Cyber-Security of Networked Control Systems

Karl Henrik Johansson
ACCESS Linnaeus Center
Royal Institute of Technology, Sweden

Joint work with Henrik Sandberg, André Teixeira, Kin C. Sou

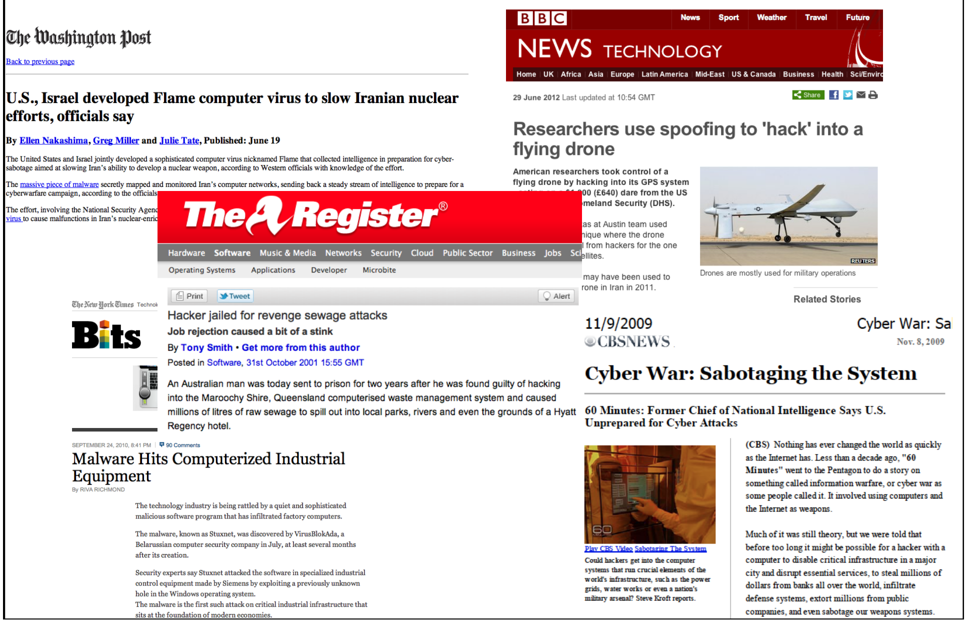
Workshop on Cooperative Estimation and Control over Networks
Academy of Mathematics and Systems Science, CAS, Beijing, Jul 3, 2012

ACCESS Linnaeus Center

- ACCESS was established at KTH from an Excellence Grant from the Swedish Research Council
- Developed into **one of Europe's largest university research centers** in networked systems
  - 36 faculty, 20 postdocs, >100 PhD students
  - Basic funding from VR on 1.3 MEUR per year
  - Total research budget 2010 over 12 MEUR
- Graduate school with **>40 graduated PhD's**
- **Faculty renewal** and mobility programs
- Extensive international and industrial collaborations
- External communication and dissemination

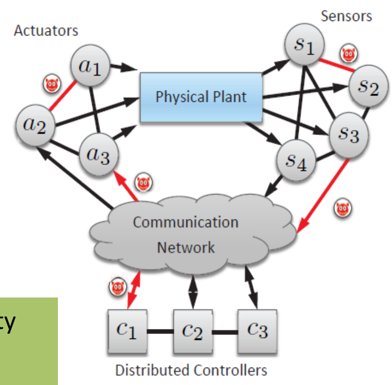
2

Recent Cyber-Attacks on Control Systems



Cyber-Security of Networked Control Systems

- Networked control systems are to a growing extent based on **open communication and software technology**
- Leads to **increased vulnerability** to cyber-threats with many potential points of attacks



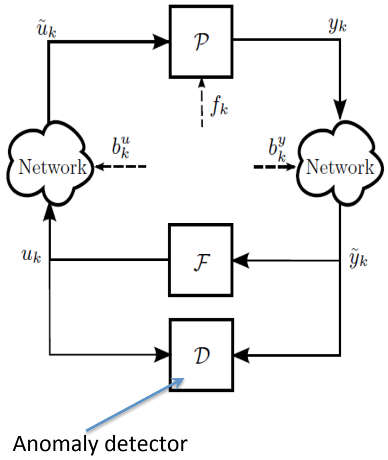
- How to model attacks?
- How to measure vulnerability?
- How to compute consequences?
- How to design secure control systems

- Traditional computer and information security do not provide answers these questions
- Need for a theory for secure control systems

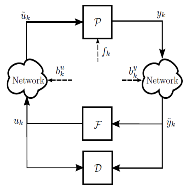
Outline

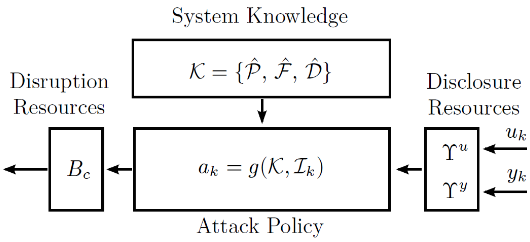
- Introduction
- Attack model for control systems
- Attack on power network state estimator
- Stealthy minimum-effort attacks
- Security index
- Conclusions
- Biography

Networked Control System under Attack



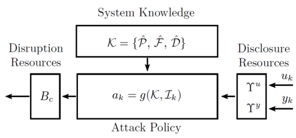
Adversary Model

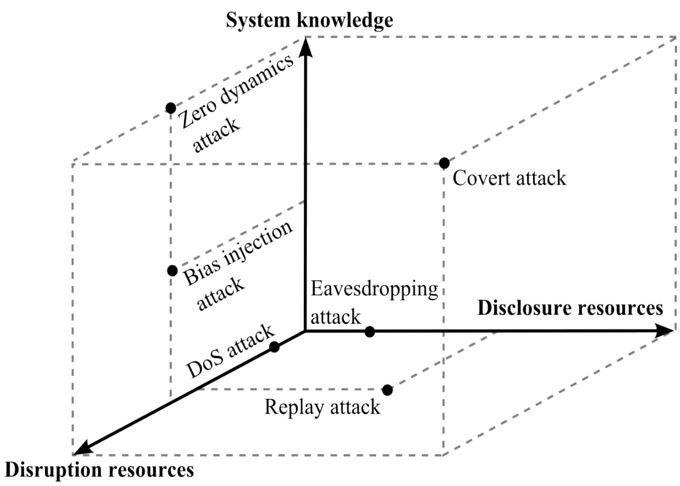




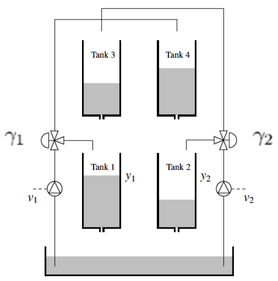
- Adversary constrained by limited resources
- Attack policy depends on adversary goals and constraints

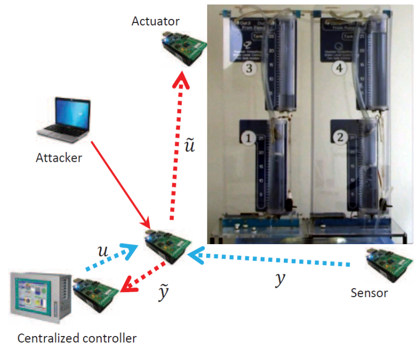
Attack Space





Experimental Set-Up





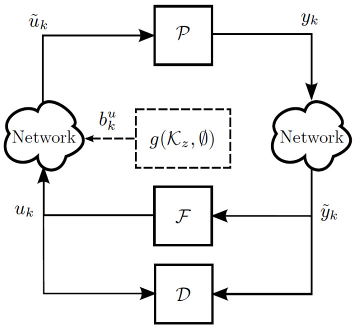
$$\frac{dx}{dt} = \begin{bmatrix} -\frac{1}{T_1} & 0 & \frac{A_3}{A_1 T_3} & 0 \\ 0 & -\frac{1}{T_2} & 0 & \frac{A_4}{A_2 T_4} \\ 0 & 0 & -\frac{1}{T_3} & 0 \\ 0 & 0 & 0 & -\frac{1}{T_4} \end{bmatrix} x + \begin{bmatrix} \frac{\gamma_1 k_1}{A_1} & 0 \\ 0 & \frac{\gamma_2 k_2}{A_2} \\ 0 & \frac{(1-\gamma_2)k_2}{A_3} \\ \frac{(1-\gamma_1)k_1}{A_4} & 0 \end{bmatrix} u$$

$$y = \begin{bmatrix} k_c & 0 & 0 & 0 \\ 0 & k_c & 0 & 0 \end{bmatrix} x$$

Quadruple-tank process has non-minimum-phase zero if $0 < \gamma_1 + \gamma_2 < 1$

[J, 2000]

Zero Dynamics Attack



- Zero dynamics are characterized by:

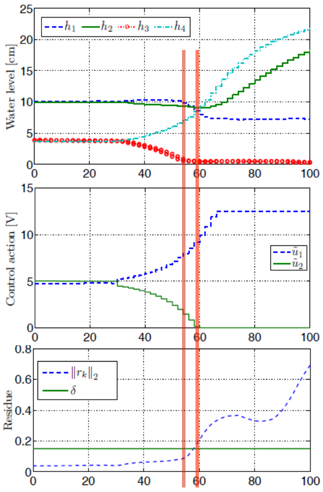
$$\begin{bmatrix} \nu I - A & -B \\ C & 0 \end{bmatrix} \begin{bmatrix} x_0 \\ g \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

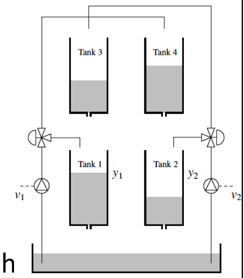
- Suggests attack on actuators with policy:

$$a_k = g\nu^k$$

- If the zero is unstable, then the plant state can be moved by this attack without detection
- Requires system knowledge (zero dynamics) but no disclosure resources

Experimental Validation





- **Attack goal:** Empty Tank 3
- Zero dynamics attack on both actuators starts at t=30s
- Tank 3 becomes empty at t=55s
- The attack is detected at t=58s
- Actuator 2 saturates at t=60s

Teixeira et al, HiCoNS, 2012

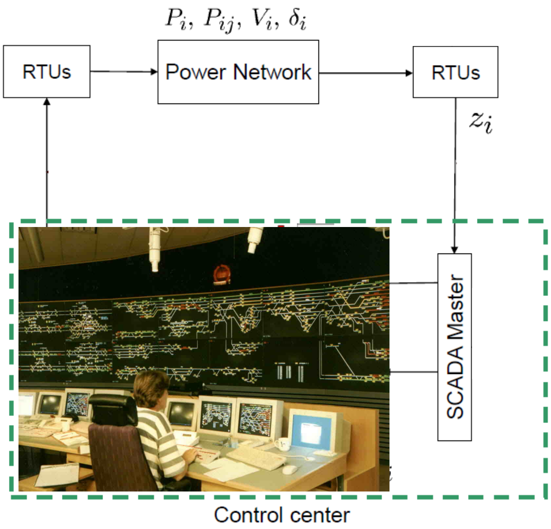
Outline

- Introduction
- Attack model for control systems
- Attack on power network state estimator
- Stealthy minimum-effort attacks
- Security index
- Conclusions
- Biography

Motivation

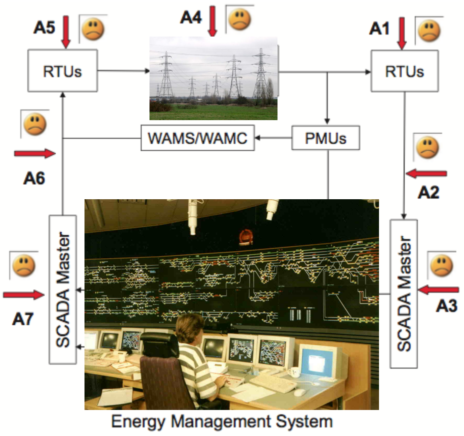
- Northeast blackout Aug 14, 2003: 55 million people affected
- Software bug in energy management system **stalled alarms in state estimator for over an hour**
- Cyber-attacks against the power network control systems with similar consequences pose a substantial threat

SCADA/EMS Systems



(SCADA/EMS = Supervisory Control and Data Acquisition/Energy Management Systems)

Attacks on Power Systems

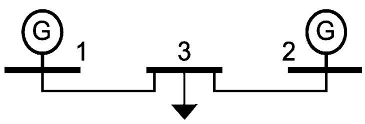


- Many attack opportunities
  - Sensor and actuators
  - Communication systems
  - Software systems (e.g., control)
  - Human operators
  - Physical infrastructure
- How strengthen the systems against cyber-attacks?

SCADA = Supervisory Control and Data Acquisition **EMS** = Energy Management System **WAMS** = Wide Area Monitoring System
RTUs = Remote Terminal Units (Sensors/Actuators) **PMUs** = Phasor Measurement Units (Sensors)

(Static) Power Network Model

- Local states at bus i:
  - θ_i – phase angle
  - V_i – voltage magnitude



- Active and reactive power injections:

$$\begin{aligned} P_i &= V_i \sum_{j\in N_i} V_j \left(G_{ij}\cos\theta_{ij} + B_{ij}\sin\theta_{ij}\right) \\ Q_i &= V_i \sum_{j\in N_i} V_j \left(G_{ij}\sin\theta_{ij} - B_{ij}\cos\theta_{ij}\right) \end{aligned}$$

- Active and reactive power flows:

$$\begin{aligned} P_{ij} &= V_i^2(g_{si}+g_{ij}) - V_iV_j\left(g_{ij}\cos\theta_{ij} + b_{ij}\sin\theta_{ij}\right) \\ Q_{ij} &= -V_i^2(b_{si}+b_{ij}) - V_iV_j\left(g_{ij}\sin\theta_{ij} - b_{ij}\cos\theta_{ij}\right) \end{aligned}$$

where

$$\theta_{ij} = \theta_i - \theta_j$$

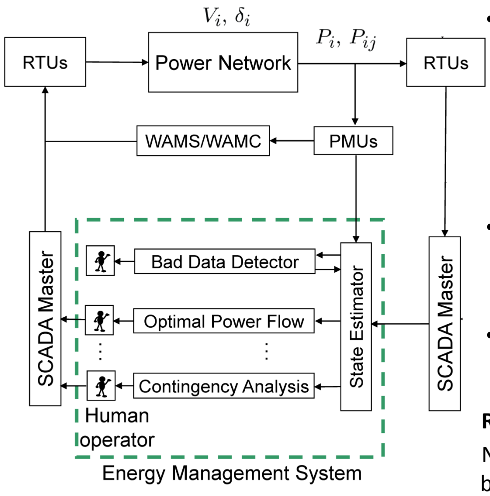
- Measurement model:

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

  - $x \in \mathbb{R}^n$: network states
  - $z \in \mathbb{R}^m$: power flow measurements
  - ϵ : measurement noise

Static model because the power grid time constant ~10 ms is beyond existing measurement technology. Typical sampling time ~1 s.

Energy Management System for Power Networks



- SCADA-EMS provides power network state information to
  - Identify faulty equipment
  - Optimize power flows
  - Analyze reliability (contingency)
  - Etc
- Large system with slow sampling
  - 100-1 000's of RTUs sampled in sec's
  - 10K-40K measurements
- Decisions taken by human operators

Remark

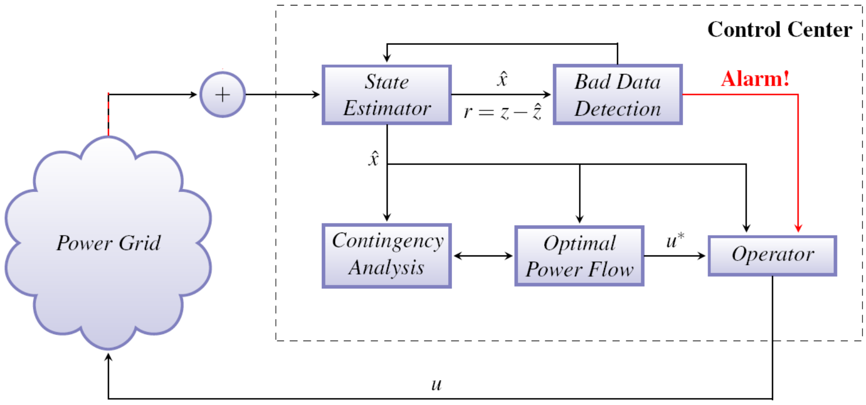
New WAMCs based on high-rate PMUs are better protected but constitute only a small portion of the overall network

WAMC = Wide Area Monitoring and Control System

RTUs = Remote Terminal Units (Sensors/Actuators)

PMUs = Phasor Measurement Units (Sensors)

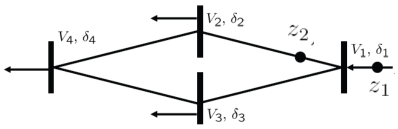
Energy Management System



- The **state estimator** has a crucial role in the EMS
- If the **bad data detector** identifies a faulty sensor, the corresponding measurement is removed from the state estimator
- Bad data detection is typically done under the assumption of **uncorrelated faults**, which does not hold for intelligent attacks

State Estimator

- Steady-state models:



$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} \frac{V_1V_2}{X_{12}} \sin(\delta_1 - \delta_2) + \frac{V_1V_3}{X_{13}} \sin(\delta_1 - \delta_3) \\ \frac{V_1V_2}{X_{12}} \sin(\delta_1 - \delta_2) \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = h(x) + e \in \mathbb{R}^m$$

- WLS estimates of bus phase angles δ_i (in vector $\widehat{x}$):

$$\hat{x}^{k+1} = \hat{x}^k + (H_k^T R^{-1} H_k)^{-1} H_k^T R^{-1}(z - h(\hat{x}^k))$$

$$H_k := \frac{\partial h}{\partial x}(\hat{x}_k) \qquad R := \mathbf{E} ee^T$$

- Linear DC approximation (≈ ML estimate):

$$\hat{x} = (H^T R^{-1} H)^{-1} H^T R^{-1} z \qquad \boldsymbol{H} := \left.\frac{\partial \boldsymbol{h(x)}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=0}$$

E.g., [Schweppe and Wildes, 1970; Abur and Exposito, 2004]

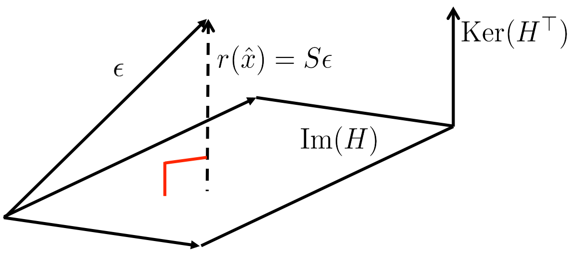
Bad Data Detector

$$H = \left.\frac{\partial h(x)}{\partial x}\right|_{x=\hat{x}}$$

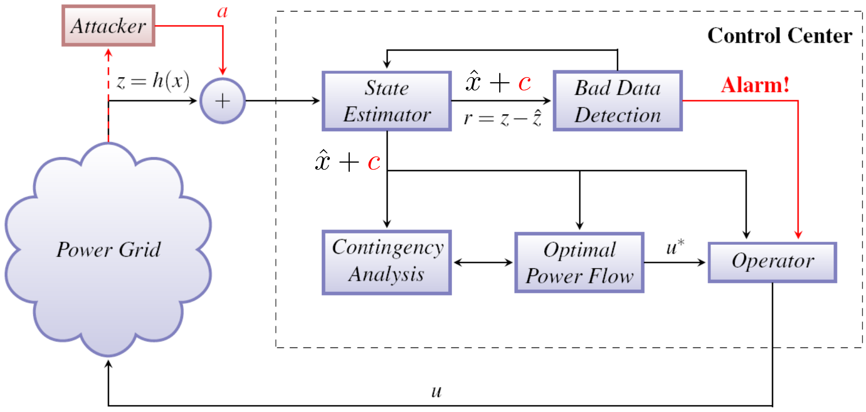
- Today's BDD is based on measurement residual $r(\hat{x}) = z - h(\hat{x})$

$$\|Wr(\hat{x})\|_p \underset{H_1}{\overset{H_0}{\lessgtr}} \tau$$

- For the Gauss-Newton method: $r(\hat{x}) \approx (I - H(H^\top H)^{-1} H^\top)\epsilon = S\epsilon$
- Note that $S = \mathbf{P}_{\mathrm{Ker}(H^\top)}$ is the orthogonal projection onto $\mathrm{Ker}(H^\top)$
- Can be exploited by an attacker



Attack Model

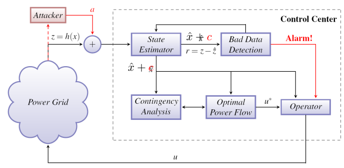


- **Scenario:** Attacker injects **malicious data** a to corrupt analog measurements in the power grid, in order to change state estimates without generating bad data detection alarm
- How characterize the set of **undetectable** malicious data a?

Outline

- Introduction
- Attack model for control systems
- Attack on power network state estimator
- Stealthy minimum-effort attacks
- Security index
- Conclusions
- Biography

Bad-Data Detection and Stealthy Attacks



- Bad-data detection trigger alarm when residual r is large

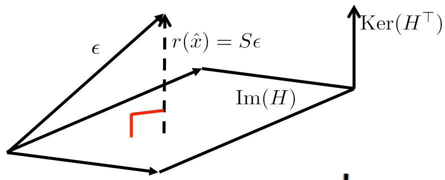
$$r := z - \hat{z} = z - H\hat{x} = z - H(H^T R^{-1} H)^{-1} H^T R^{-1} z$$

- Characterization of undetectable malicious data a

$$z_a := z + a$$

$$a = Hc \in \mathrm{Im}(H)$$

$$r = z - \hat{z} = z_a - \hat{z}_a$$



- The attacker has a lot of freedom in the choice of a!
- Attacker likely to seek sparse solutions a , i.e., manipulate only few measurements

[Liu *et al.*, 2009]

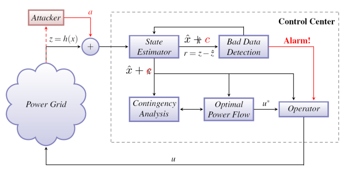
Stealthy Minimum-Effort Attack

- Attack single measurement z_k

$$\min_a \|a\|_2$$

$$\text{s.t. } a \in \mathcal{U} \cap \mathcal{G}_k \cap \mathcal{C}$$

- $\mathcal{U} = \mathrm{Im}(H)$
- $\mathcal{G}_k = \{a \in \mathbb{R}^m : a_k = 1\}$
- $\mathcal{C} = \mathbb{R}^m$



- Optimal attack

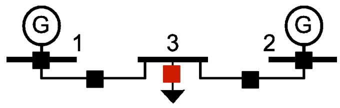
$$a^* = K_{kk}^{-1} K e_k$$

$$K = HH^\dagger$$

- a^* is typically not sparse, so many sensors need to be corrupted
- Consider **0-norm instead of 2-norm**

Stealthy Minimum-Effort Attack

$$\min_{a} \|a\|_0$$
$$\text{s.t. } a \in \mathcal{U} \cap \mathcal{G}_k \cap \mathcal{C}$$



- P_3 is the target measurement
- A few possible attacks:
  - ~~$\{P_3\}, \{P_3, \star\}$~~ not stealthy
  - $\{P_1, P_{13}, P_3\}$ minimum effort
  - $\{P_2, P_{23}, P_3\}$
  - $\{P_1, P_{13}, P_3, P_{23}, P_2\}$

Security Index ρ_k

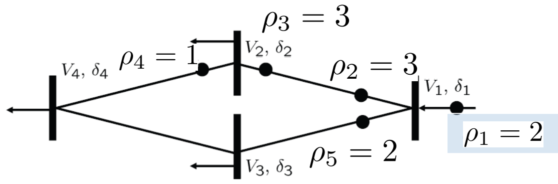
- Security index for measurement k: $\rho_k = \|a^*\|_0$
  - a^* is the optimal solution of

$$\min_{a} \|a\|_0$$
$$\text{s.t. } a \in \mathcal{U} \cap \mathcal{G}_k \cap \mathcal{C}$$

- $\mathcal{U} = \text{Im}(H)$ Stealthy
- $\mathcal{G}_k = \{a \in \mathbb{R}^m : a_k = 1\}$ Corrupted
- $\mathcal{C} = \{a \in \mathbb{R}^m : a_i = 0 \quad \forall i \in \mathcal{P}\}$ Protected

- ρ_k is the minimum number of measurements to manipulate for a successful attack

Example



- Sparse attack corresponding to ρ_k:

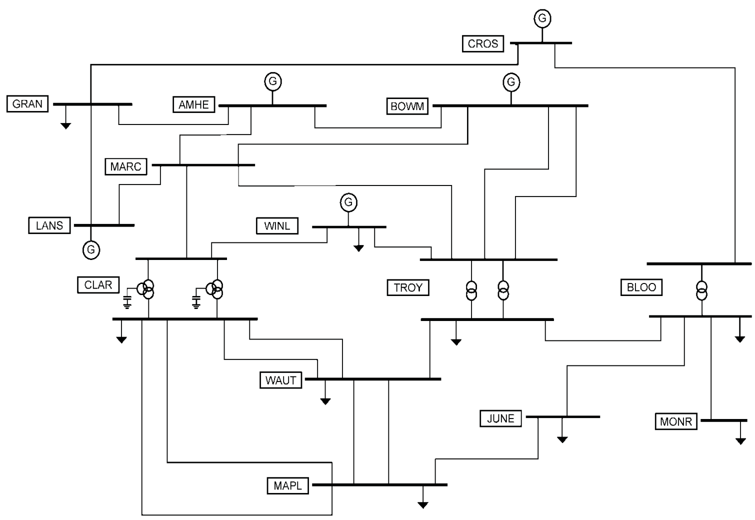
$$\begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

- Compare with the "hat matrix":

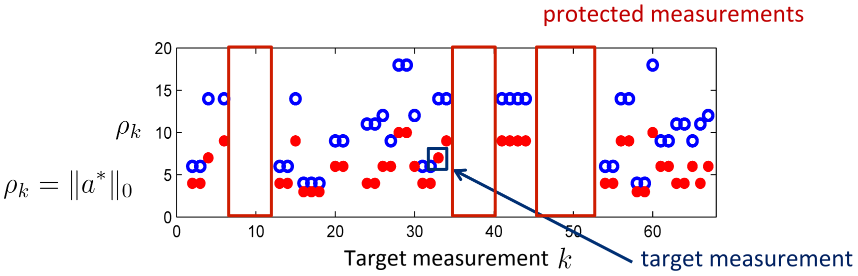
$$\begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \\ \hat{z}_3 \\ \hat{z}_4 \\ \hat{z}_5 \end{pmatrix} = \underbrace{\begin{pmatrix} 0.60 & 0.20 & -0.20 & 0 & 0.40 \\ 0.20 & 0.40 & -0.40 & 0 & -0.20 \\ -0.20 & -0.40 & 0.40 & 0 & 0.20 \\ 0 & 0 & 0 & 1.00 & 0 \\ 0.40 & -0.20 & 0.20 & 0 & 0.60 \end{pmatrix}}_{=H(H^T R^{-1} H)^{-1} H R^{-1}} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \end{pmatrix}$$

- Hat matrix misleading for judging sparsity of attacks!

VIKING 40-bus Benchmark



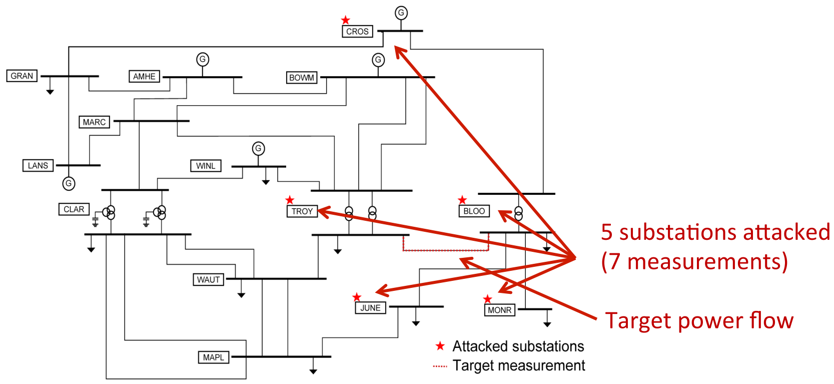
VIKING Benchmark: Security Index



Existing measurement configuration

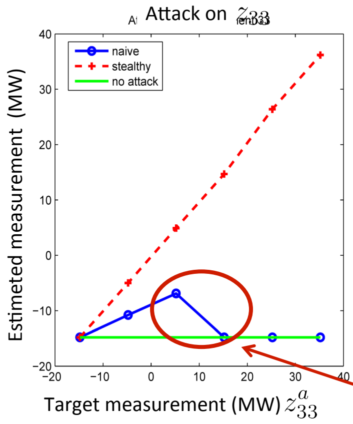
Extended measurements configuration

VIKING Benchmark: Experimental Results



- Target measurement: flow between TROY and BLOO, z_{33}
- Nonlinear models are used by the state estimator and bad data detector
- Attacker knows the linear DC model accurately

VIKING Benchmark: Experimental Results



$$z^a = z + a$$

Table 1: Results from the stealthy attack for large bias

Target bias, a_{33}	False value (MW), z_{33}^a	Estimate (MW), $\hat{z}_{33}^a$	#BDD Alarms
0	-14.8	-14.8	0
50	35.2	36.2	0
100	85.2	86.7	0
150	135.2	137.5	0
200	185.2	-	-

- 150 MW was not detected (56% of nominal value)
- State estimator did not converge for 200 MW

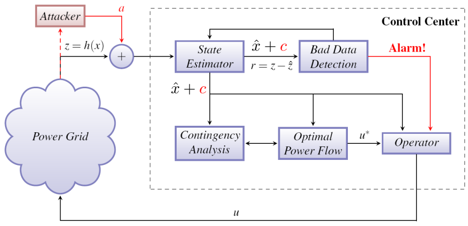
Bad data detected and removed

Outline

- Introduction
- Attack model for control systems
- Attack on power network state estimator
- Stealthy minimum-effort attacks
- Security index
- Conclusions
- Biography

Conclusions

- **Cyber-attack models** for networked control systems
- Undetectable false-data attack against power systems state estimator possible, both in theory and practice
- New **security index** ρ_k to estimate vulnerabilities
- Suggests locations of counter measures
- Many open problems in secure networked control theory



http://www.ee.kth.se/~kallej

Bibliography

- A. Giani, S. Sastry, K. H. Johansson, and H. Sandberg. The VIKING project: an initiative on resilient control of power networks. In Proc. 2nd Int. Symp. on Resilient Control Systems, pages 31–35, Idaho Falls, ID, USA, August 2009.
- H. Sandberg, A. Teixeira, and K. H. Johansson, "On Security Indices for State Estimators in Power Networks". In First Workshop on Secure Control Systems, Stockholm, Sweden, 2010.
- A. Teixeira, H. Sandberg, and K. H. Johansson, "Networked Control Systems under Cyber Attacks with Applications to Power Networks". In American Control Conference, Baltimore, MD, USA, 2010.
- G. Dan and H. Sandberg, "Stealth attacks and protection schemes for state estimators in power systems," in Proceedings of IEEE International Conference of Smart Grid Communications, 2010.
- A. Teixeira, S. Amin, H. Sandberg, K. H. Johansson, and S. Sastry, "Cyber Security Analysis of State Estimators in Electric Power Systems". In Proceedings of the 49th Conference on Decision and Control, Atlanta, GA, USA, 2010.
- A. Teixeira, G. Dán, H. Sandberg, and K. H. Johansson, "Cyber Security Study of a SCADA Energy Management System: Stealthy Deception Attacks on the State Estimator". In 18th IFAC World Congress, Milan, Italy, 2011.
- O. Vuković, K. C. Sou, G. Dán, H. Sandberg, "Network-layer Protection Schemes against Stealth Attacks on State Estimators in Power Systems", in Proceedings of IEEE International Conference of Smart Grid Communications, 2011.
- K. C. Sou, H. Sandberg, and K. H. Johansson. Electric power network security analysis via minimum cut relaxation. In *Proceedings of the 50th IEEE Conference on Decision and Control,* December 2011.
- A. Teixeira, "Toward Secure and Reliable Networked Control Systems", Licentiate Thesis, KTH, Sweden, 2011.
- A. Teixeira, H. Sandberg, G. Dán, and K. H. Johansson. "Optimal Power Flow: Closing the Loop over Corrupted Data". American Control Conference, 2012. Submitted.

http://www.ee.kth.se/~kallej

Bibliography (cont'd)

- F. C. Schweppe and J. Wildes. Power system static-state estimation, part I: Exact model. IEEE Transactions on Power Apparatus and Systems, 89(1):120–125, January 1970.
- F. F. Wu. Power system state estimation: a survey. Int. J. Elec. Power and Energy Systems, April 1990.
- M. Shahidehpour, F. Tinney, and Y. Fu. Impact of security on power systems operation. Proceedings of the IEEE, 93(11):2013–2025, nov 2005a.
- Y. Liu, M. K. Reiter, and P. Ning. False data injection attacks against state estimation in electric power grids. In Proc. 16th ACM Conf. on Computer and Communications Security, pages 21–32, New York, NY, USA, 2009.
- L. Jia, R. J. Thomas, and L. Tong. Malicious data attack on real-time electricity market. In Proc. of IEEE ICASSP, May 2011.
- O. Kosut, L. Jia, R. Thomas, and L. Tong. Malicious data attacks on smart grid state estimation: Attack strategies and countermeasures. In Proc. of IEEE SmartGridComm, October 2010.
- L. Xie, Y. Mo, and B. Sinopoli. False data injection attacks in electricity markets. In First IEEE International Conference on Smart Grid Communications, Oct. 2010

- A. Abur and A.G. Exposito. Power System State Estimation: Theory and Imple- mentation. Marcel-Dekker, 2004.
- A. Monticelli. State Estimation in Electric Power Systems: A Generalized Approach. Kluwer Academic Publishers, 1999.
- P. Kundur. Power System Stability and Control. McGraw-Hill Professional, 1994.
- E. A. Blood, From Static to Dynamic Electric Power Network State Estimation: The Role of Bus Component Dynamics, PhD Thesis, ECE, CMU, 2011