

Optimal False Data Injection Attack against Automatic Generation Control in Power Grids

Rui Tan¹ Hoang Hai Nguyen² Eddy. Y. S. Foo³ Xinshu Dong⁴

David K. Y. Yau^{1,5} Zbigniew Kalbarczyk² Ravishankar K. Iyer² Hoay Beng Gooi³

¹ School of Computer Science & Engineering, Nanyang Technological University

² Coordinated Science Lab, University of Illinois at Urbana-Champaign

³ School of Electrical & Electronic Engineering, Nanyang Technological University

⁴ Advanced Digital Sciences Center, Illinois at Singapore

⁵ Singapore University of Technology and Design

Attacks against Power Grids



REUTERS EDITION: U.S.

HOME BUSINESS MARKETS WORLD POLITICS TECH

REUTERS VIDEO
The Latest in Business, Finance & Technology News

UPDATE 1-Malicious virus shuttered U.S. power plant -DHS

Wed Jan 16, 2013 5:53pm EST

12 COMMENTS | Tweet 409 | in Share 72 | Share this | +1 25 | Email | Print



BBC News Sport Weather Earth Future

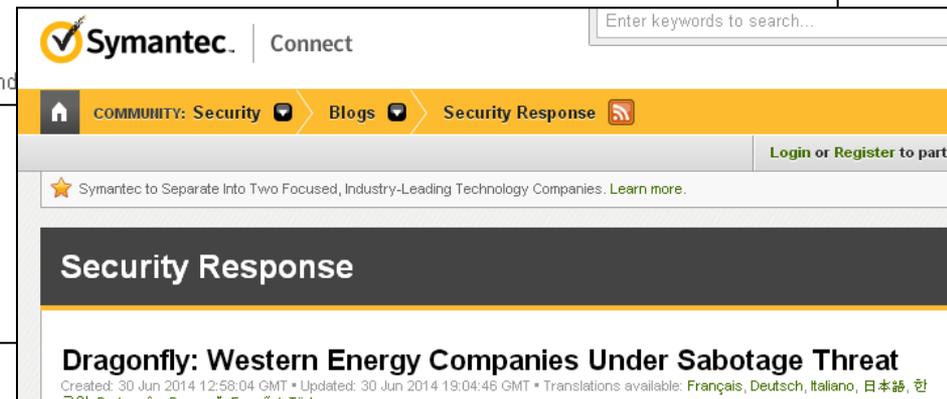
NEWS TECHNOLOGY

Home World Asia Australia India China UK Business Health Science/Environment Technology

16 October 2014 Last updated at 12:18

Smart meters can be hacked to cut power bills

By Mark Ward
Technology correspondent



Symantec. Connect

Enter keywords to search...

COMMUNITY: Security Blogs Security Response

Login or Register to participate

Symantec to Separate Into Two Focused, Industry-Leading Technology Companies. Learn more.

Security Response

Dragonfly: Western Energy Companies Under Sabotage Threat

Created: 30 Jun 2014 12:58:04 GMT • Updated: 30 Jun 2014 19:04:46 GMT • Translations available: Français, Deutsch, Italiano, 日本語, 中文

- **59%** attacks on critical infrastructures target grids [DHS'13]
 - *Night Dragon*: grid operation **data exfiltration** [McAfee'11]
 - *Dragonfly*: **Trojan horses** in grid control systems [Symantec'14]

Frequency Control

- Maintain freq. at 50 or 60 Hz when loads change

The diagram shows the power balance equation $P_L + D \cdot \omega = P_G$. Callouts identify the terms: P_L is 'Frequency-insensitive load', $D \cdot \omega$ is 'Frequency-sensitive load', and P_G is 'Generation'. The citation '[Kundur 94]' is located to the right of the equation.

$$P_L + D \cdot \omega = P_G$$

[Kundur 94]

D : load-damping constant
 ω : system frequency

- Widespread and costly impact of failure
 - System frequency is **global**
 - **-0.5 Hz**: load shedding (regional blackout)
 - **±2.0 Hz**: permanent equipment damage

Automatic Generation Control (AGC)

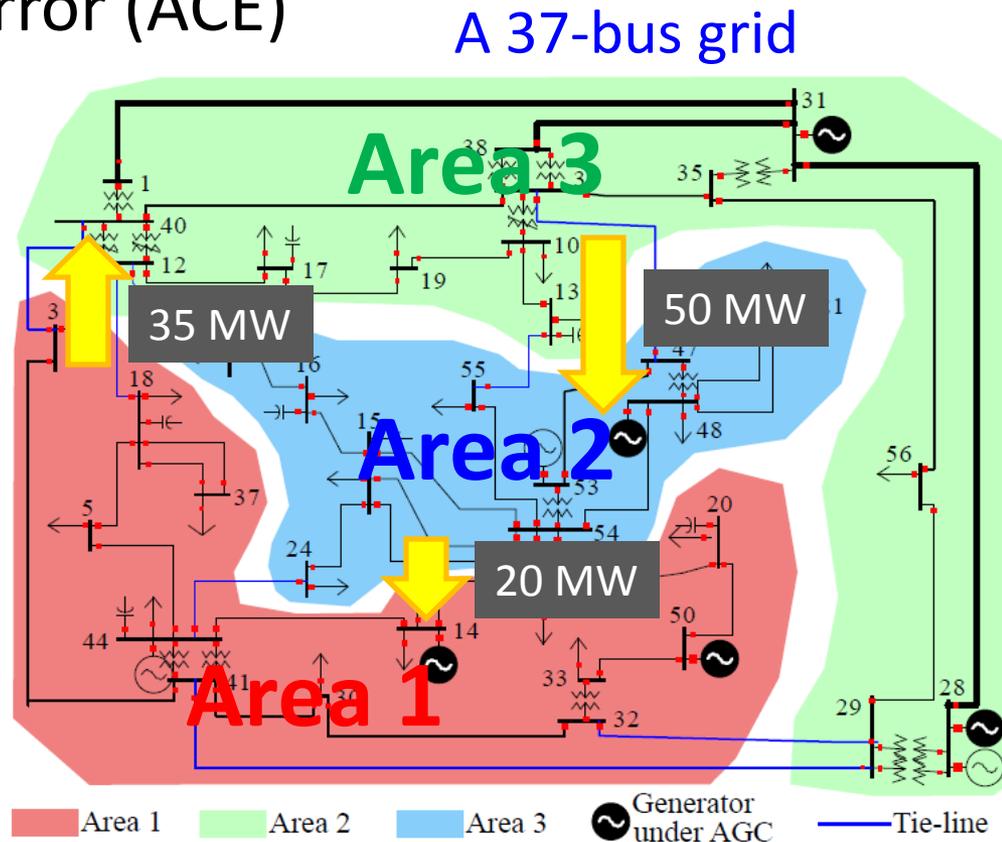
- More than frequency control
 - Regulates power exchanges btw *control areas*
 - **Input:** freq. deviation, area power export deviations
 - **Output:** Area Control Error (ACE)

$$ACE = \alpha \cdot \Delta\omega + \beta \cdot \Delta p_E$$

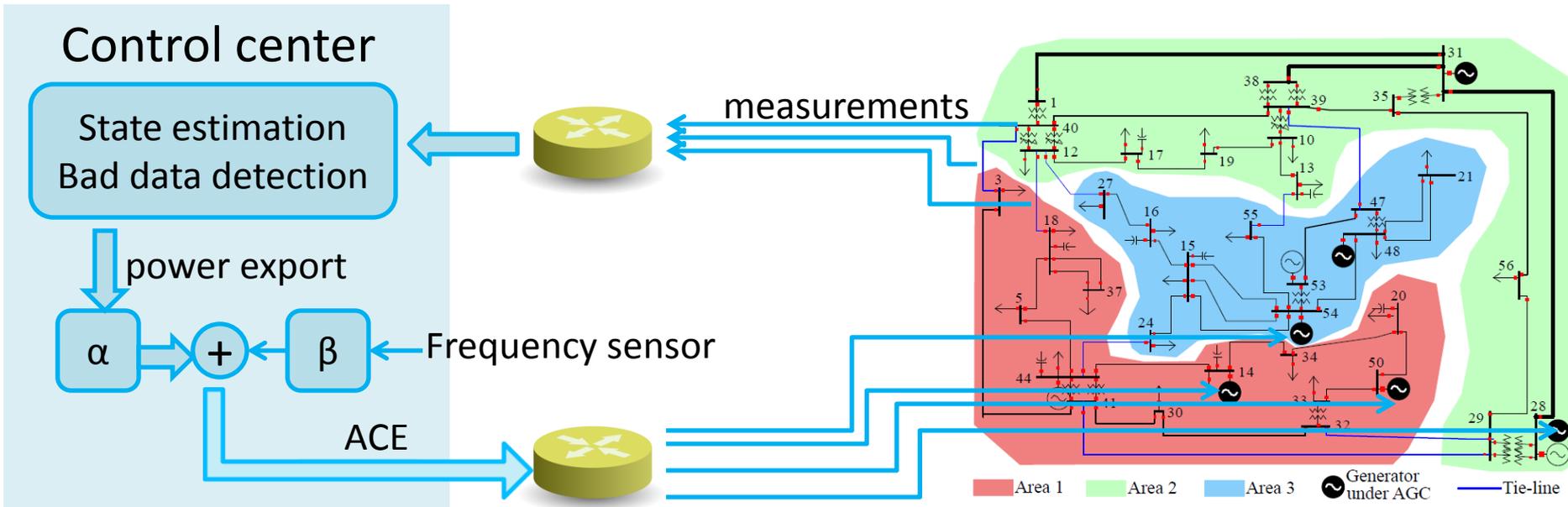
Frequency deviation

Power export deviation

- Further dispatched to generators
- **Cycle:** 2 to 4 secs

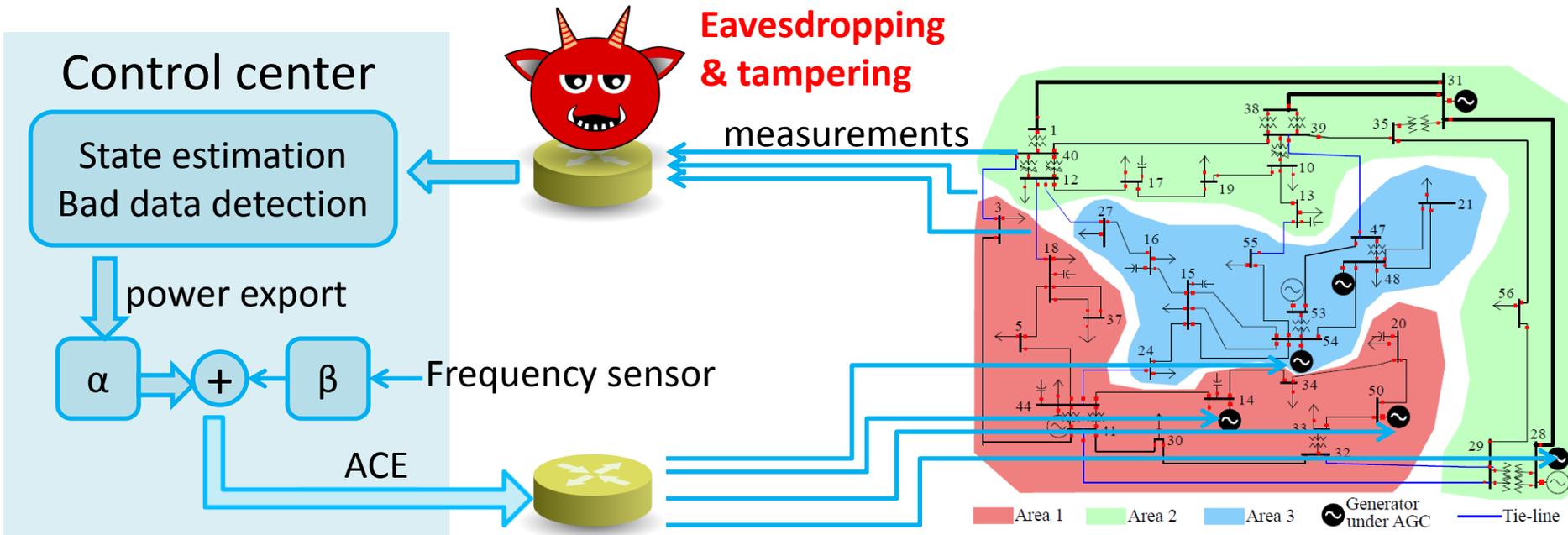


Implementation & Threat



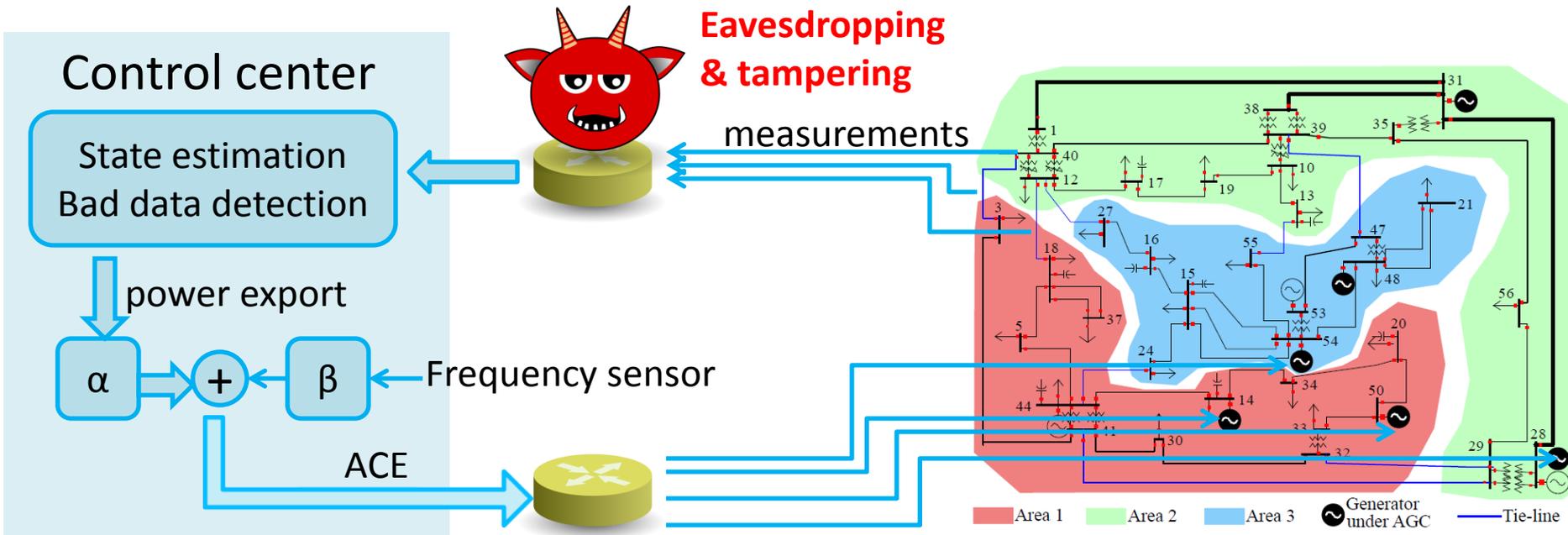
- Networked control system
 - Distributed, networked sensors
 - Control center
 - Generators

Implementation & Threat



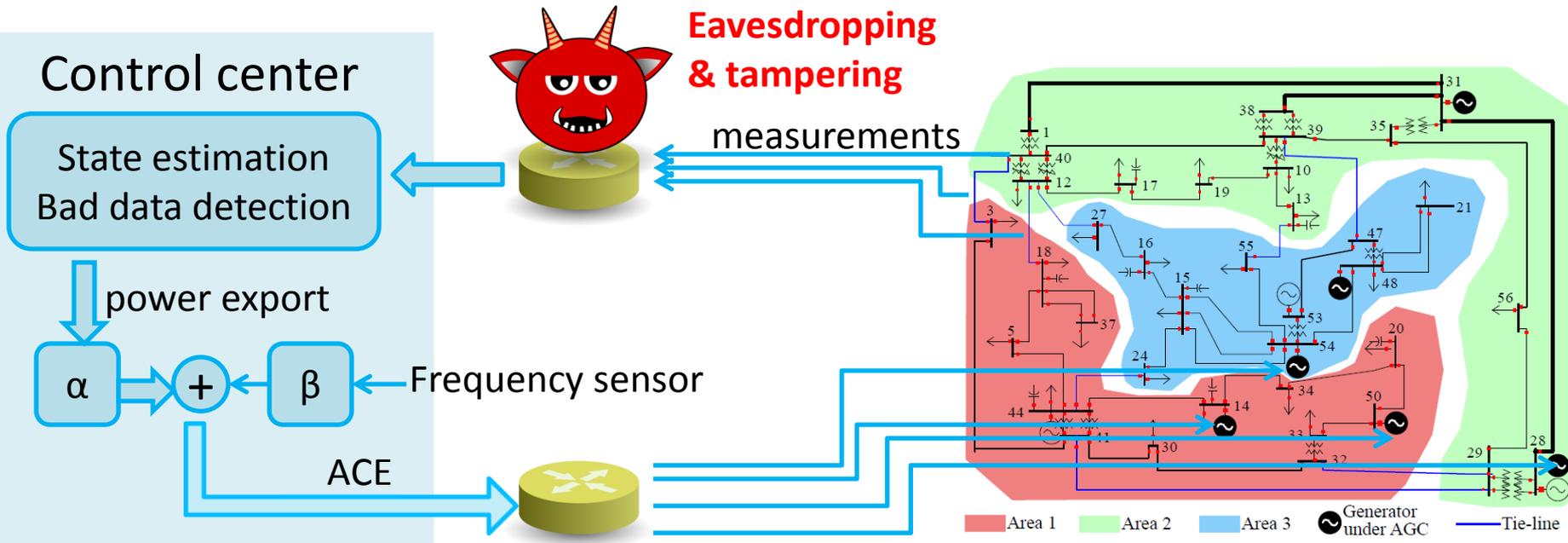
- Networked control system
 - Distributed, networked sensors
 - Logically isolated links (e.g., VPN) in existing networks
 - Control center
 - Generators

Implementation & Threat



- Networked control system
 - Distributed, networked sensors
 - Logically isolated links (e.g., VPN) in existing networks
 - Control center
 - Well protected ✓
 - Generators

Implementation & Threat

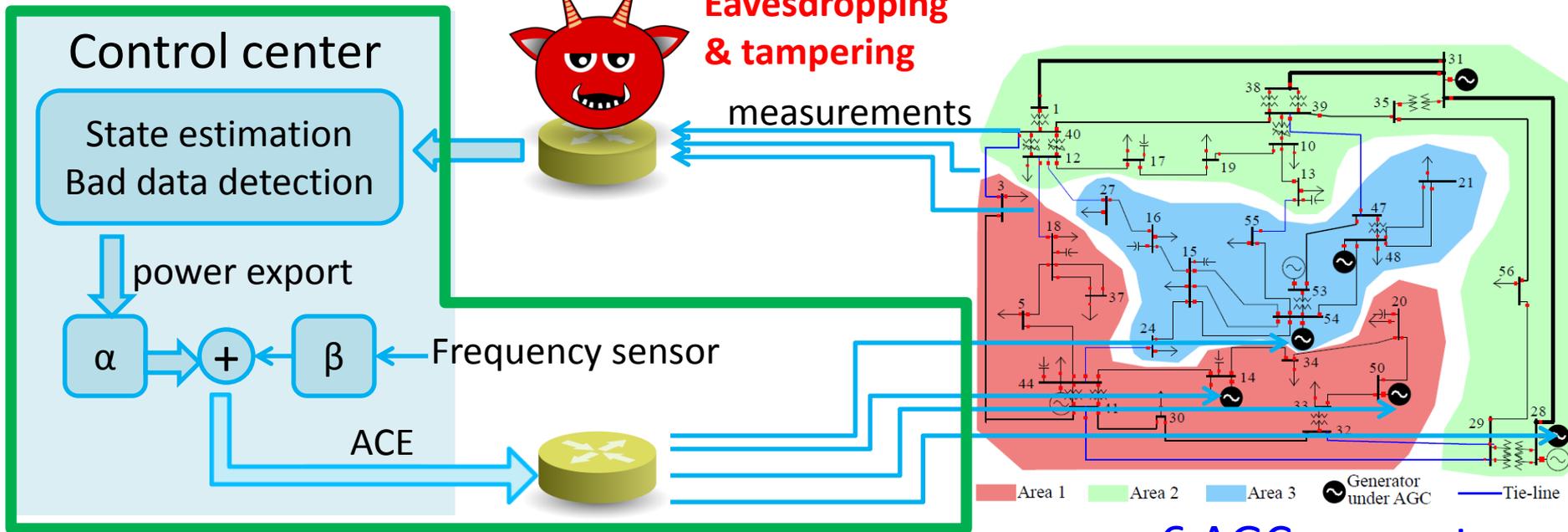


6 AGC generators
vs. 81 sensors

- Networked control system
 - Distributed, networked sensors
 - Logically isolated links (e.g., VPN) in existing networks
 - Control center
 - Well protected ✓
 - Generators
 - Limited #, physically isolated links ✓

Implementation & Threat

Trusted



6 AGC generators
vs. 81 sensors

- Networked control system
 - Distributed, networked sensors
 - Logically isolated links (e.g., VPN) in existing networks
 - Control center
 - Well protected ✓
 - Generators
 - Limited #, physically isolated links ✓

Outline

- Motivation & Background
- **Attack Model**
- Optimal Attack
- Simulations & Testbed Experiments

False Data Injection (FDI)

- Corrupt measurements

$$\mathbf{z}' = \mathbf{z} + \mathbf{a}$$

Attack vector

Measurement vector

- Can read \mathbf{z} and corrupt a subset of \mathbf{z} elements
- Stealthy to fault/attack detectors
 - Bypass bad data detection [Liu CCS'09]

$$\mathbf{a} = \mathbf{H}\mathbf{c}$$

Arbitrary vector

Measurement matrix

- Bypass data quality checks

$$- \mathbf{a}_{\max} \prec \mathbf{a} \prec \mathbf{a}_{\max}$$

Research Question

- **Optimal attack** via worst-case analysis
 - A sequence of FDIs to deviate frequency to unsafe level **in shortest time**
Protection assessment given attack response delay
 - **How to compute?**
 - **Achievable in practice?**
- Existing work on AGC security
 - Simulations based on predefined attack templates
scaling, ramps, surges, random noises, time delays
[Bose 2004 2005] [Sridhar 2010 2014]
 - Reachability analysis [Esfahani 2010]

Outline

- Motivation & Background
- Attack Model
- **Optimal Attack**
 - **How to compute?**
 - Achievable?
- Simulations & Testbed Experiments

Attack Impact Model

In Laplace domain: $\Delta \omega = \Phi \cdot \Delta \mathbf{p} + \Phi \Lambda \mathbf{T} \cdot \mathbf{a}$

Freq. deviation

Vector of load
changes of areas

Attack vector

- **T**: constant integer matrix
 - From grid topology
 - Can be obtained by attacker (social engineering)
- **Φ, Λ**
 - Transfer functions of turbines, generators, transmission lines, etc
 - Closed-forms unknown

Optimal Attack

- Time-to-emergency (TTE): remaining time before

$$\Delta \omega \notin (\Delta \omega_{\min}, \Delta \omega_{\max})$$

– $\Delta \omega_{\min} = -0.5 \text{ Hz}$: load shedding (regional blackout)

- Compute a series of \mathbf{a} to minimize TTE subject to

– Write access

– Stealthiness

- $\Delta \omega = \Phi \cdot \Delta \mathbf{p} + \Phi \Lambda \mathbf{T} \cdot \mathbf{a}$ complex differential eqns

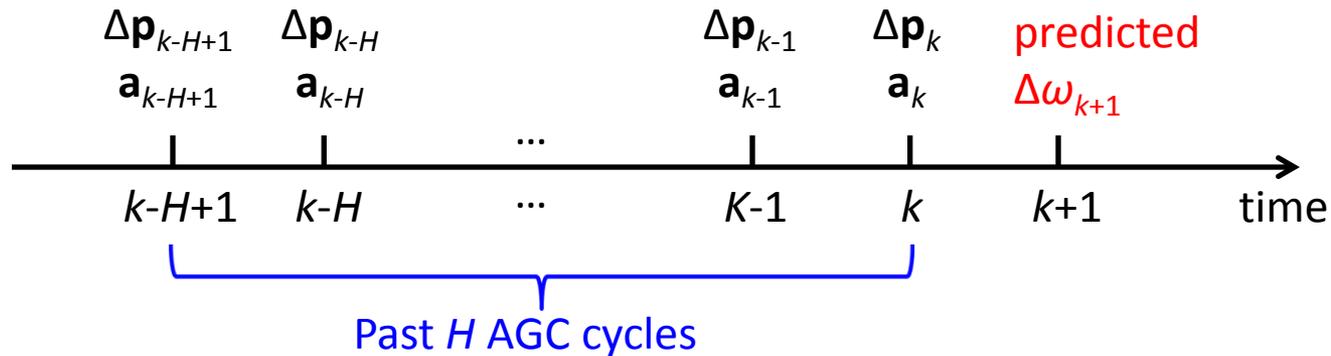
– Exhaustive search with prohibitive complexity

Regression

Laplace domain: $\Delta \omega = \Phi \cdot \Delta \mathbf{p} + \Phi \Lambda \mathbf{T} \cdot \mathbf{a}$



Time domain: $\Delta \omega_{k+1} = \sum_{i=0}^{H-1} \mathbf{u}_i \cdot \Delta \mathbf{p}_{k-i} + \mathbf{v}_i \cdot \mathbf{a}_{k-i}$



Regression

Laplace domain: $\Delta \omega = \Phi \cdot \Delta \mathbf{p} + \Phi \Lambda \mathbf{T} \cdot \mathbf{a}$



Time domain: $\Delta \omega_{k+1} = \sum_{i=0}^{H-1} \mathbf{u}_i \cdot \Delta \mathbf{p}_{k-i} + \mathbf{v}_i \cdot \mathbf{a}_{k-i}$

- Coefficients \mathbf{u} and \mathbf{v}
 - Trained using data generated by Laplace-domain model

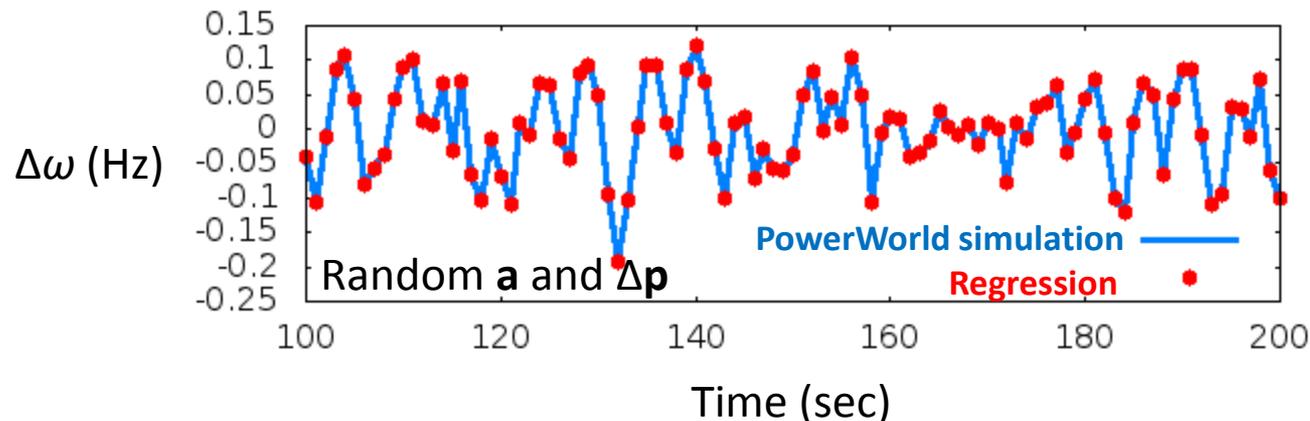
Regression

Laplace domain: $\Delta \omega = \Phi \cdot \Delta \mathbf{p} + \Phi \Lambda \mathbf{T} \cdot \mathbf{a}$



Time domain: $\Delta \omega_{k+1} = \sum_{i=0}^{H-1} \mathbf{u}_i \cdot \Delta \mathbf{p}_{k-i} + \mathbf{v}_i \cdot \mathbf{a}_{k-i}$

- Coefficients \mathbf{u} and \mathbf{v}
 - Trained using data generated by Laplace-domain model



Optimal Attack Algorithm

At prediction horizon h :

$$\Delta \omega_{k+h} = \begin{bmatrix} \mathbf{u}_{H-1} \\ \vdots \\ \mathbf{u}_h \\ \mathbf{u}_{h-1} \\ \vdots \\ \mathbf{u}_0 \end{bmatrix}^T \cdot \begin{bmatrix} \Delta \mathbf{p}_{k-H+h+1} \\ \vdots \\ \Delta \mathbf{p}_k \\ \Delta \hat{\mathbf{p}}_{k+1} \\ \vdots \\ \Delta \hat{\mathbf{p}}_{k+h} \end{bmatrix} + \begin{bmatrix} \mathbf{v}_{H-1} \\ \vdots \\ \mathbf{v}_h \\ \mathbf{v}_{h-1} \\ \vdots \\ \mathbf{v}_0 \end{bmatrix}^T \cdot \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{a}_{k+1} \\ \vdots \\ \mathbf{a}_{k+h} \end{bmatrix}$$

Eavesdropped past load changes

Predicted future load changes

Attack vectors to be scheduled

- Increasing h from 1, minimize and maximize $\Delta \omega_{k+h}$ until

$$\Delta \omega_{k+h} \notin (\Delta \omega_{\min}, \Delta \omega_{\max})$$

- Optimal, modulo approx err of regression
- Linear programming

Outline

- Motivation & Background
- Attack Model
- **Optimal Attack**
 - How to compute?
 - **Achievable?**
 - How to learn attack impact model?*
 - What prior information needed?*
- Simulations & Testbed Experiments

A Baseline Approach

$$\Delta \omega_{k+1} = \sum_{i=0}^{H-1} \mathbf{u}_i \cdot \Delta \mathbf{p}_{k-i} + \mathbf{v}_i \cdot \mathbf{a}_{k-i}$$

- Inject small attack vectors to collect training data to learn the coefficients
- Less stealthy

Passive Monitoring

$$\Delta \omega = \Phi \cdot \Delta \mathbf{p} + \Phi \Lambda \mathbf{T} \cdot \mathbf{a}$$

- Learn Φ , Λ from eavesdropped measurements

Passive Monitoring

$$\Delta \omega = \Phi \cdot \Delta \mathbf{p} + \Phi \Lambda \mathbf{T} \cdot \mathbf{a}$$

- Learn Φ , Λ from eavesdropped measurements
 - Φ : from $\Delta \omega$ and $\Delta \mathbf{p}$

$$\Delta \omega = \Phi \cdot \Delta \mathbf{p}$$

Passive Monitoring

$$\Delta \omega = \Phi \cdot \Delta \mathbf{p} + \Phi \Lambda \mathbf{T} \cdot \mathbf{a}$$

- Learn Φ , Λ from eavesdropped measurements

- Φ : from $\Delta \omega$ and $\Delta \mathbf{p}$

$$\Delta \omega = \Phi \cdot \Delta \mathbf{p}$$

- Λ : no training data

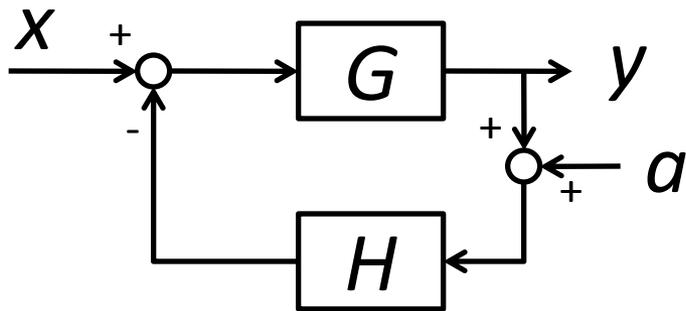
Passive Monitoring

$$\Delta \omega = \Phi \cdot \Delta \mathbf{p} + \Phi \Lambda \mathbf{T} \cdot \mathbf{a}$$

- Learn Φ , Λ from eavesdropped measurements
 - Φ : from $\Delta \omega$ and $\Delta \mathbf{p}$

$$\Delta \omega = \Phi \cdot \Delta \mathbf{p}$$

- Λ : no training data



$$y = \frac{G}{1 + GH} x - \frac{GH}{1 + GH} a$$

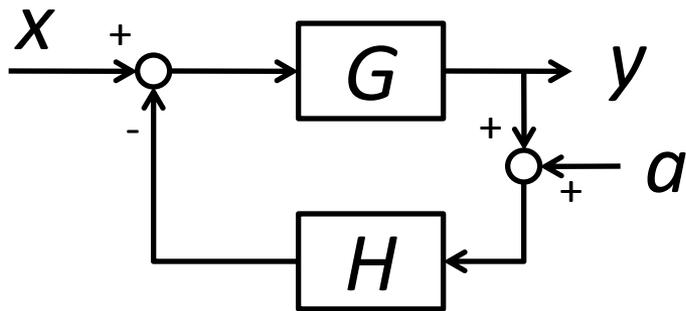
Passive Monitoring

$$\Delta \omega = \Phi \cdot \Delta \mathbf{p} + \Phi \Lambda \mathbf{T} \cdot \mathbf{a}$$

- Learn Φ , Λ from eavesdropped measurements
 - Φ : from $\Delta \omega$ and $\Delta \mathbf{p}$

$$\Delta \omega = \Phi \cdot \Delta \mathbf{p}$$

- Λ : no training data



$$y = \frac{G}{1 + GH} x - \frac{GH}{1 + GH} a$$

Estimated as a whole,
but not individual G and H

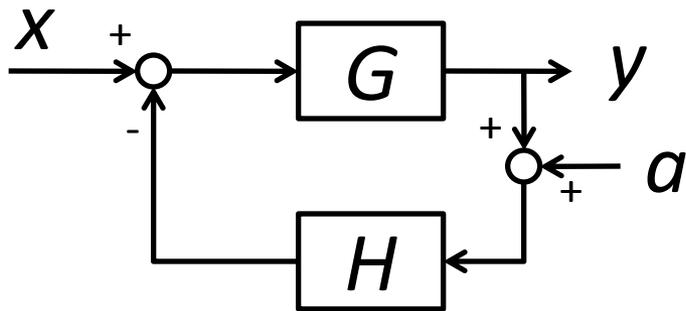
Passive Monitoring

$$\Delta \omega = \Phi \cdot \Delta \mathbf{p} + \Phi \Lambda \mathbf{T} \cdot \mathbf{a}$$

- Learn Φ , Λ from eavesdropped measurements
 - Φ : from $\Delta \omega$ and $\Delta \mathbf{p}$

$$\Delta \omega = \Phi \cdot \Delta \mathbf{p}$$

- Λ : no training data



$$y = \frac{G}{1 + GH} x - \frac{GH}{1 + GH} a$$

Estimated as a whole,
but not individual G and H

?

Passive Monitoring

$$\Delta \omega = \Phi \cdot \Delta \mathbf{p} + \Phi \Lambda \mathbf{T} \cdot \mathbf{a}$$

- Learn Φ , Λ from eavesdropped measurements
 - Φ : from $\Delta \omega$ and $\Delta \mathbf{p}$

$$\Delta \omega = \Phi \cdot \Delta \mathbf{p}$$

- Λ : from $\Delta \omega$, \mathbf{z} in normal state and 4 system constants:
 - ACE weight parameters α and β : **public**
 - Load damping constant and total inertia of generators

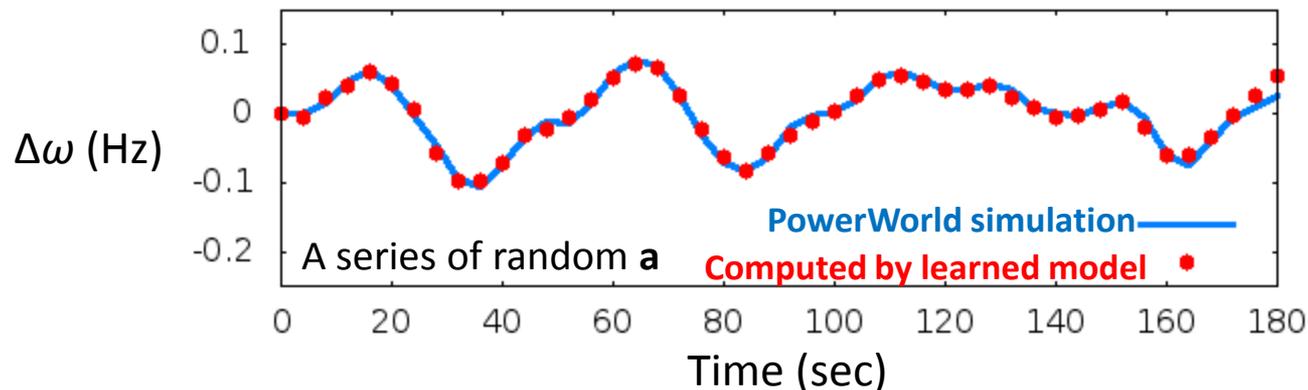
Passive Monitoring

$$\Delta \omega = \Phi \cdot \Delta \mathbf{p} + \Phi \Lambda \mathbf{T} \cdot \mathbf{a}$$

- Learn Φ , Λ from eavesdropped measurements
 - Φ : from $\Delta \omega$ and $\Delta \mathbf{p}$

$$\Delta \omega = \Phi \cdot \Delta \mathbf{p}$$

- Λ : from $\Delta \omega$, \mathbf{z} in normal state and 4 system constants:
 - ACE weight parameters α and β : public
 - Load damping constant and total inertia of generators

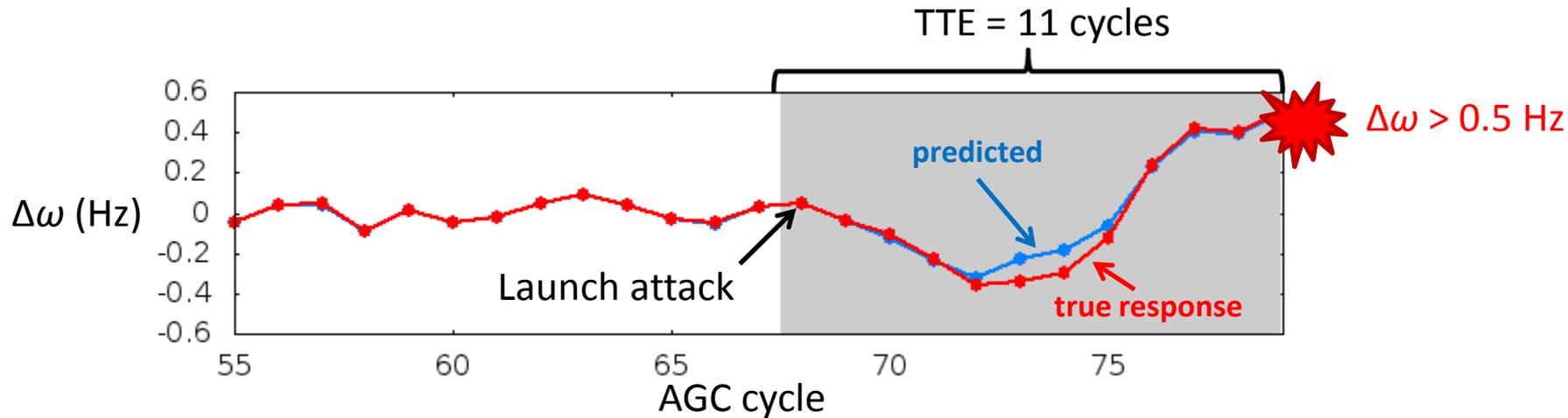


Outline

- Motivation and Background
- Attack Model
- Optimal Attack
- **Simulations & Testbed Experiments**

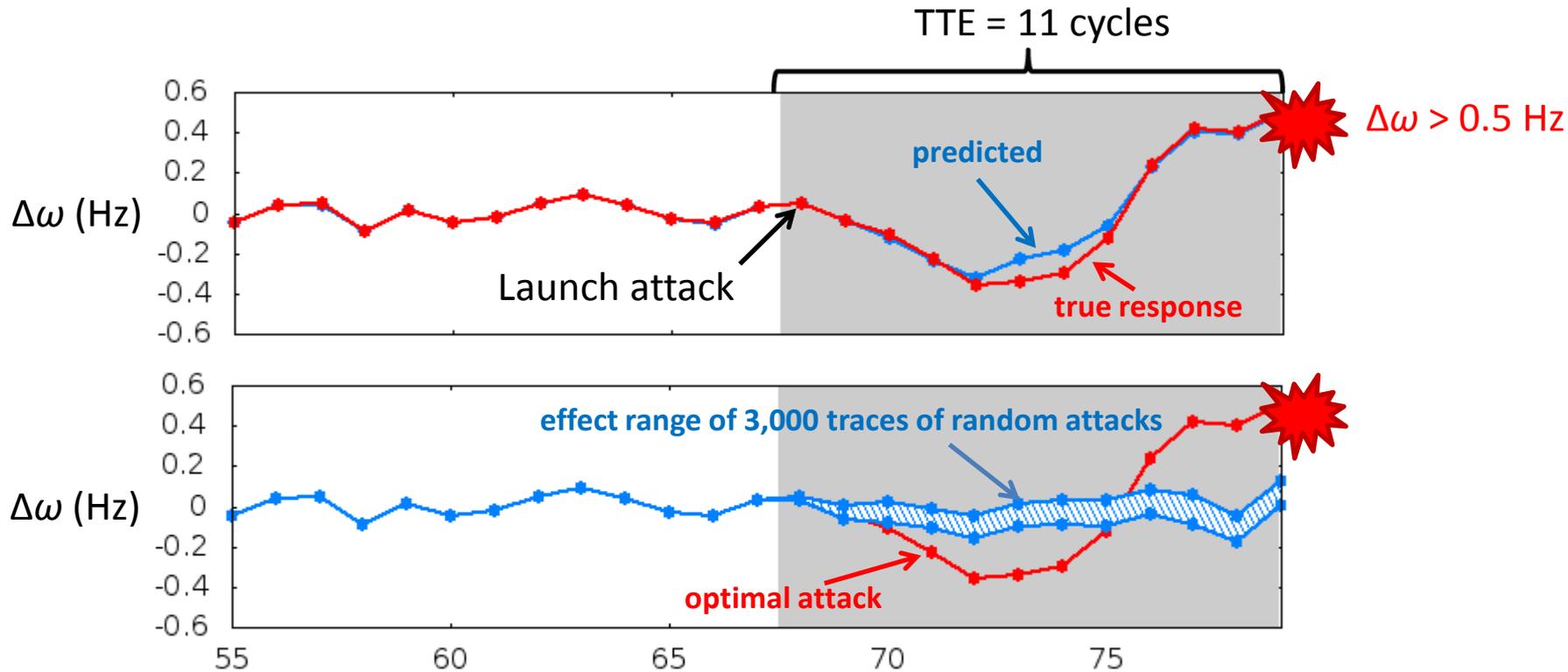
PowerWorld Simulations

37-bus 3-area grid, 81 sensors (all compromised), frequency safety range: (-0.5, +0.5) Hz



PowerWorld Simulations

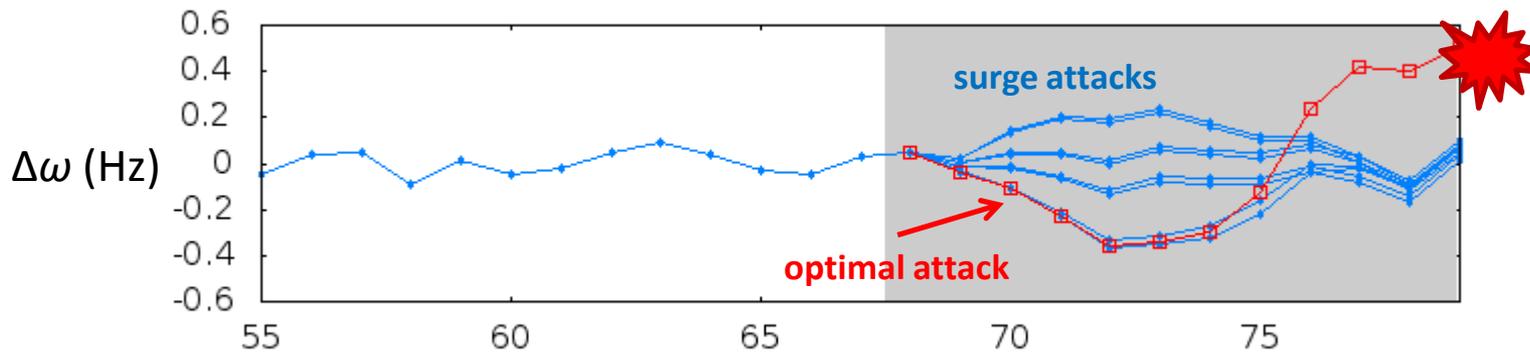
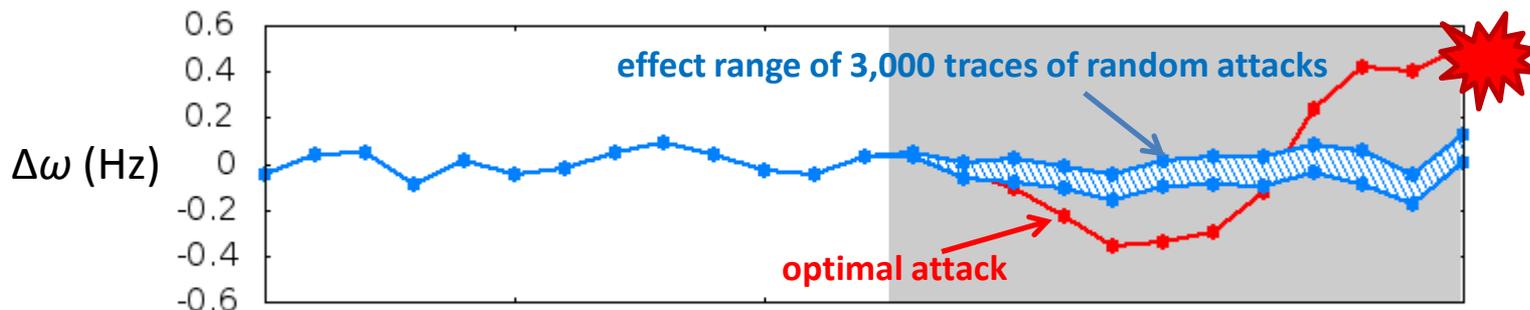
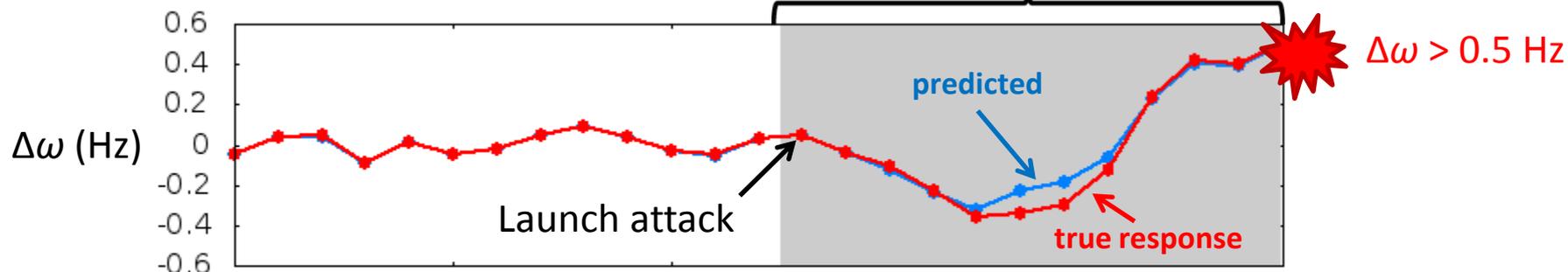
37-bus 3-area grid, 81 sensors (all compromised), frequency safety range: (-0.5, +0.5) Hz



PowerWorld Simulations

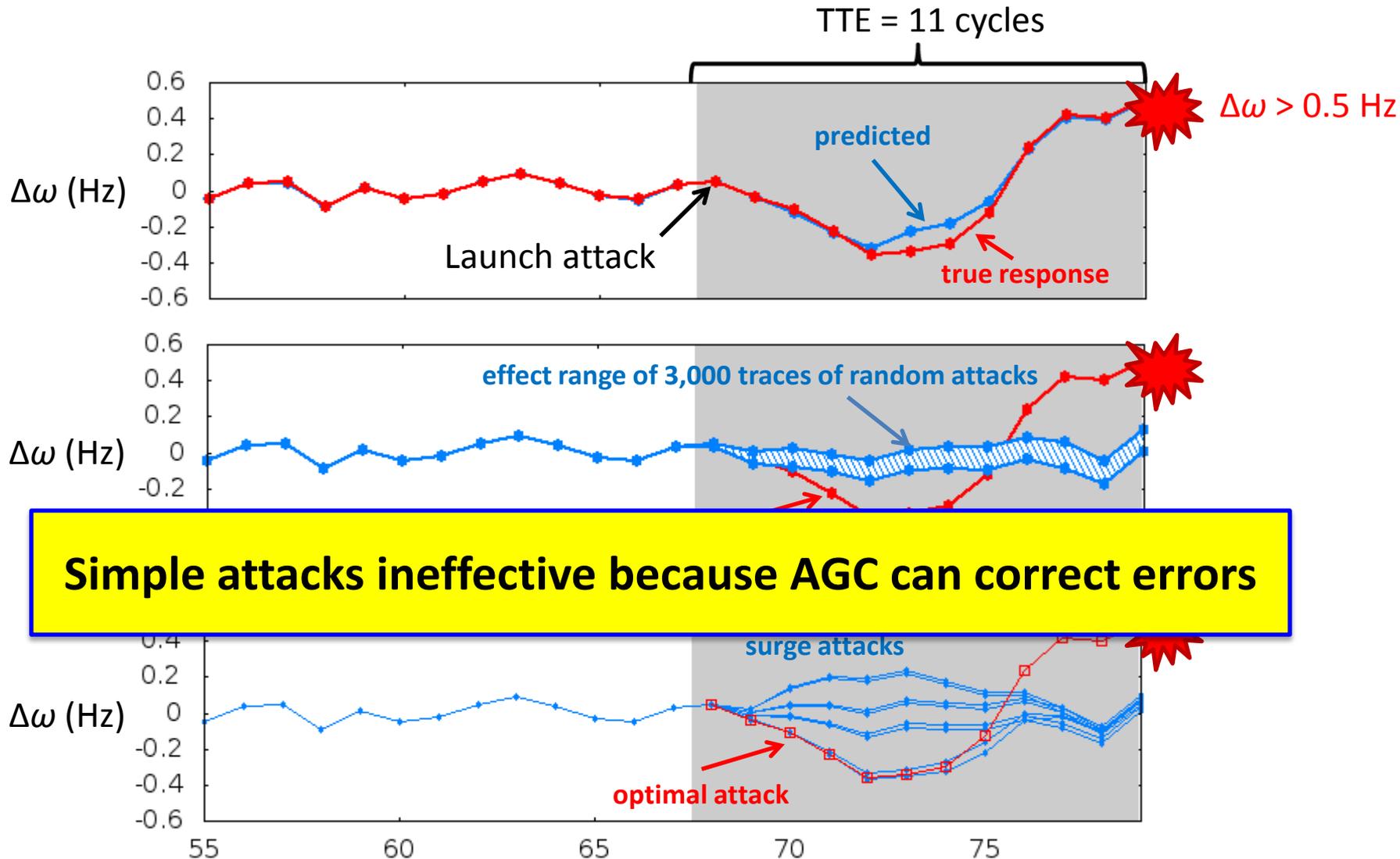
37-bus 3-area grid, 81 sensors (all compromised), frequency safety range: (-0.5, +0.5) Hz

TTE = 11 cycles



PowerWorld Simulations

37-bus 3-area grid, 81 sensors (all compromised), frequency safety range: (-0.5, +0.5) Hz



Simple attacks ineffective because AGC can correct errors

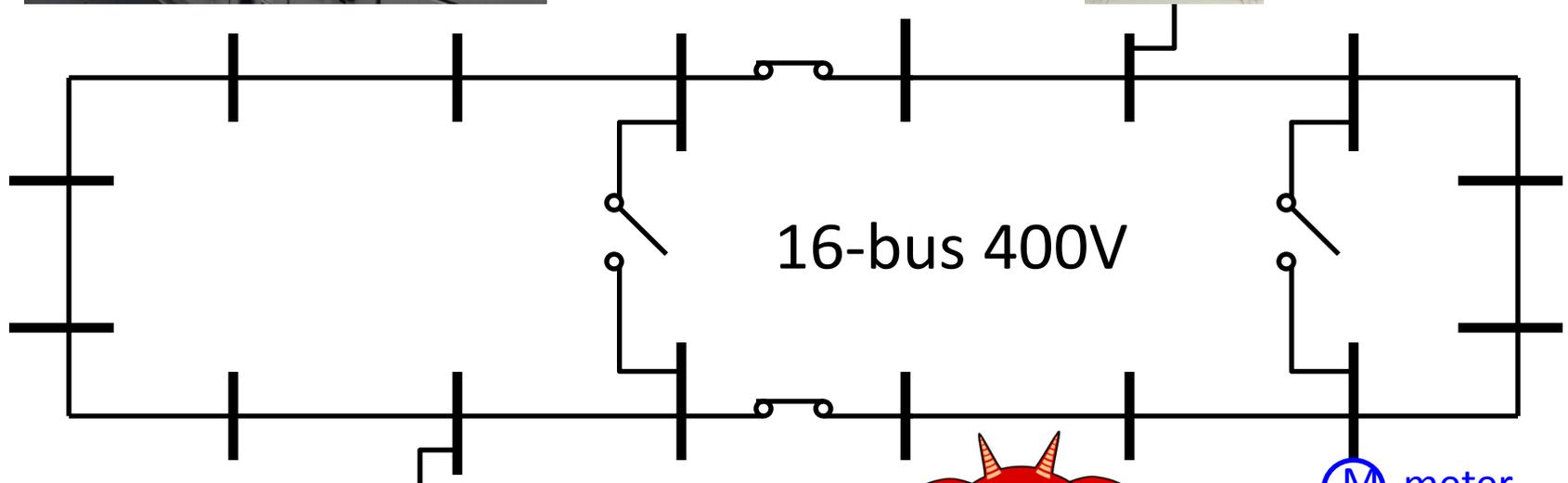
Microgrid Testbed



Microgrid
switchboard



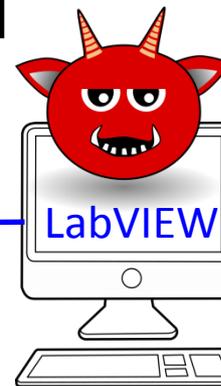
variable load



13.5kVA
generator



command

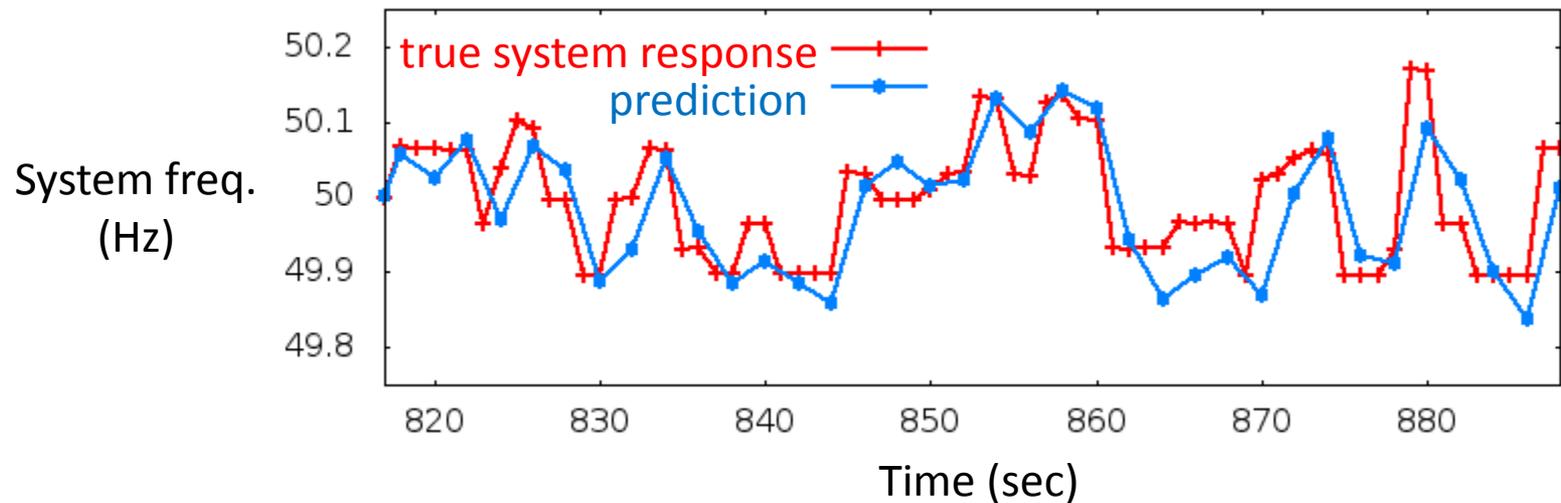
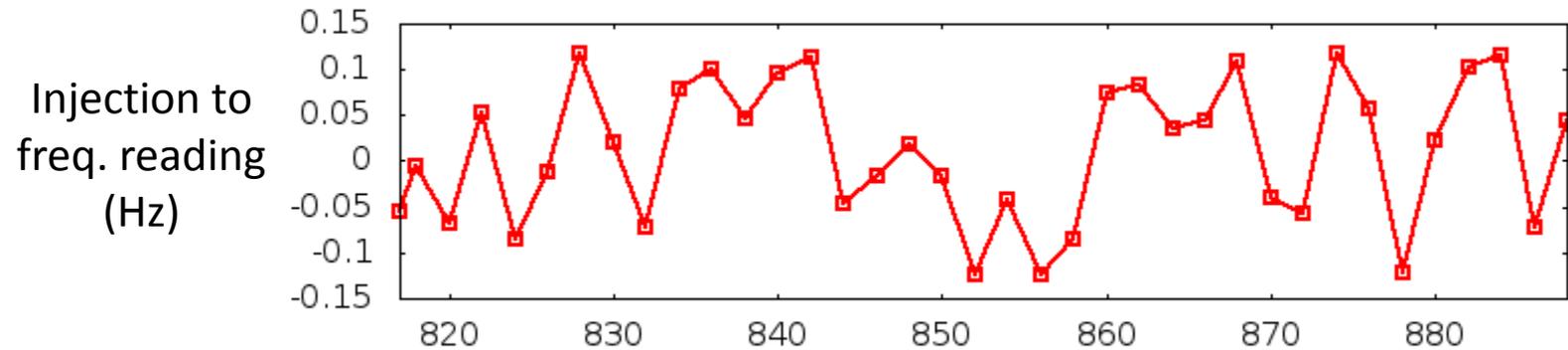


frequency readings
over Modbus/TCP
20/22

M meter

Accuracy of Learned Attack Model

A random bounded injection sequence to avoid damage to testbed



From learned model, 30 seconds to achieve 0.5 Hz deviation

Conclusion

- **FDI attack against AGC**
 - Attack impact model
 - Learned using data in normal state
 - Minimize time-to-emergency
- **Evaluation**
 - PowerWorld simulations
 - Experiments on a real power system
- **Ongoing work**
 - Attack detection, identification (which measurements compromised?), mitigation