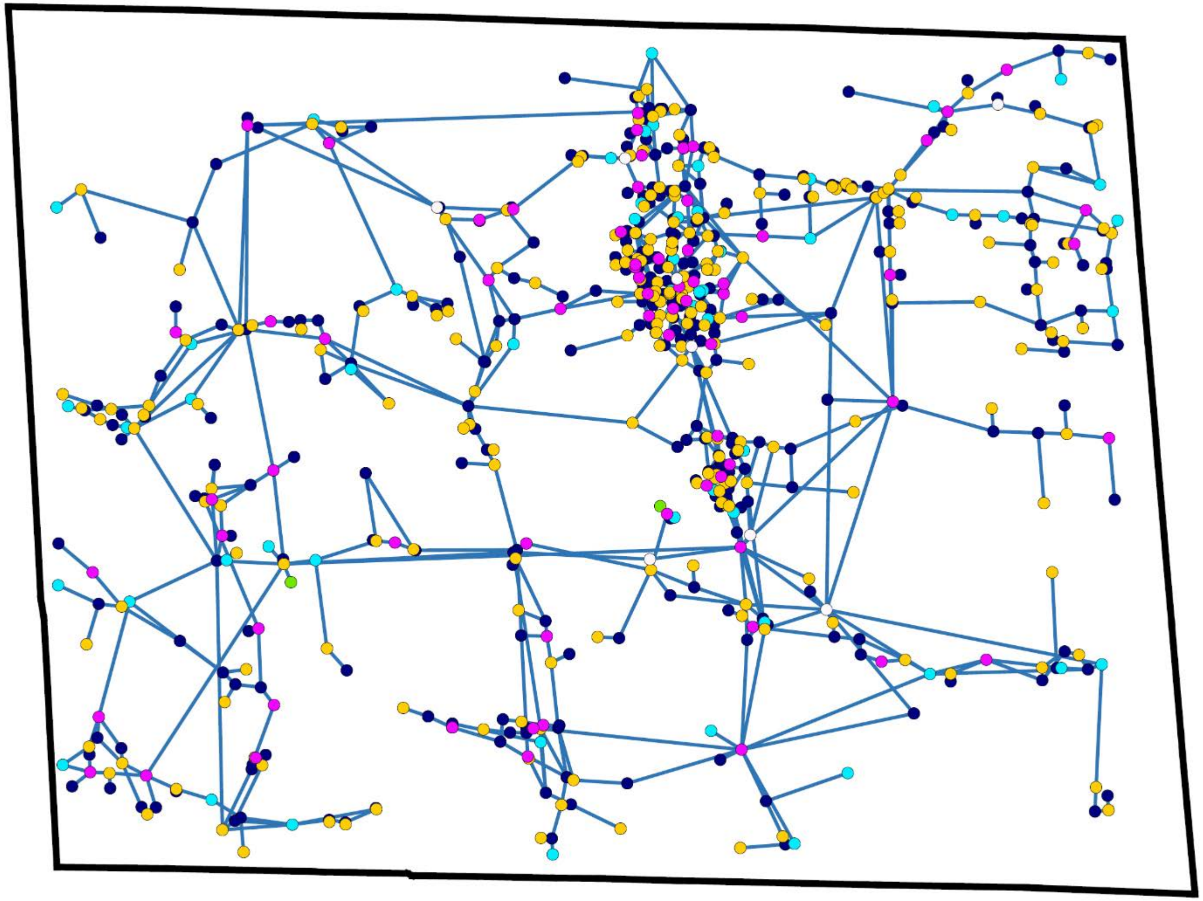
H-inner-connected

Conditions and Constraints

External Conditions	Internal Conditions	Attack Constraints
Matching	Acyclic	None
Matching	Planar	Less than half of the edges in each cycle are failed
Partial Matching	Acyclic	Less than half of the edges connected to an internal node are failed
Partial Matching	Planar	Two above conditions



Partitioning of the Colorado state grid into 6 attack-resilient zones



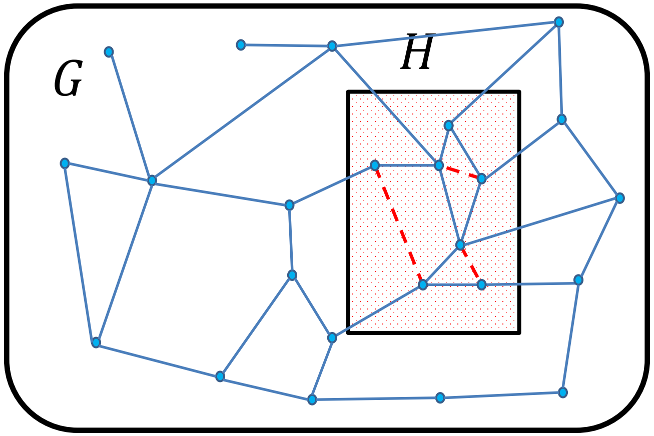
- NP-Hard
- Developed good approximations

Outline

- State estimation under a cyber and physical attack
- State estimation in the presence of measurement noise and uncertainty
- State estimation under the AC power flows
- Attack identification when the affected area is unknown

Measurement Noise and Uncertainty

- Assume $A\left(\vec{\theta}\ - \vec{\mathrm{n}}\right) = \vec{p}$
  $\vec{\mathrm{n}}$ is a Gaussian measurement noise
- $SNR = 20\log\frac{\|\vec{\theta}\|_2}{\|\vec{\mathrm{n}}\|_2}$
- Relax the constraints as in $(***)$

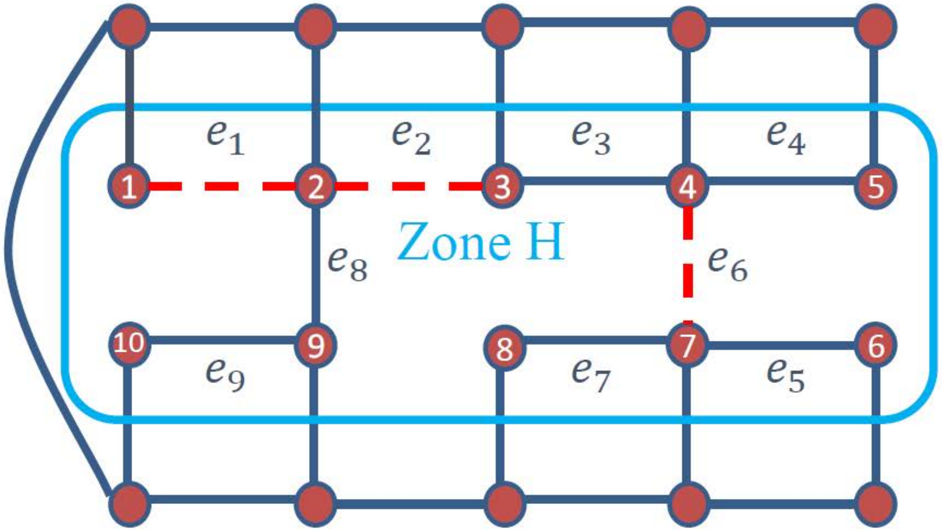


$$\min \| \vec{x} \|_1 \ \ s.t.$$
$$||D_H\vec{x} - A_{H|G}(\vec{\theta} - \vec{\theta}')||_2 < \epsilon_1$$
$$||A_{\overline{H}|G}(\vec{\theta} - \vec{\theta}')||_2 < \epsilon_2$$

$(***)$

- Second order cone programming

Numerical Results

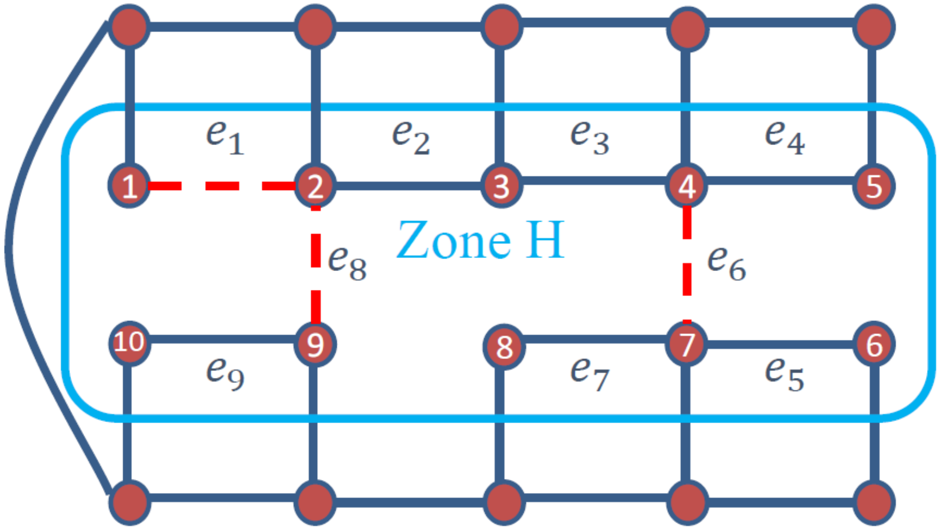


Actual $\vec{\theta}'_H$	Recovered $\vec{\theta}'_H$
0.1062	0.1030
0.0882	0.0872
0.0042	0.0075
0.0062	0.0036
0.0342	0.0334
-0.0758	-0.0752
-0.0971	-0.0937
-0.0925	-0.0919
0.0442	0.0441
0.0322	0.0321

SNR = 50dB

$\vec{x}$	-0.0074	-0.0068	0	0	0	-0.0857	0	0	0
supp($\vec{x}$)	1	1	0	0	0	1	0	0	0

Numerical Results



Actual $\vec{\theta}'_H$	Recovered $\vec{\theta}'_H$
0.1071	0.1203
0.1428	0.1242
0.1085	0.0997
0.1013	0.0884
0.1271	0.1273
-0.1409	-0.1381
-0.1623	-0.1496
-0.1576	-0.1563
-0.0429	-0.0299
-0.0329	0.0321

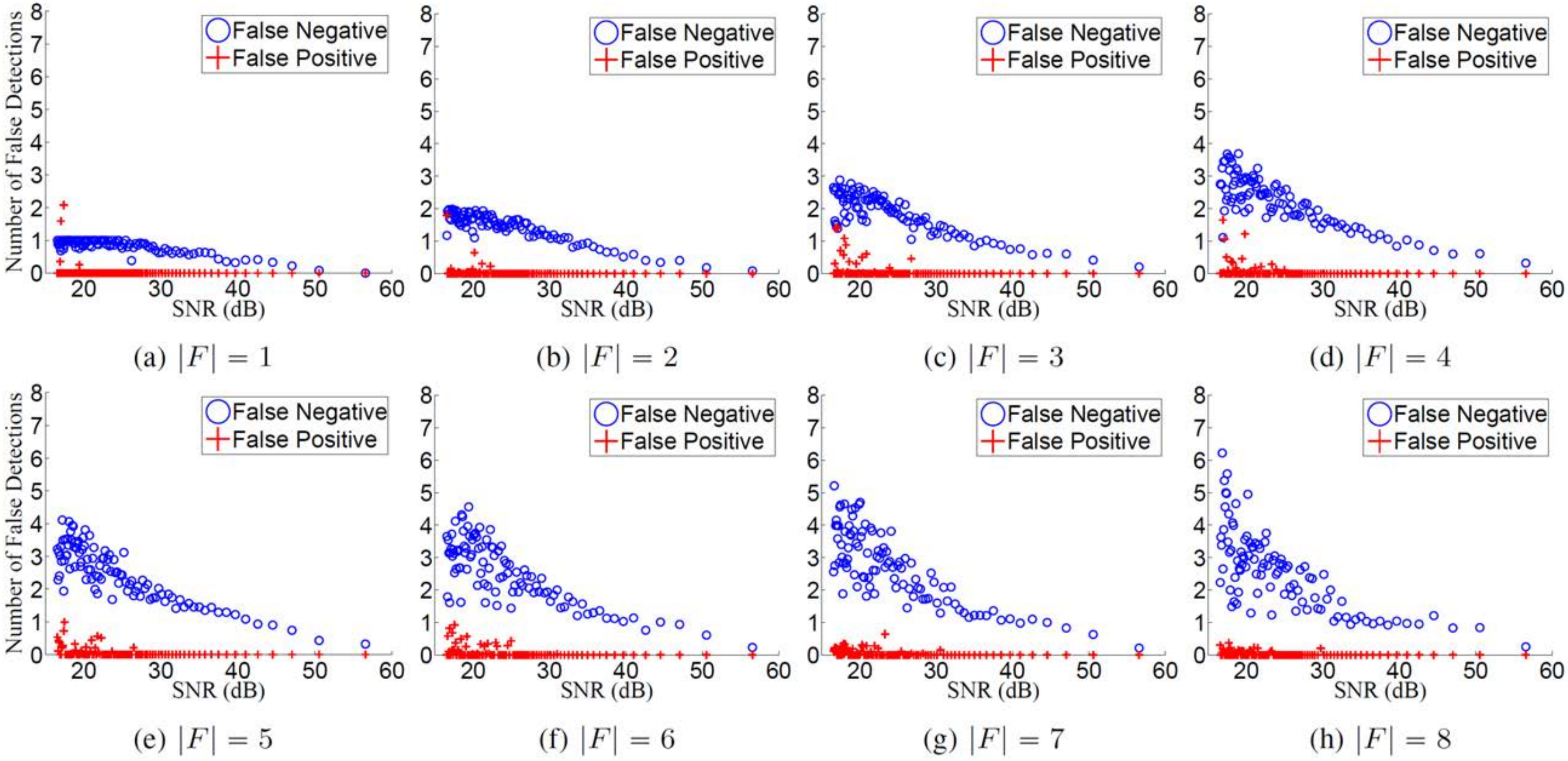
SNR = 30dB

Could not detect failure in e_1

$\vec{x}$	0	0	0	0	0	-0.1938	0	-0.1164	0
supp($\vec{x}$)	0	0	0	0	0	1	0	1	0

Numerical Results

False Negative: Not detecting a failed line
False Positive: Detecting an operating line as failed



(a) $|F| = 1$

(b) $|F| = 2$

(c) $|F| = 3$

(d) $|F| = 4$

(e) $|F| = 5$

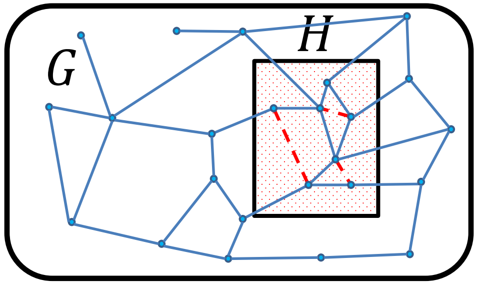
(f) $|F| = 6$

(g) $|F| = 7$

(h) $|F| = 8$

State Estimation under the AC Power Flows

- Phase angle of the nodes are computed under *the AC power flows*

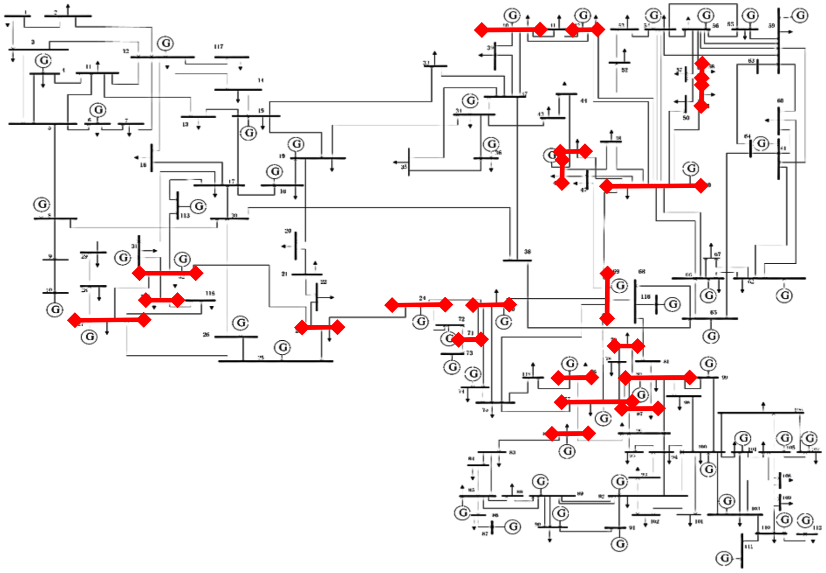


$$\min \| \vec{x} \|_1 \;\; s.t.$$
$$||D_H\vec{x} - A_{H|G}(\vec{\theta} - \vec{\theta}')||_2 < \epsilon_1$$
$$||A_{\overline{H}|G}(\vec{\theta} - \vec{\theta}')||_2 < \epsilon_2$$

$(***)$

- Use different ϵ_1 and ϵ_2 values and statistically detect the line failures

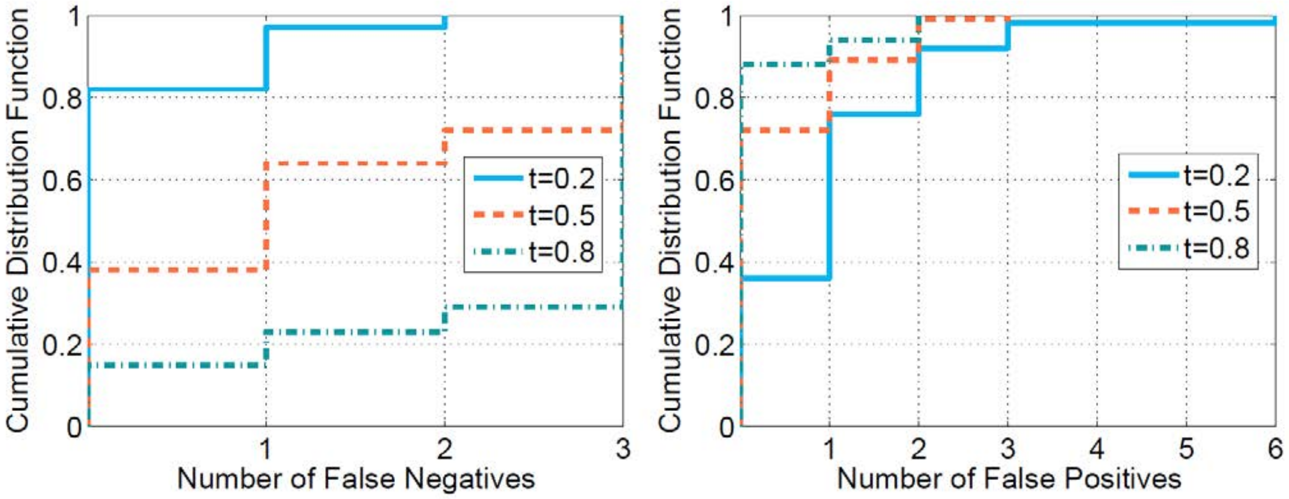
- IEEE 118-bus system - an attacked area with 21 nodes and 22 lines



Saleh Soltan and Gil Zussman, “Power Grid State Estimation after a Cyber-Physical Attack under the AC Power Flow Model,” Proc. IEEE PES-GM’17, July 2017.

State Estimation under the AC Power Flows

- The phase angles can be estimated with less than 1% error for 1-, 2-, and 3-line failures
- Line failures can also be detected with less than 20% error (false positives or negatives)

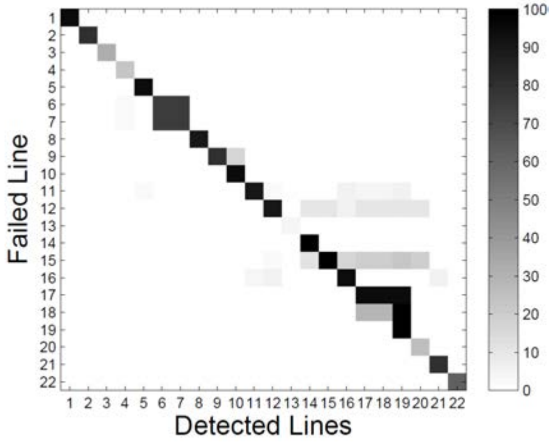


The CDF of the number of false negatives and positives in detecting triple line failures (100 cases).

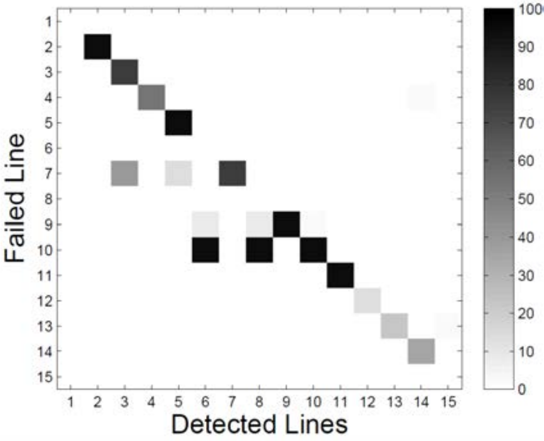
Dealing with AC Directly

- Recovering the voltages
  - Similar conditions
  - Non-linear but convex
- DC-based method:

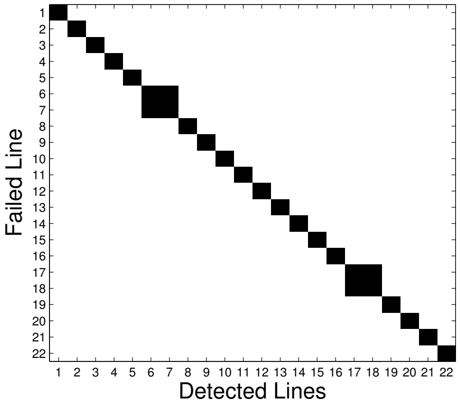
118-bus

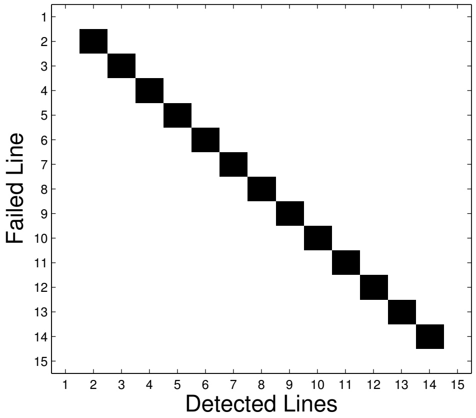


300-bus



- AC-based method:





- Evaluation when Zone does not satisfy conditions

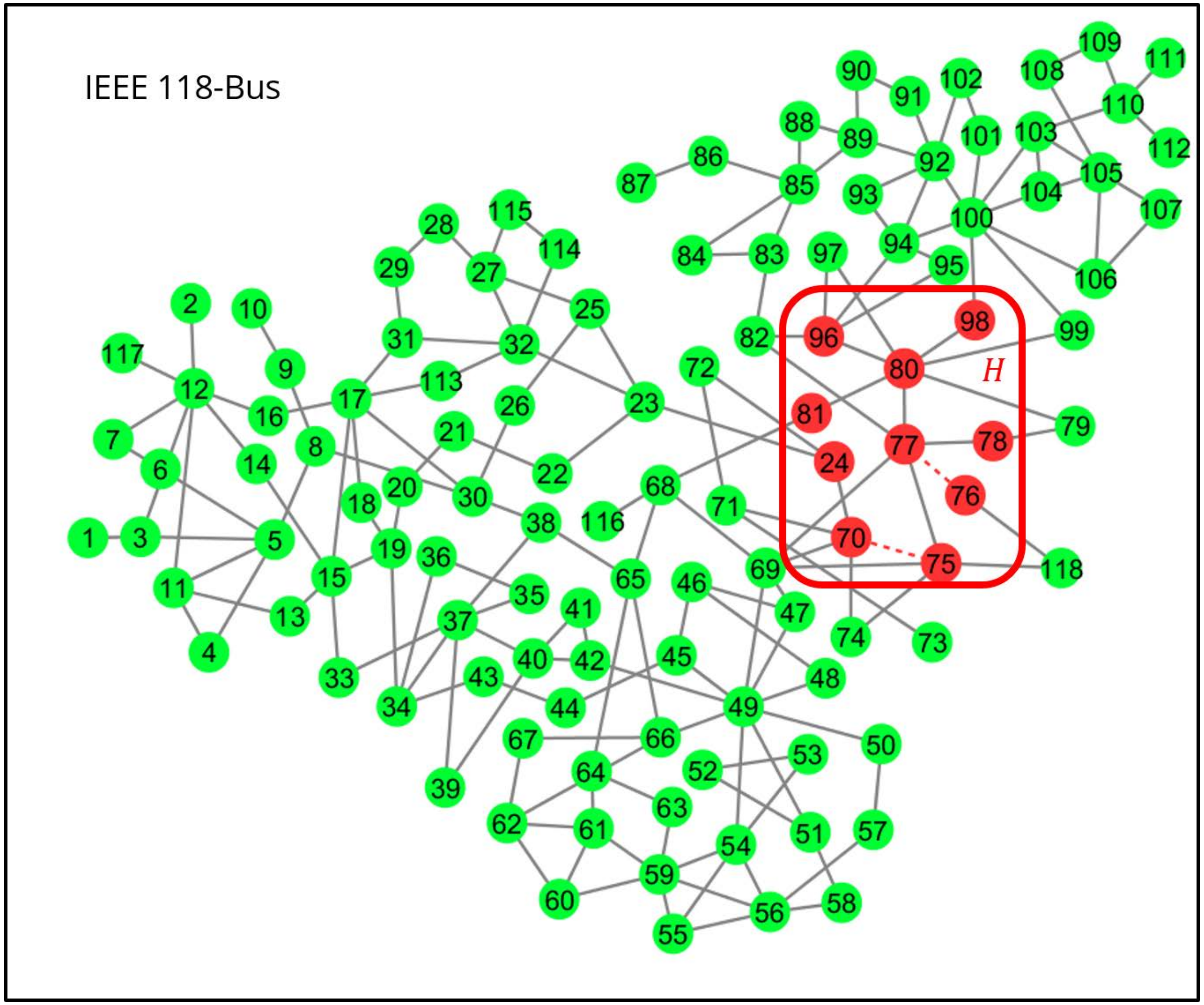
Saleh Soltan, Mihalis Yannakakis, Gil Zussman, “REACT to Cyber Attacks on Power Grids,” in preparation.

Outline

- State estimation under a cyber and physical attack
- State estimation in the presence of measurement noise and uncertainty
- State estimation under the AC power flows
- Attack identification when the affected area is unknown

Back to DC - Location Unknown

- Detect the line failures as well as the attacked area H after a cyber-physical attack



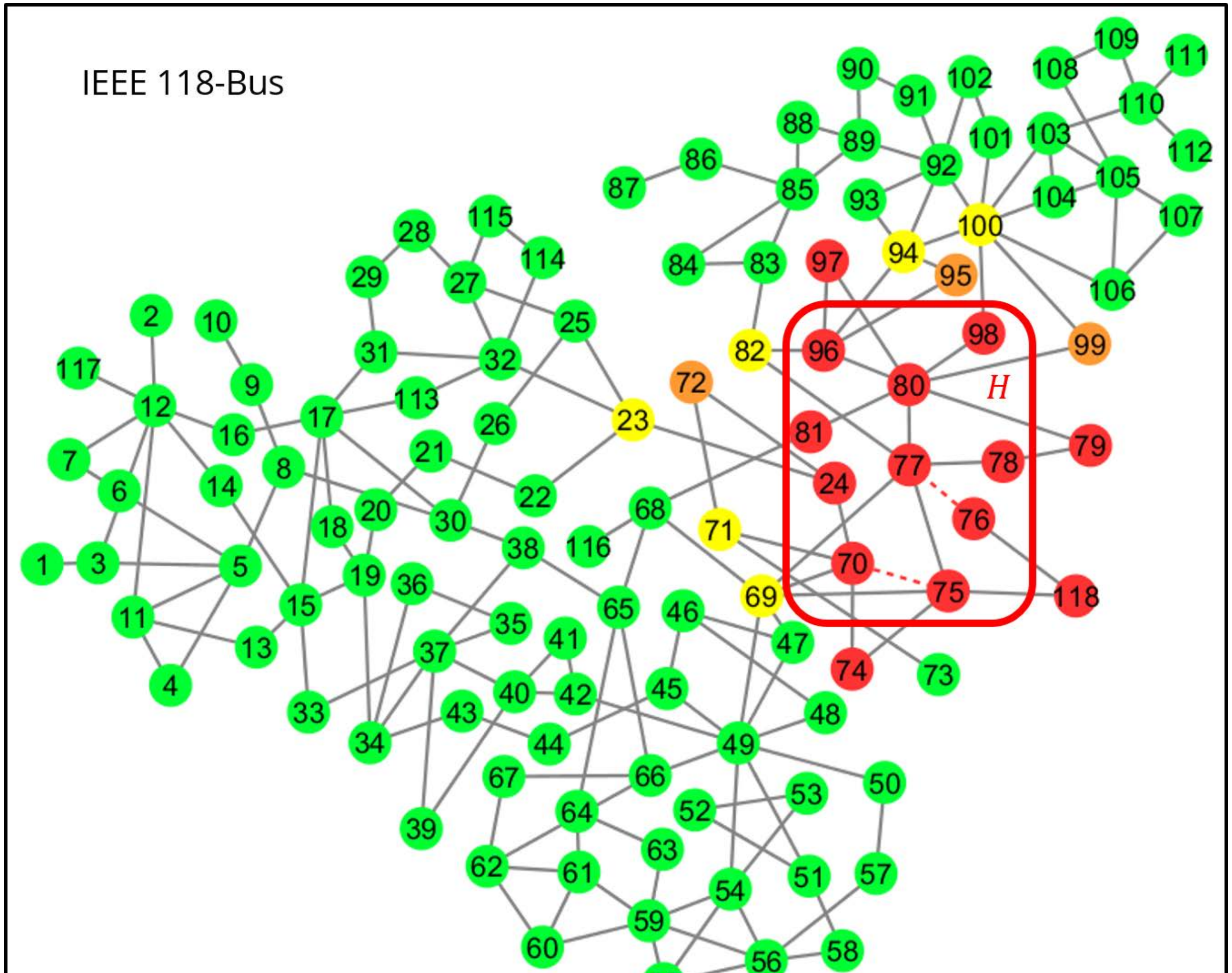
Saleh Soltan, Mihalis Yannakakis, Gil Zussman, “REACT to Cyber Attacks on Power Grids,” submitted.

Location Unknown - Cyber Attacks

- Two types of cyber attacks
  - Data distortion
  - Data Replay
- $\vec{\theta}^{\star}$ is the observed phase angles after the attack which is different from the actual $\vec{\theta}'$
- NP-Hard
- Approximate solutions

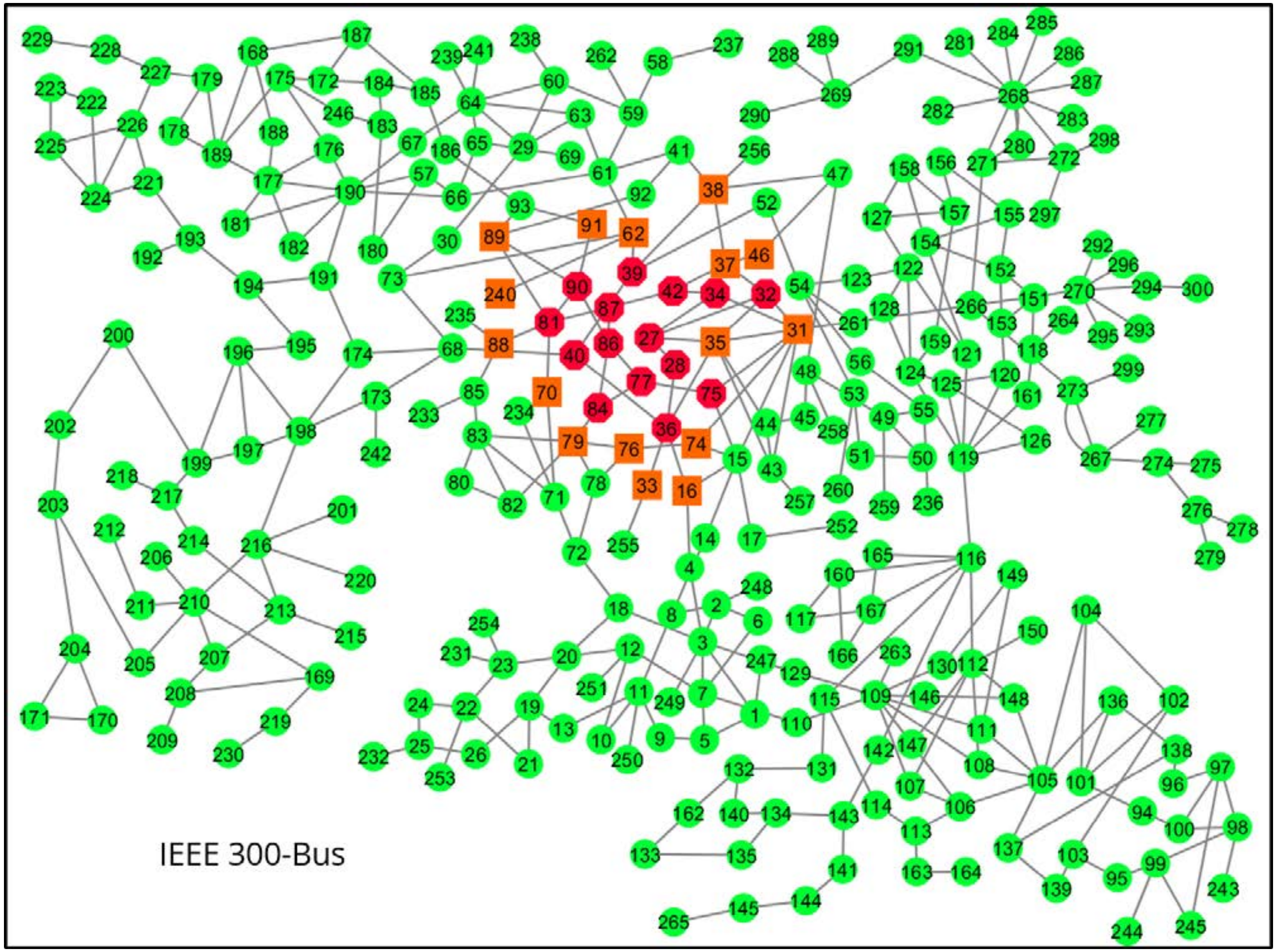
Example

- Approximately detect the attacked area in 3 steps
- Identify line failures with some confidence

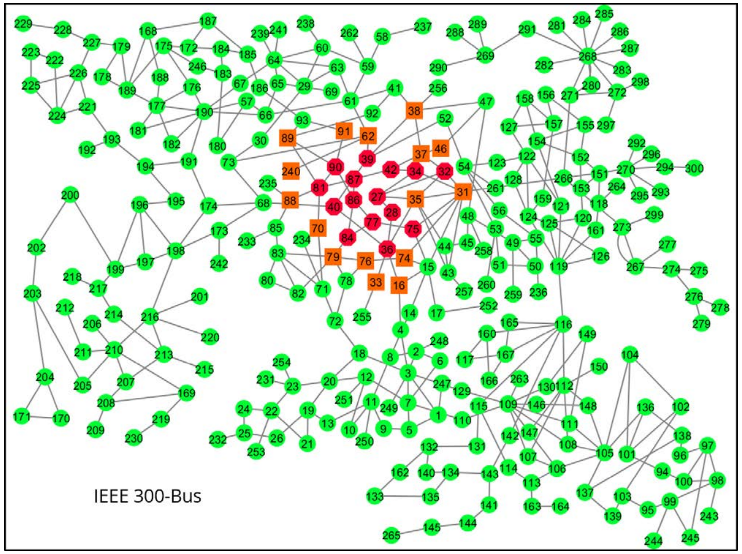


Simulations

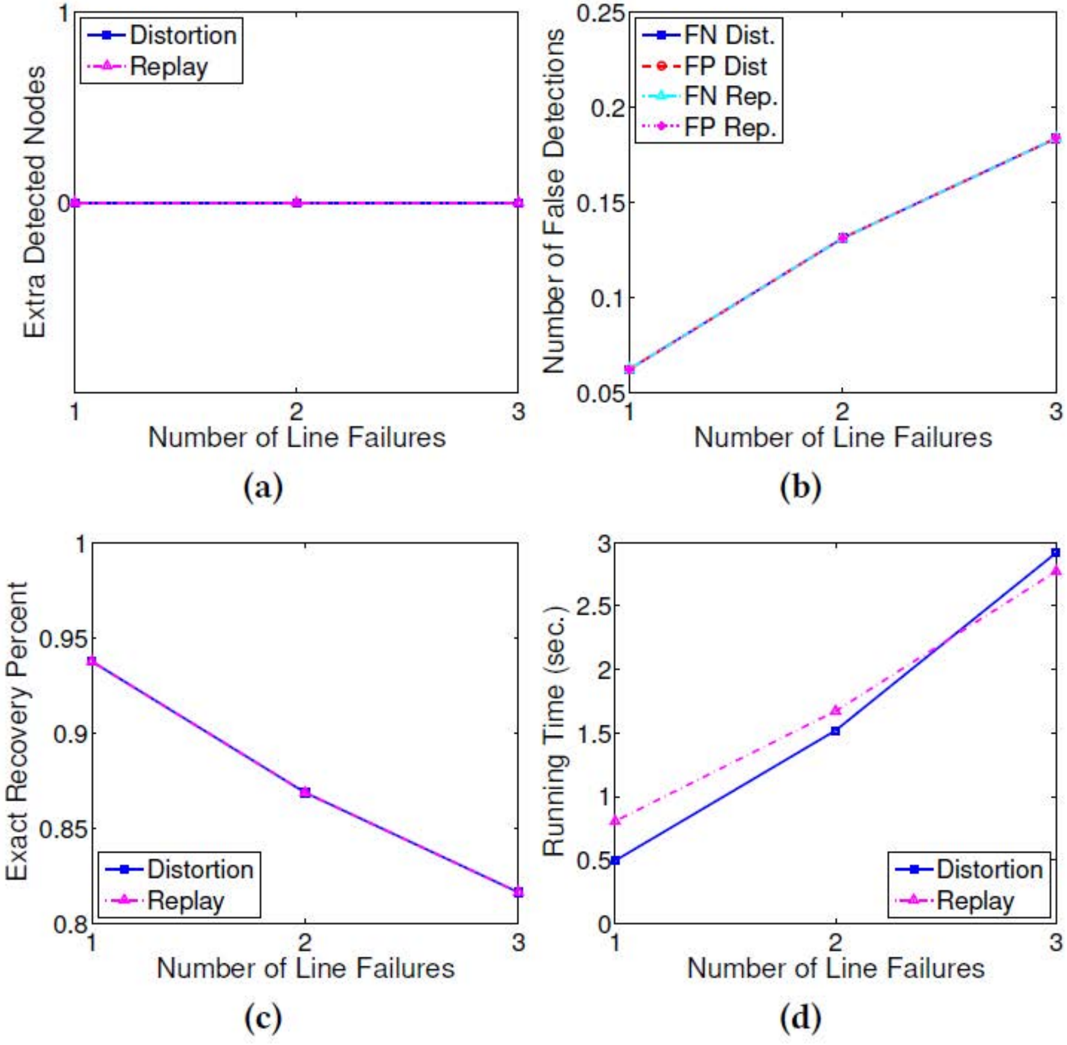
- Simulations on two attacked areas within 300-bus system
  - Attacked areas with 15 and 31 nodes



Performance - Small Area



100 1,2,3-line failure samples



(a)

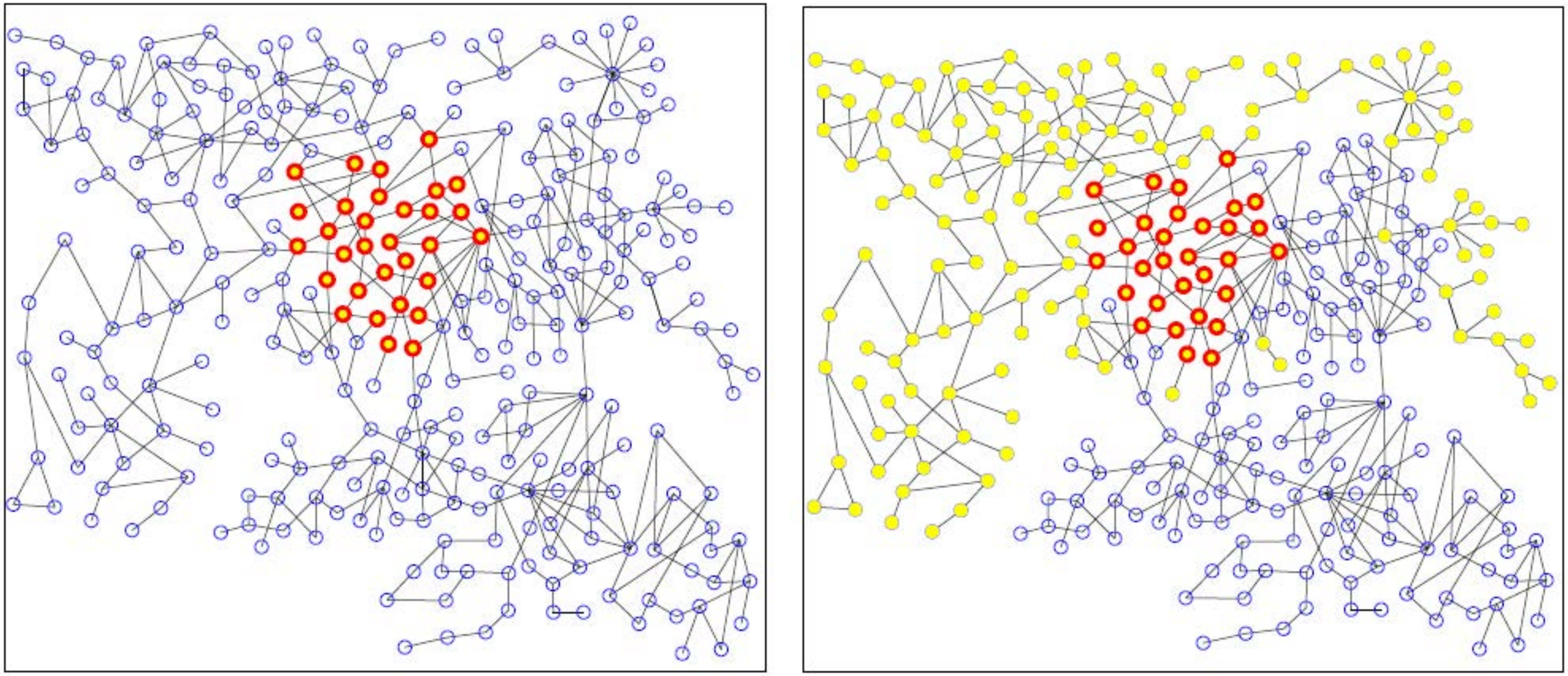
(b)

(c)

(d)

Distortion vs. Replay Attacks

- Replay attacks are much harder to cope with

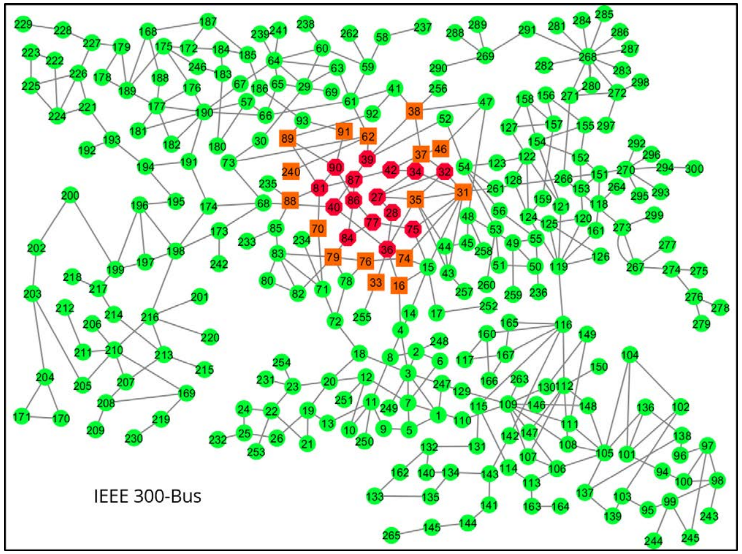


(a) Data Distortion Attack

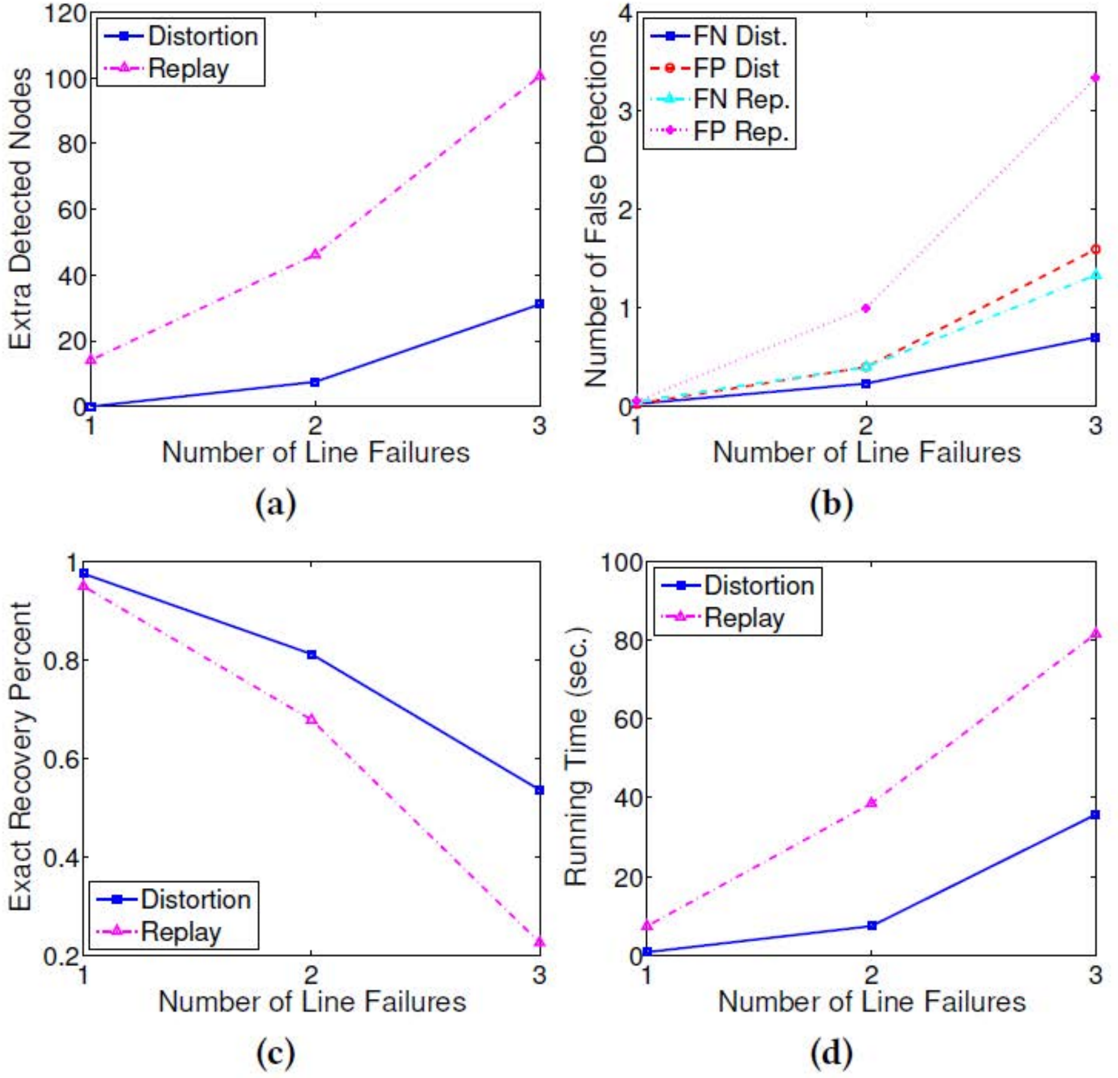
(b) Data Replay Attack

- IEEE 300-bus
- 3 line failures

Performance - Large Area



100 1,2,3-line failure samples



(a)

(b)

(c)

(d)

Summary

- Developed schemes for detecting line failures following an attack using partial information
- Identified tradeoffs between the structural complexity of the attack area and the accuracy of data recovery
- Evaluated under measurement noise and AC model
- Developed an AC-based algorithm
- Considered the case in which attacked area is unknown and there are false data injections
- (near) future work: deal with a (streaming) simulated attack on a ~1500 bus grid…

BSF
United States – Israel
Binational Science Foundation

Defense Threat Reduction Agency
United States of America

NSF

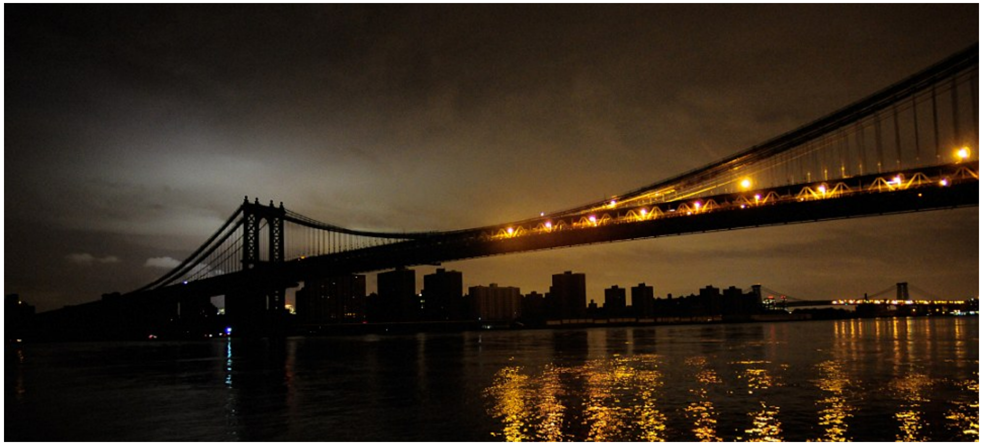
DARPA

GRID
MODERNIZATION
LABORATORY
CONSORTIUM
U.S. Department of Energy

ICRC
Blavatnik Interdisciplinary
Cyber Research Center

Postdoctoral Positions – Power Grid Resilience

- The groups of D. Bienstock and G. Zussman (Columbia University)
- DARPA, DOE, DTRA, and ARPA-E projects
- Background and experience in some of the following: optimization, power, machine learning, control, algorithm design, and computational implementation
- dano@columbia.edu , gil@ee.columbia.edu
- wimnet.ee.columbia.edu



COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK