Attacks Against Process Control Systems: Risk Assessment, Detection, and Response

A.Cardenas, S. Amin, Z. Lin, Y. Huang,
C. Huang and S. Sastry
ASIACCS 2011

Presented by
Siddharth Murali

Control Systems

- Computer based systems that monitor and control physical processes
- Other names
  - Process Control Systems (PCS)
  - Supervisory Control and Data Acquisition (SCADA)
  - Distributed Control Systems (DCS)
  - Cyber-Physical Systems (CPS)

Attacks against Control Systems

- Computer-based accidents
- Non-targeted attack
- Targeted attacks – Stuxnet
  - Uses 0-day exploits, rootkits, stolen certs
  - Searches for WinCC/Step 7, and infects PLC
  - Uses a PLC rootkit to hide changes
  - Changed rotational speed of motors to 1410Hz to 2Hz and back to original speed
  - Shut down 984 centrifuges in Natanz

Current efforts and challenges

- Current Efforts
  - Focus on safety and reliability
  - Guidelines have been published
- Challenges
  - Patching and updates are not suited for control systems
  - Legacy systems
  - Real-time availability

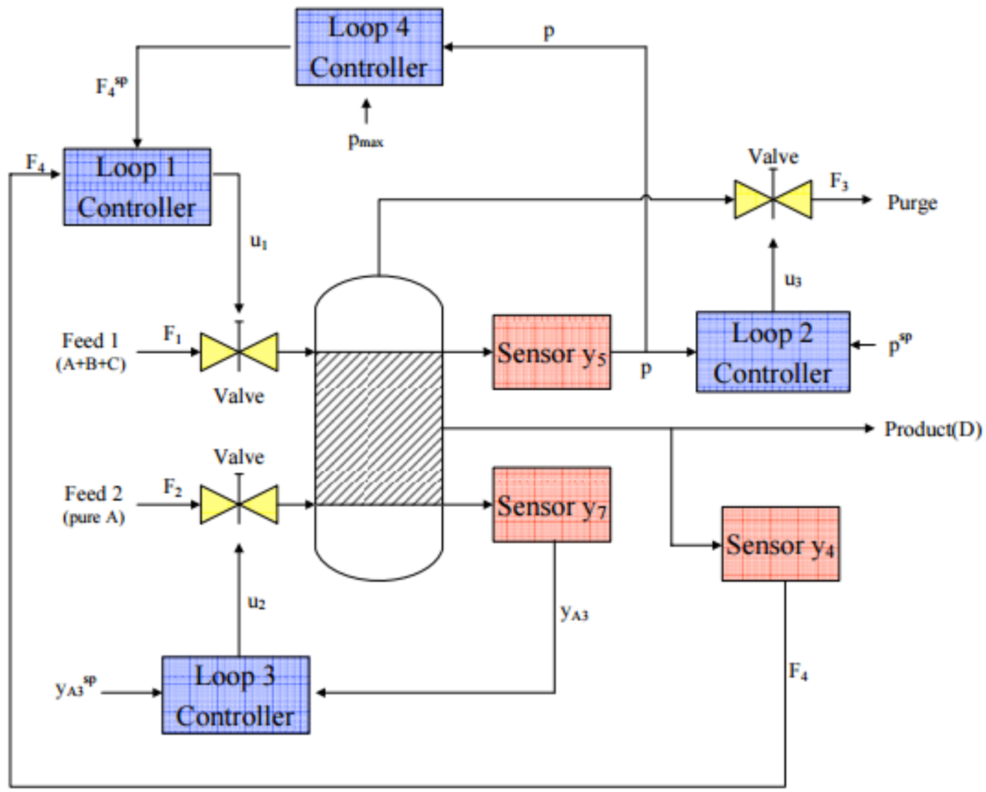
Contributions

- Risk Assessment
  - Understanding attack strategy of adversary
- New attack-detection algorithms
  - Detecting attacks based on compromised measurement
- New attack-resilient architecture
  - Design control systems to survive an attack with no loss of critical functions

Risk Assessment

- Attack model
  - Integrity attack
  - DoS attack
- Experiment
  - Goal is to make the reactor operate over 3000kPa
  - Attacker has access to a single sensor at a time

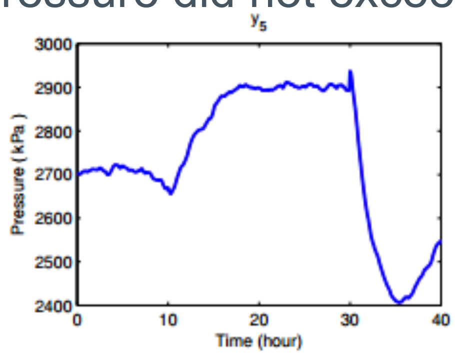
Experiment

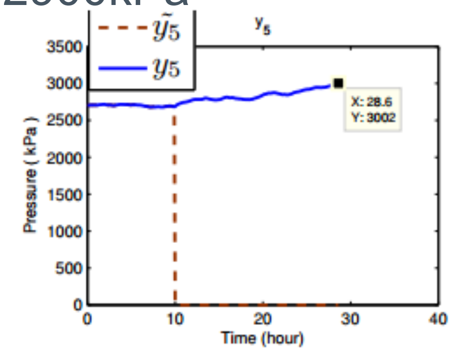


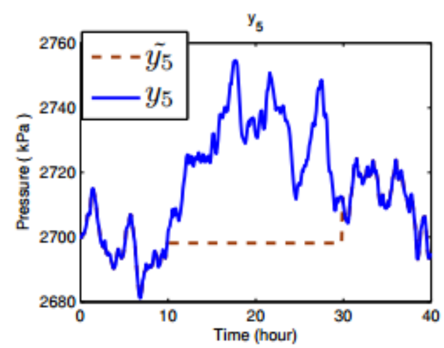
Loop 4 Controller
p
F_4^{sp}
p_{max}
F_4
Loop 1 Controller
Valve
F_3
Purge
u_1
u_3
Feed 1 (A+B+C)
F_1
Valve
Sensor y_5
p
Loop 2 Controller
p^{sp}
Product(D)
Valve
Feed 2 (pure A)
F_2
Sensor y_7
Sensor y_4
u_2
y_{A3}
F_4
y_{A3}^{sp}
Loop 3 Controller

Experiment Results

- Attacking the sensors (integrity attack) results in the controller responding with incorrect signals, but unable to force system into unsafe state
- Reducing the purge value did cause the pressure to increase past 3000kPa, takes 20 hours
- DoS attacks do not affect the plant, for a 20 hour DoS attack, pressure did not exceed 2900kPa







Detection of Attacks

- Optimal stopping problems
  - Given a time series sequence z(1), z(2), . . . , z(N) and hypotheses H0 (normal behavior) and H1 (attack)
  - Goal is to determine the minimum number of samples, N, the anomaly detection scheme should observe before making a decision
- Types of problems
  - Sequential detection
  - Change detection

Detection of Attacks

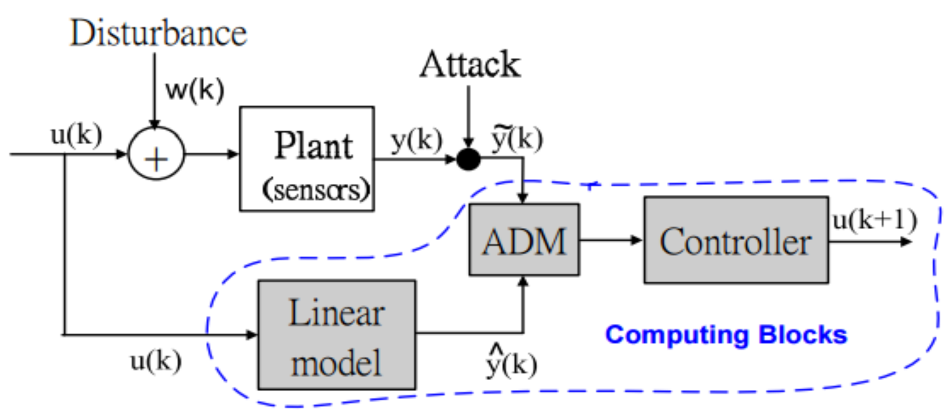
- Sequential Detection
  - Observation z(i) is generated either by H0 or H1
  - Goal is to decide which hypothesis is true in minimum time
  - Sequential Probability Ratio Test
- Change Detection
  - Observation z(i) starts under H0, but at a given time k, it changes to H1
  - Goal is to detect change as soon as possible
  - Cumulative sum(CUSUM)

Stealthy Attacks

- Goal is to raise pressure in the tank without being detected
- Surge Attacks
  - Attacker tries to maximize the damage as soon as possible
- Bias Attacks
  - Attacker adds a small constant to the system at each time step
- Geometric Attacks
  - The attacker wants to drift the value very slowly at the beginning and maximize the damage at the end

Response to Attacks

› Anomaly Detection Module
  – Replaces sensor measurements with measurements generated by the linear model if anomaly detection algorithm sounds alarm



Response to Attacks – Experiments

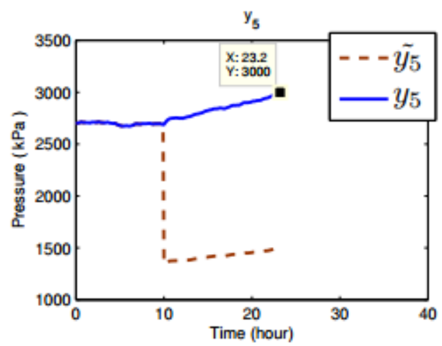
› Experiment ran for 40 hours

Alarms	Avg y_5	Std Dev	Max y_5
0	2700.4	14.73	2757

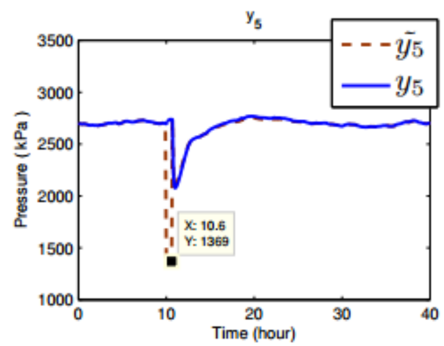
Table 1: For Thresholds $\tau_{y_4} = 50, \tau_{y_5} = 10000, \tau_{y_7} = 200$

	Alarms	Avg y_5	Std Dev	Max y_5
y_4	61	2710	30.36	2779
y_5	106	2705	18.72	2794
y_7	53	2706	20.89	2776

Table 2: Behavior of the plant after response to a false alarm



9(a) Without ADM



9(b) ADM detects and responds to the attack at $T = 10.7$ (hr)

Discussion

› Can these algorithms be applied to other CPS?

› How do you design a security protocol for control systems, keeping in mind the constraints?

› Will a system like this work against an attack like the Stuxnet worm?

› Is it enough to ensure integrity of a control system, or should we aim to prevent attackers from gaining access to the system as well?