Security Challenges in Cyber Physical Systems

Slides by David Brumley, Adrian Perrig and Bruno Sinopoli

Presented by Nicolas Christin

Carnegie Mellon CyLab

Smart Grid Components

Advanced Metering Infrastructure

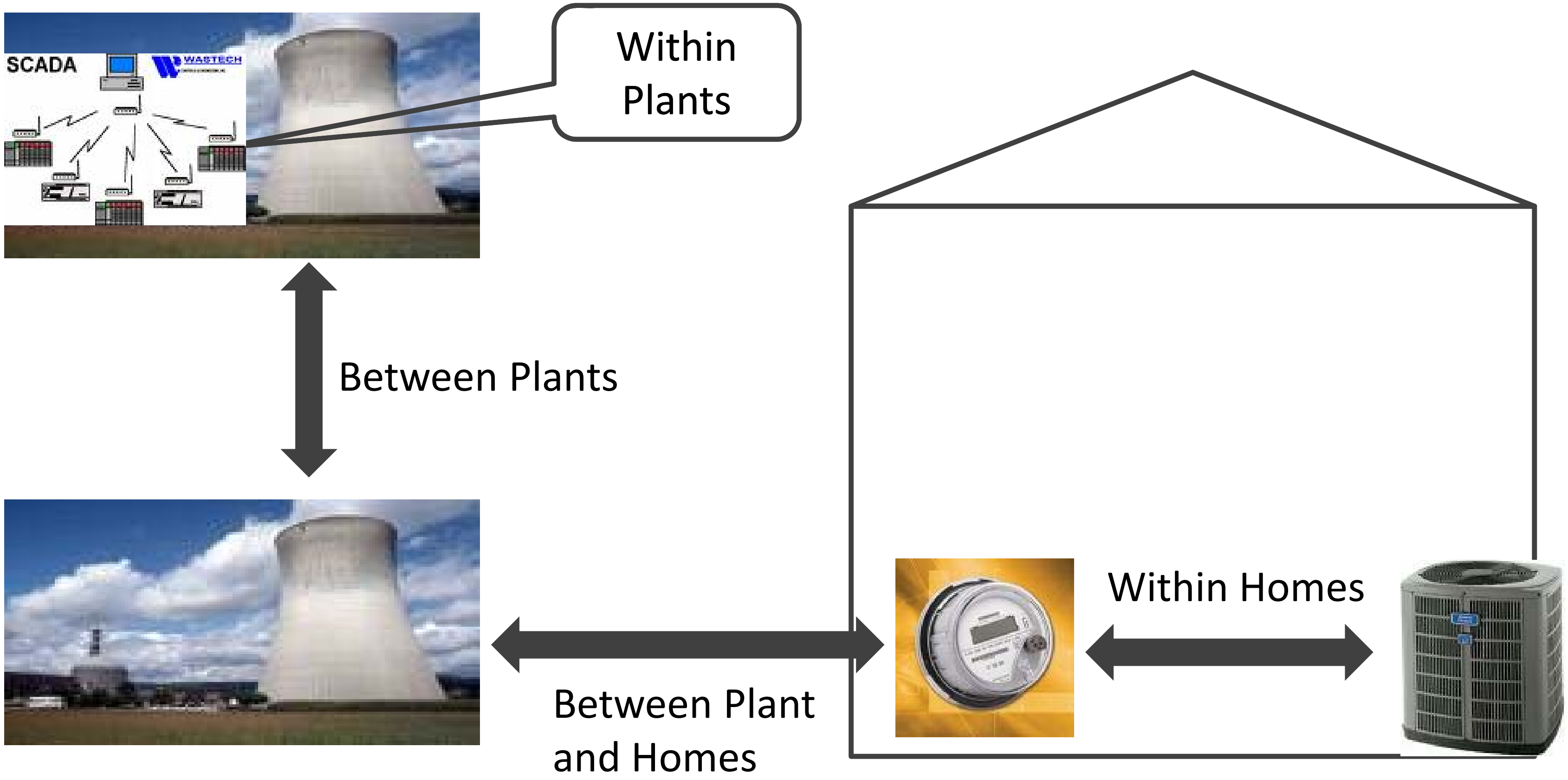
Home Area Network

SCADA

Distributed Control Systems

Physical Plants

Unprecedented Levels of Communication and Coordination



The Smart Grid: An Attractive Target For Attackers

”One of the easiest penetration tests I'd ever done. By the first day, we had penetrated the network. Within a week, we were controlling a nuclear power plant. I thought, 'Gosh. This is a big problem.'”
- SCADA Penetration test by Lunsford in 2007

“An attacker with $500 of equipment ... could take command and control of the [advanced meter infrastructure] allowing for the en masse manipulation of service to homes and businesses.”
- IOActive, March 21, 2009

“Hacking The Smart Grid”
- DefCon 2009, by Tony Flick

Security is Needed for Business

"it is paramount that smart grid devices and interoperability standards include protections against cyber intrusions .. that are designed from the start *(not patches added on)*."
- Patricia Hoffman, US DOE, in Testimony to House, July 23 2009

"*Merit Review Criteria:*
…Addressing Interoperability and *Cyber Security* (20%)"
Stimulus Package Allocated 4.5 Billion for the Smart Grid. One vehicle, the Smart Grid Investment Grant Program Funding Opportunity Number: DE-FOA-0000058

Security in the Smart Grid

- Advanced Metering Infrastructure
- Home Area Network
- SCADA
- Distributed Control Systems
- Physical Plants

AMI Security Specification 1.0.1

Security Services

- What is the current usage (Availability)
- Is the meter lying? (Integrity)
- No unauthorized party should read my meter? (Confidentiality)
- When was the last time I heard from a meter? (Accounting)
- Does this look right? Is the load unusual? (Anomaly & Attack Detection)

Security in the Smart Grid

Advanced Metering Infrastructure

Home Area Network

SCADA

Distributed Control Systems

Physical Plants

ZigBee & HAN Security

- Are you authorized to shut down the alarm? (Available & Authentication)
- Can others snoop on my HAN settings? (Secrecy)
- I am a new fire alarm. Lets talk. (Key Management, enrollment)
- How do I set the system up? (Usability)
- Only dad can turn up the air conditioner (Access Control)

Security in the Smart Grid

- **Advanced Metering Infrastructure**
- **Home Area Network**
- **SCADA**
- **Distributed Control Systems**
- **Physical Plants**

"Securing DCS/SCADA is a national priority." – The National Strategy to Secure Cyber-Space (2003)

- Are all those commands real or a DoS? (Availability)
- Will that break the system? (Survivable and Available)
- Are you a real technician? (Authentication)
- Are you authorized to disable that functionality? (Access Control)
- Are you a hacker? (Intrusion Detection)

Security in the Smart Grid

Advanced Metering Infrastructure

Home Area Network

SCADA

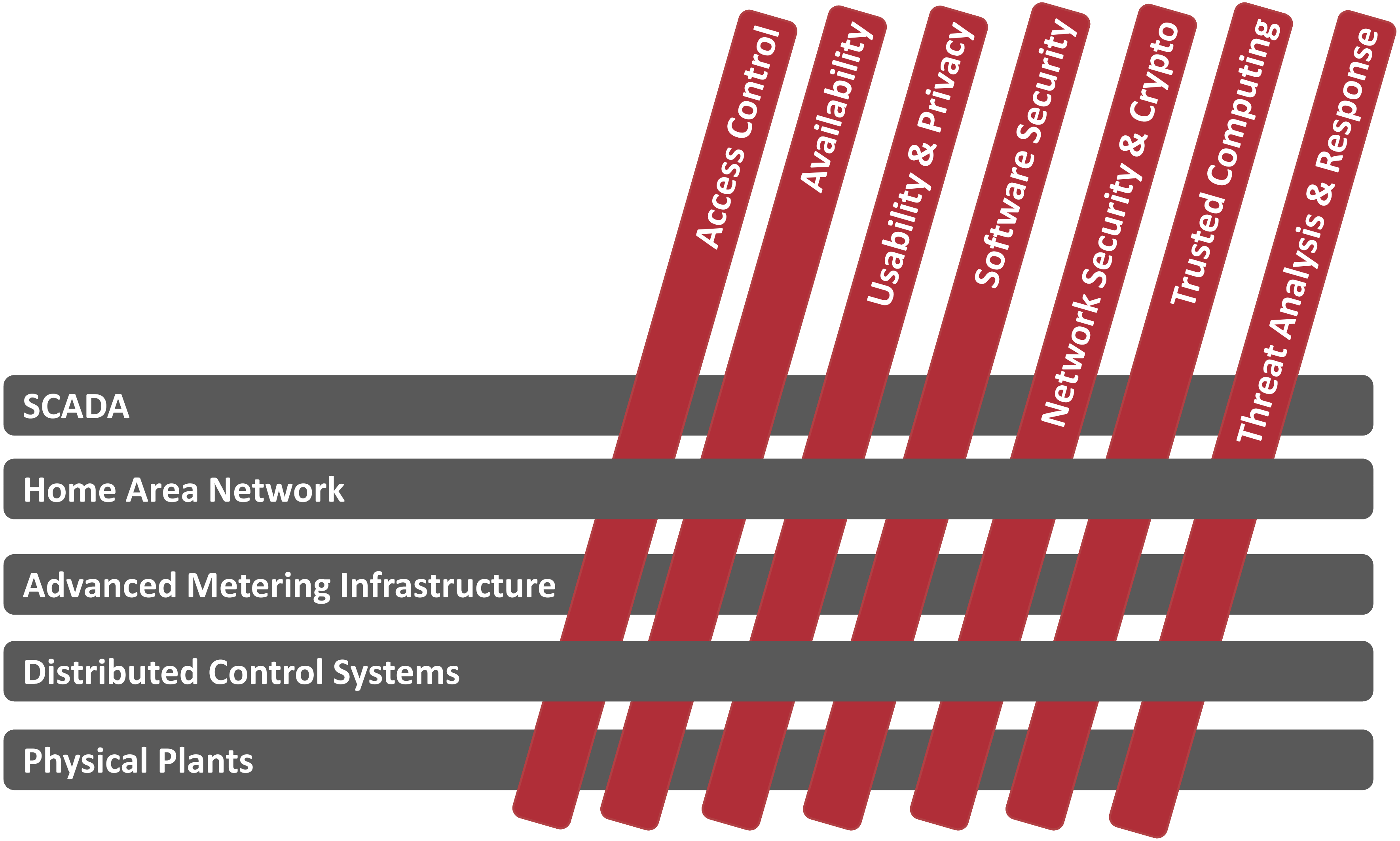
Distributed Control Systems

Physical Plants

NERC Cyber Security Standard Compliance for Bulk Power Systems

- How can I update the software without downtime? (Availability)
- Who are you? (Authentication)
- Should I have this modem connected? (Vulnerability Assessment)
- System 2.0 came out. Should I install it? (Up-to-date patches)
- Has our system been hacked? (Intrusion detection)
- What did the hacker do? (Accountability)

Smart Grid Security: @ Intersection of Fields



Sensor replay attack

- General attacker strategy:
  - Attack/compromise a system
  - Modify sensor data so that victim doesn’t know they are under attack
- Classic example replay attack:
  - Bank robbers replace video feed by taped segment showing no malicious activity, and then proceed to steal the money
  - Security guards do not notice the coffers are being emptied

System model

- We model the underlying physical system as a linear time-invariant system:

$$x_{k+1} = Ax_k + w_k$$

- Sensors are used to monitor the system:

$$y_k = Cx_k + v_k$$

- Each element in y_k represents the reading of a certain sensor at time k.

Failure Detector

- A failure detector is used to detect abnormality in the system, which triggers an alarm based on the following condition:

$$g_k > threshold$$

where

$$g_k = g(y_k, \hat{x}_k, \ldots, y_{k-\mathcal{T}}, \hat{x}_{k-\mathcal{T}}),$$

and the function g is continuous.

Failure Detector

- For example, g_k for a chi-square detector takes the following form:

$$g_k = z_k^T \mathcal{P}^{-1} z_k$$

where
$$z_k = y_k - CA\hat{x}_{k-1},$$

and $\mathcal{P}$ is the covariance of z_k.

Replay Attack Model

(Mo and Sinopoli, Allerton conf. '09)

- The attacker can
  - Record and modify the sensors' readings y_k
  - Inject malicious control input
- Replay Attack
  - Record sufficient number of y_k without adding control inputs.
  - Inject malicious control input to the system and replay the previous y_k. We denote the replayed measurements to be y'_k.
- When replay begins, there is no information from the systems to the controller. As a result, the controller cannot guarantee any close-loop control performance. The only chance is to detect the replay.

Practical "applications"

The New York Times Reprints

This copy is for your personal, noncommercial use only. You can order pr
distribution to your colleagues, clients or customers here or use the "Repr
article. Visit www.nytreprints.com for samples and additional information. C

January 15, 2011

Israeli Test on Worn
Iran Nuclear Delay

By WILLIAM J. BROAD, JOHN MARKOFF and DAVID

This article is by ***William J. Broad, John M***

The biggest single factor in putting time on the nu
sophisticated cyberweapon ever deployed.
The worm itself now appears to have included t
Iran's nuclear centrifuges spinning wildly out of
The computer program also secretly recorded w
like, then played those readings back to plant op
bank heist, so that it would appear that everythi
were actually tearing themselves apart.

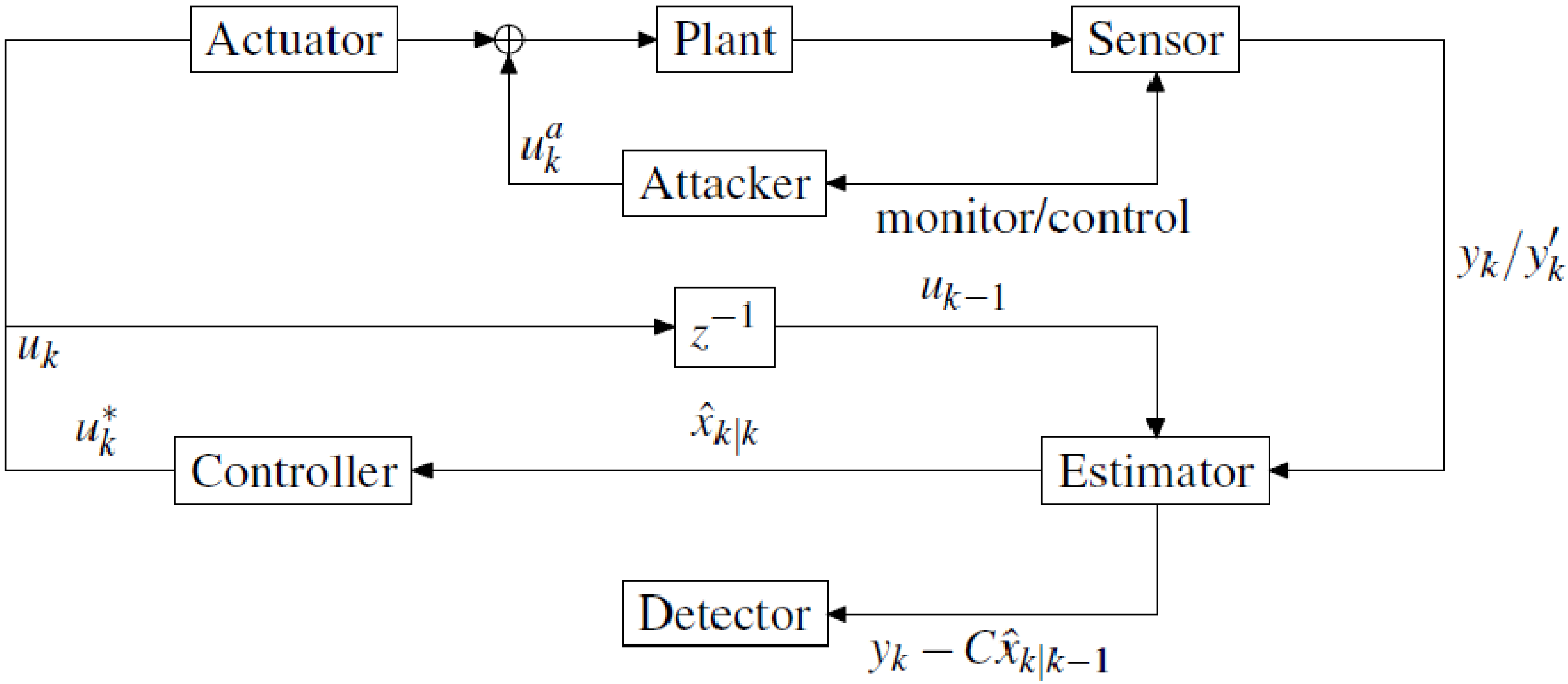
Forty-Seventh Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
September 30 - October 2, 2009

Secure Control Against Replay Attacks

Yilin Mo, Bruno Sinopoli *†

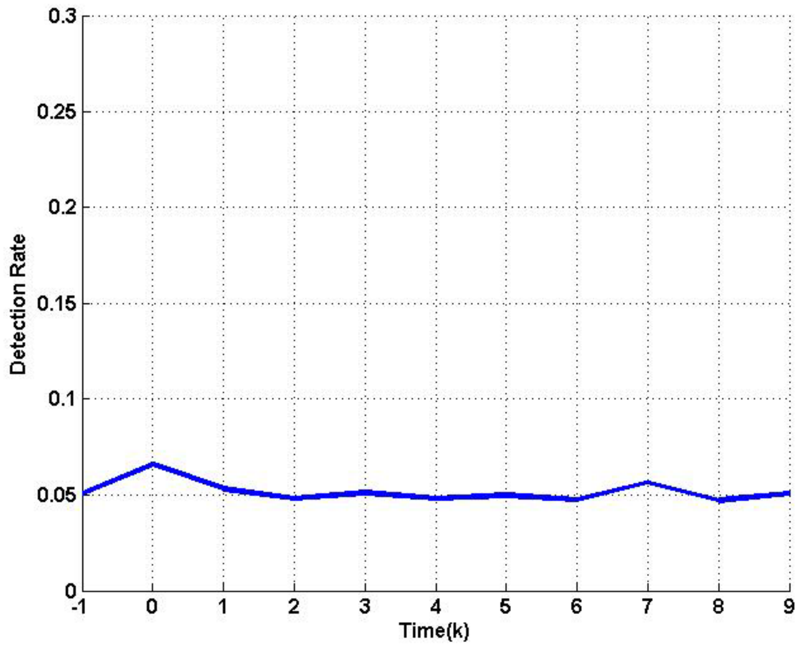
This paper analyzes the effect of replay attacks on a control system. We assume an attacker wishes to disrupt the operation of a control system in steady state. In order to inject an exogenous control input without being detected the attacker will hijack the sensors, observe and record their readings for a certain amount of time and repeat them afterwards while carrying out his attack. This is a very common and natural attack (we have seen numerous times intruders recording and replaying security videos while performing their attack ⋯

System Diagram



Simulation

- Suppose the attacker records from time –T and replay begins at time 0.



- For some systems, the Chi^2 detector cannot distinguish system under replay and system without replay.

Detection of Replay Attack

- Manipulating equations:

$$\boxed{y'_k - C\hat{x}_{k|k-1}} = \boxed{(y_{k-T} - C\hat{x}_{k-T|k-T-1})} + \boxed{C\mathcal{A}^k(\hat{x}_{0|-1} - \hat{x}_{-T|-T-1})},$$

$\uparrow$ innovation under replay; $\uparrow$ innovation without replay; $\uparrow$ converges to 0 if $\|\mathcal{A}\| < 1$

where

$$\mathcal{A} = (A + BL)(I + KC)$$

- If $\mathcal{A}^k$ converges to 0 very fast, then there is no way to distinguish the compromised system and healthy system.

Counter Measure

- Replay is feasible because the optimal estimator and controller are deterministic

- If we add random control input to the system:
  - If the system responds to this input, then there is no replay
  - If not, then there is a replay
  - Random control inputs act like time stamps
  - Cost: The controller is not optimal any more

Related challenges

- Integrity attacks
  - Changing reading of sensors w/o detection
- Price manipulation in Electricity Market Pricing
  - The price of electricity is determined by the state estimation , i.e. generation, power flow over transmission and load of the power grid.
  - If an attacker was able to compromise some sensors, then it could introduce a bias in the state estimation accordingly.
  - Eventually, over a finite time-horizon, the attacker will affect the pricing to his advantage and make a profit
- …

Conclusion

- Security has to be a first-order design principle in cyber-physical systems
  - Multiple points of attacks for smart grids
- Small attacks that run "under the radar" can have serious consequences
- A science of security for CPS systems needs to be developed
- Security needs to be integrated with system theory/knowledge