KTH ROYAL INSTITUTE OF TECHNOLOGY

Resilient Smart Grid Control: Two Case Studies

Henrik Sandberg

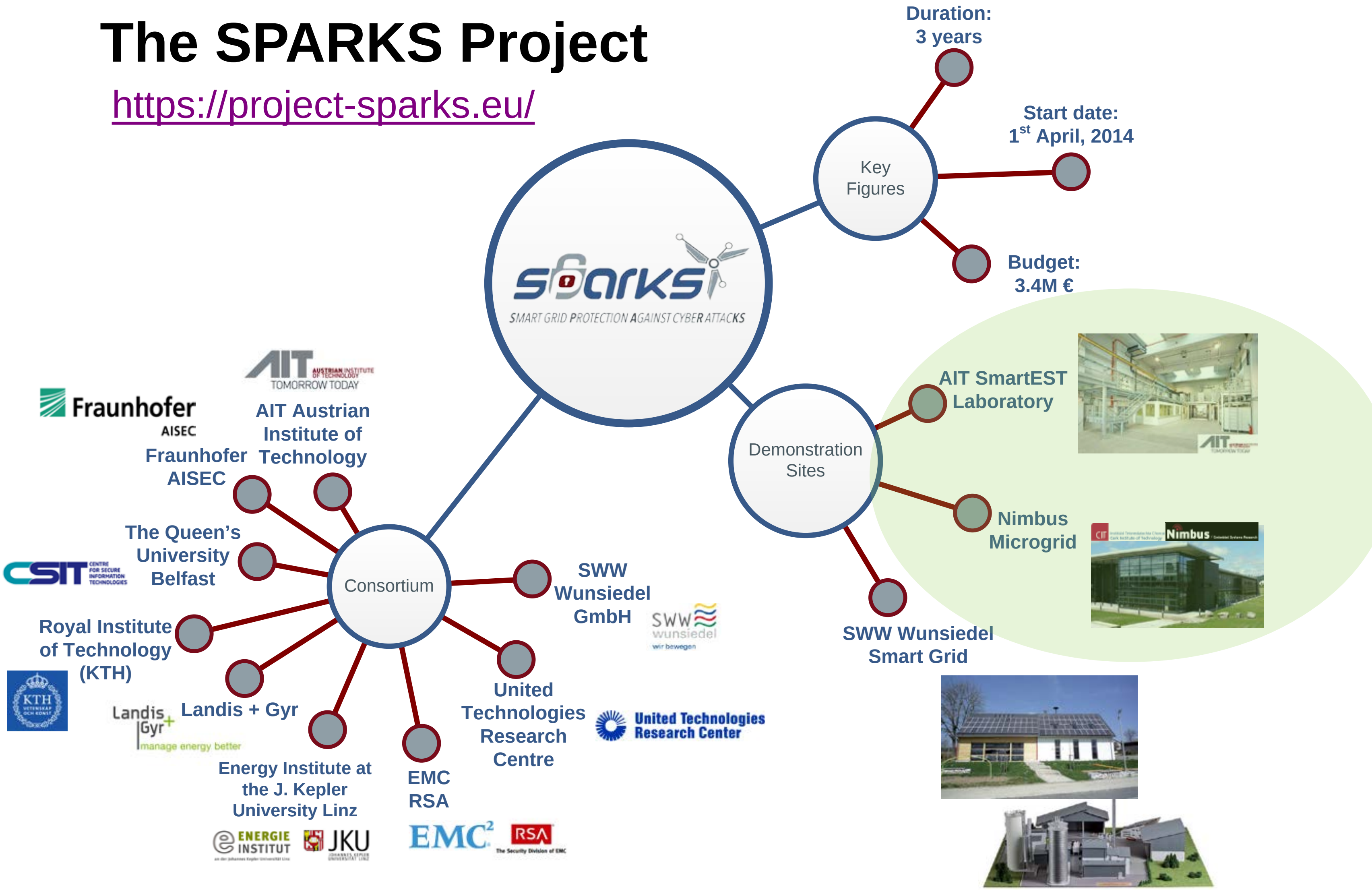
hsan@kth.se

Department of Automatic Control, School of Electrical Engineering

KTH, Stockholm, Sweden

The SPARKS Project

https://project-sparks.eu/



Joint Work With…

KTH: Kaveh Paridari, David Umsonst, Karl H. Johansson

AIT: Paul Smith, Friederich Kupzog, Mario Faschang

CSIT: BooJoong Kang, Kieran McLaughlin

Delft University of Technology: André Teixeira

Dell-EMC: Niamh O'Mahony

UTRC: Alie El-Din Mady, Rohan Chabukswar, Menouer Boubekeur

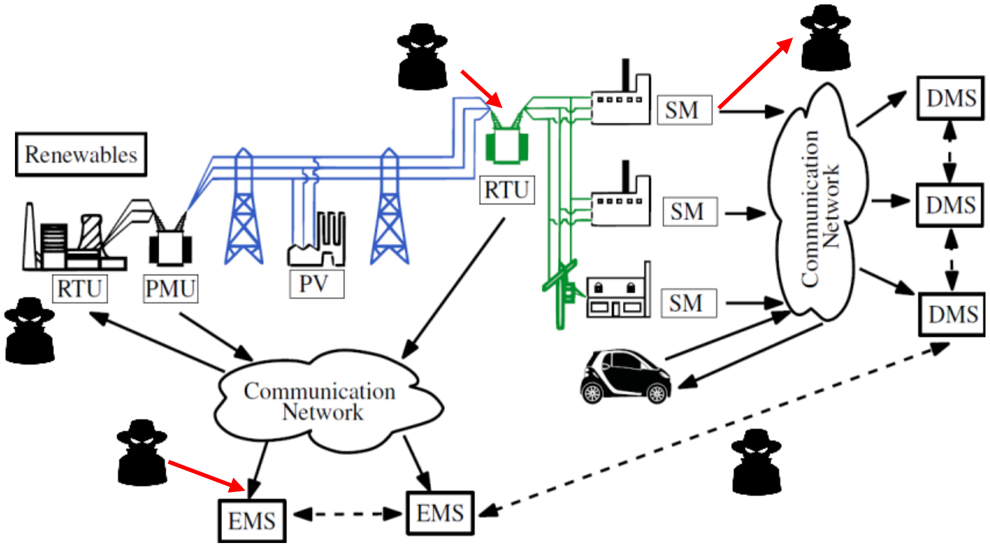
Motivation

The Smart Grid is a Cyber-Physical System

- **Power system** and **IT infrastructure** tightly coupled through SCADA and control systems. Lots of legacy equipment, but…
- Many ICT-enabled smart grid devices (photovoltaics, thermostats, battery inverters, electric vehicles, smart secondary substations, etc.)
- IT security necessary but not sufficient to secure cyber-physical systems

Today's talk

- Fault-tolerant control systems + IT-security → CPS resilience
- Integration with legacy systems
- Two attack/fault models



Outline

- Resilient control in cyber-physical systems

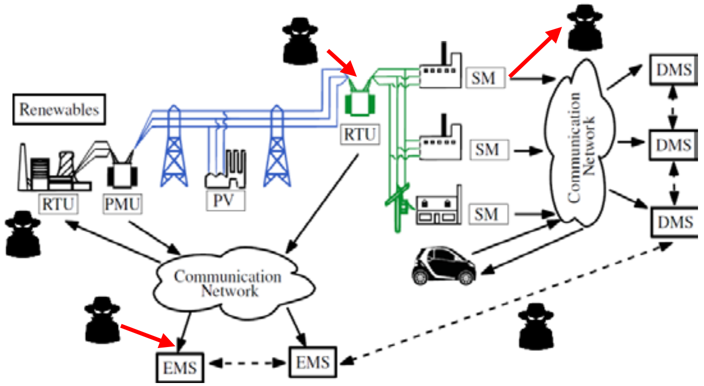
- Case Study 1: Low-level attacks against local controllers
  - Assumptions and architecture
  - Use Case: The NIMBUS Microgrid

- Case Study 2: Man-in-the-middle attacks against DERs
  - Assumptions and architecture
  - Use Case: Decentralized resilience in low-voltage grid

- Conclusions and outlook

DER = Distributed Energy Resource

Resilient Control System

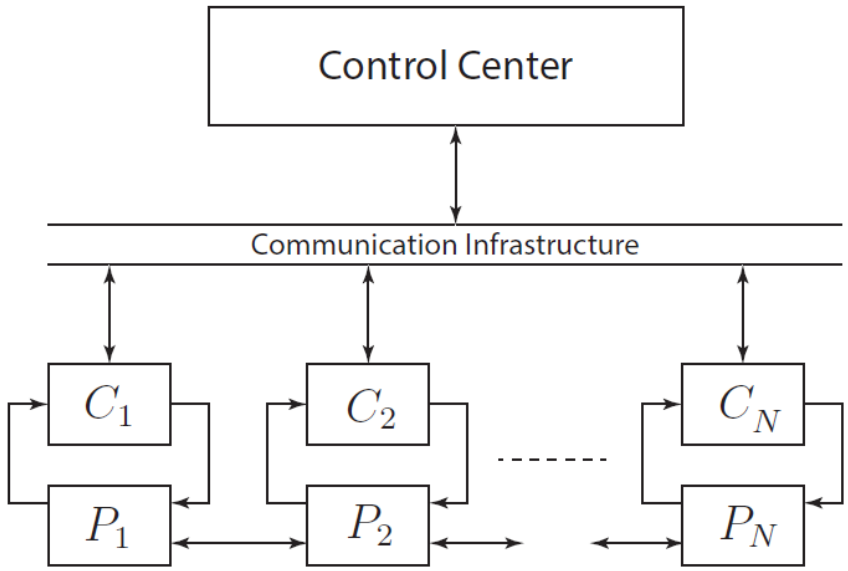


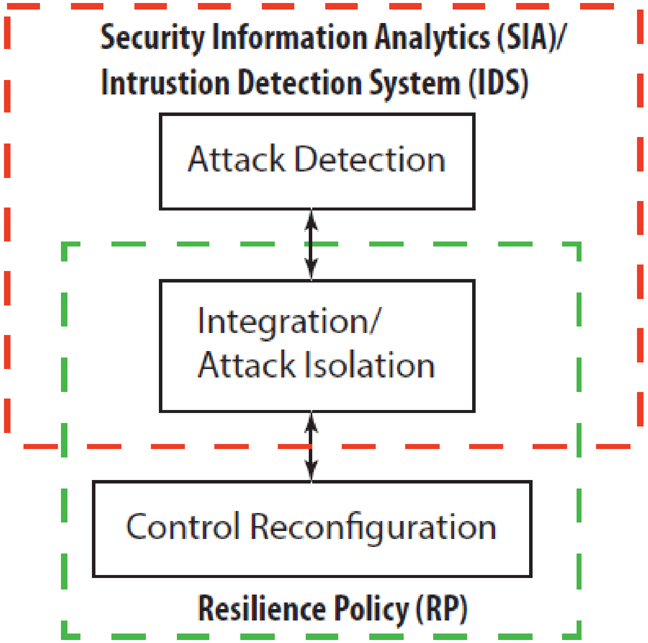
"A resilient control system is one that maintains state awareness and an accepted level of operational normalcy in response to disturbances, including threats of an unexpected and malicious nature."

- Rieger, Gertman, McQueen, 2009

- Faults and attacks will happen
- We cannot foresee them all, so aim for resilience
- Physical knowledge (often) encoded in controllers. Use it!
- Which controllers should be given more/less authority?

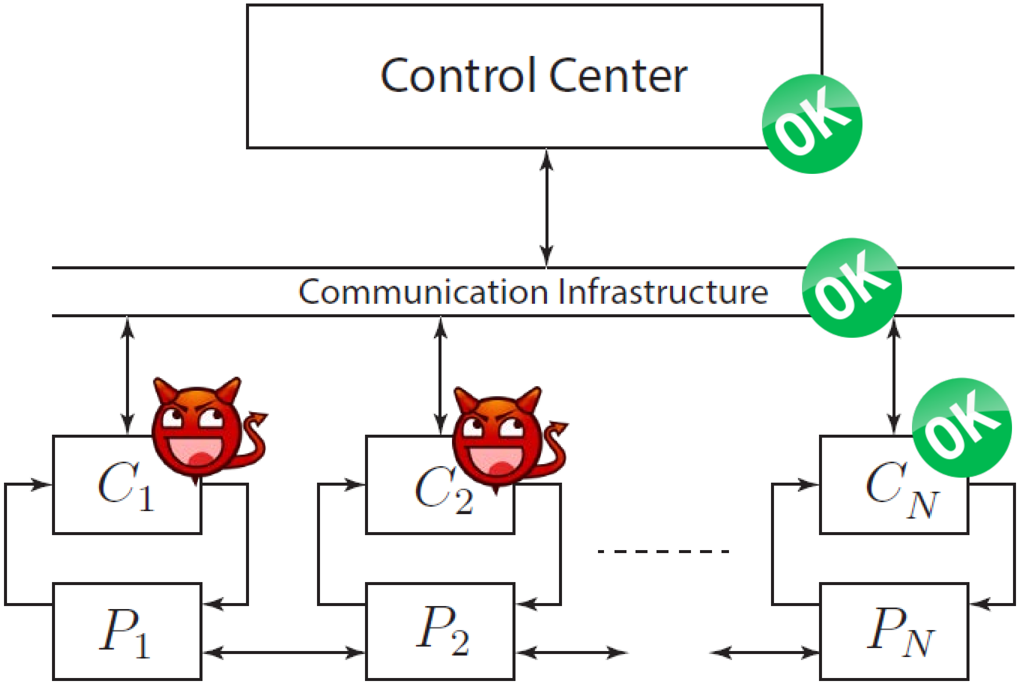
Proposed Security Architecture





- Common high-level defense architecture
- Different concrete distributed implementations to identified high-risk scenarios (NESCOR Failure Scenarios)

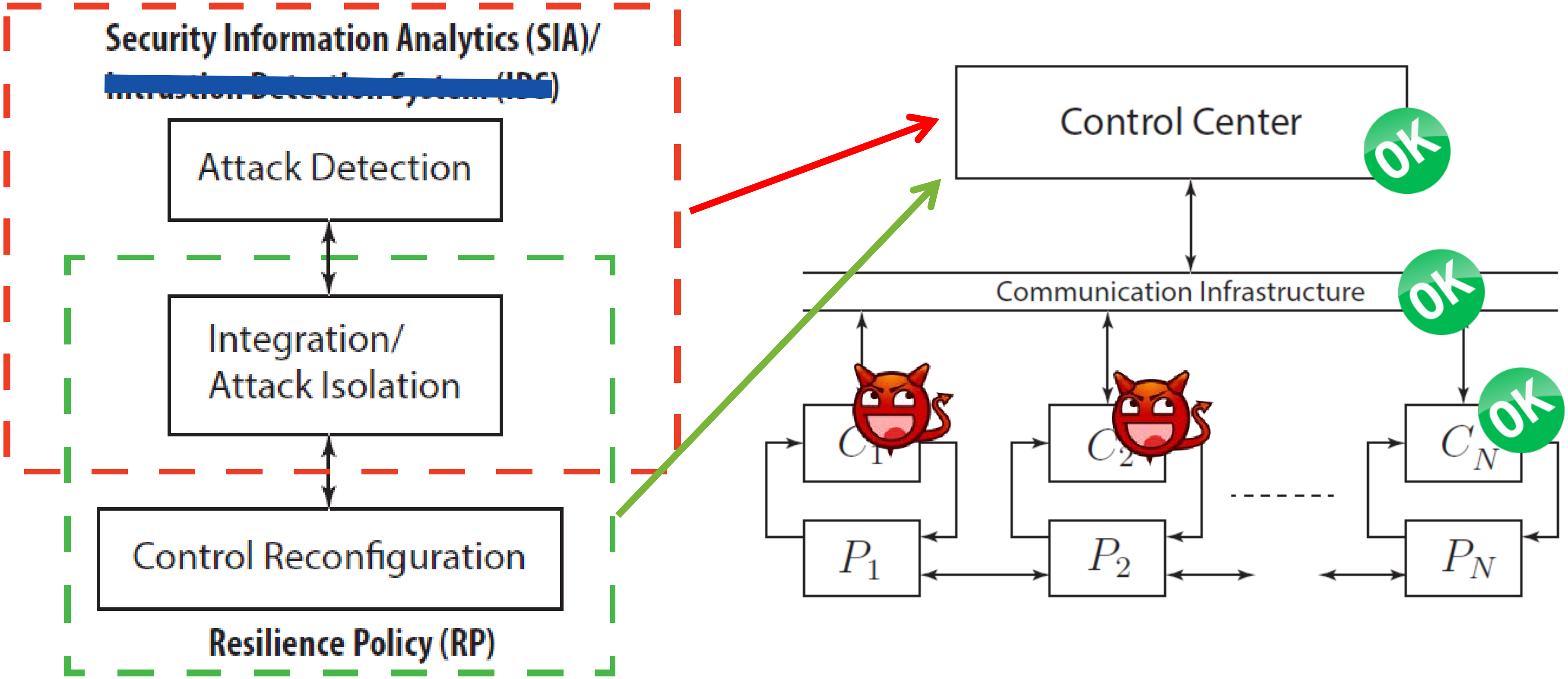
Case Study 1: Low-level Attacks Against Local Controllers



- Some, but not all, of the local controllers $(C_1, C_2, \ldots)$ are arbitrarily corrupted
- Communication Infrastructure, Control Center, and one Local Controller (C_N), are trusted
- Technical assumption: Infrastructure $(P_1, P_2, \ldots, P_N)$ *observable* from C_N

[*A Framework for Attack-resilient Industrial Control Systems,*" Proc. IEEE, 2017]
In collaboration with UTRC and Dell-EMC Corporation (Ireland)

Proposed Defense Architecture



Use Case:
NIMBUS Microgrid, Cork, Ireland

Electrical components

10kW wind turbine

35kWh (85kW peak) Li-Ion battery

50kW electrical/82kW thermal combined heat and power unit (CHP) and

Feeder management relay to manage the point of coupling between the microgrid and the rest of the building, and a set of local loads.

Battery and wind turbine interfaced through power electronics converters

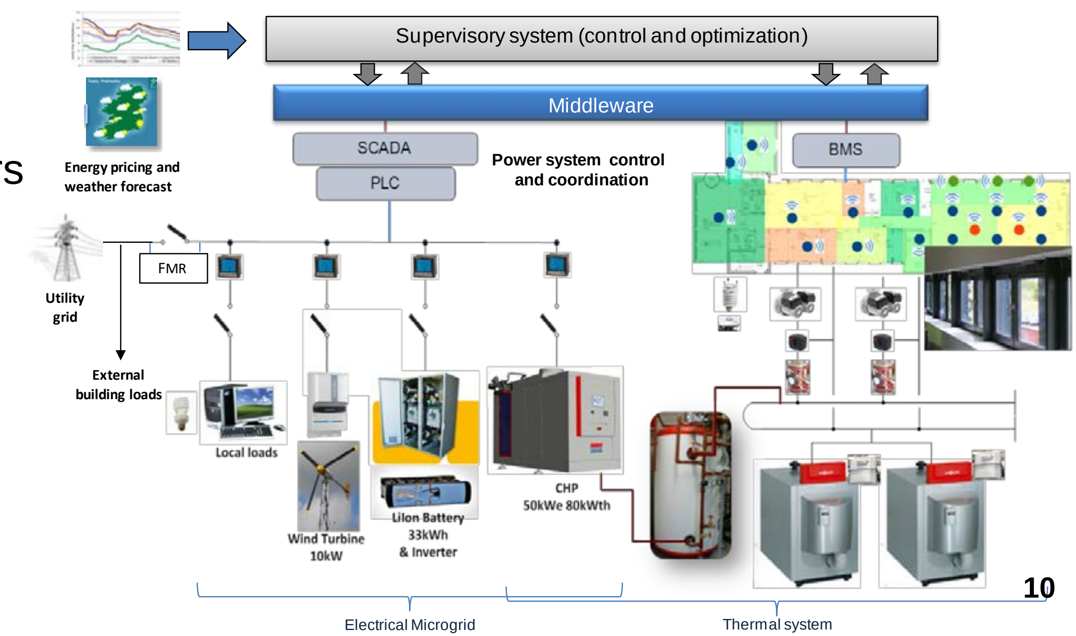
CHP with synchronous machine

IT System

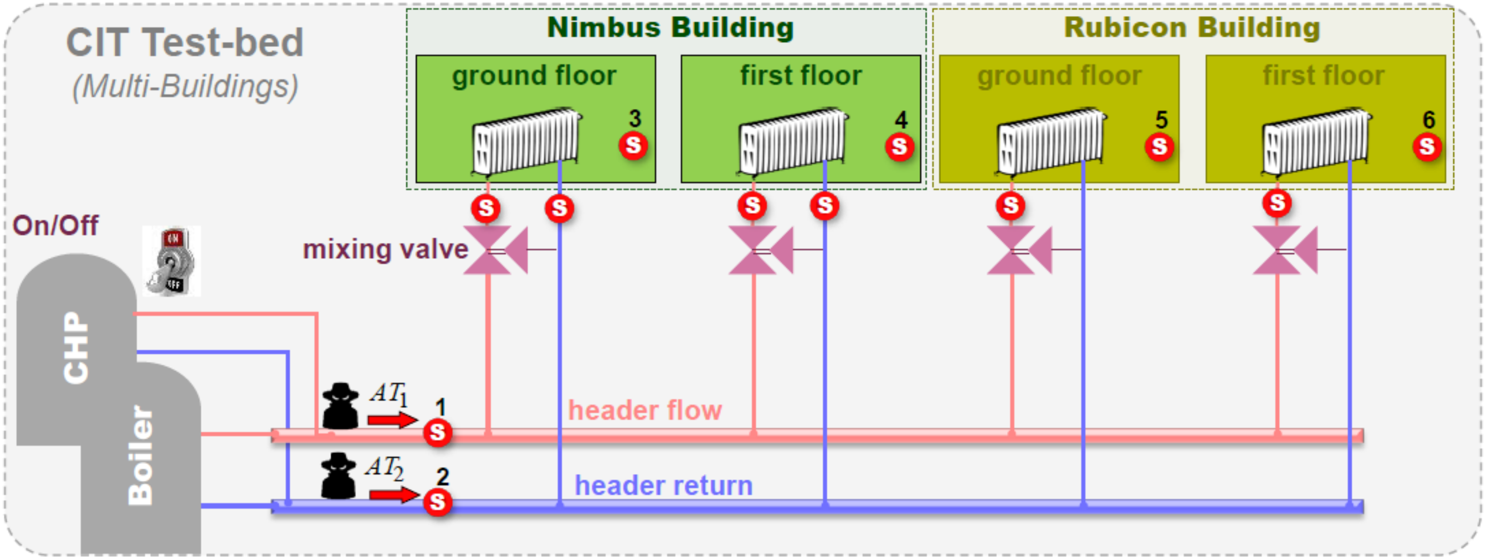
Interlinked Building Management System and Microgrid SCADA

Three-layer control systems

UTRC Middleware



Concrete Scenario: NIMBUS Microgrid

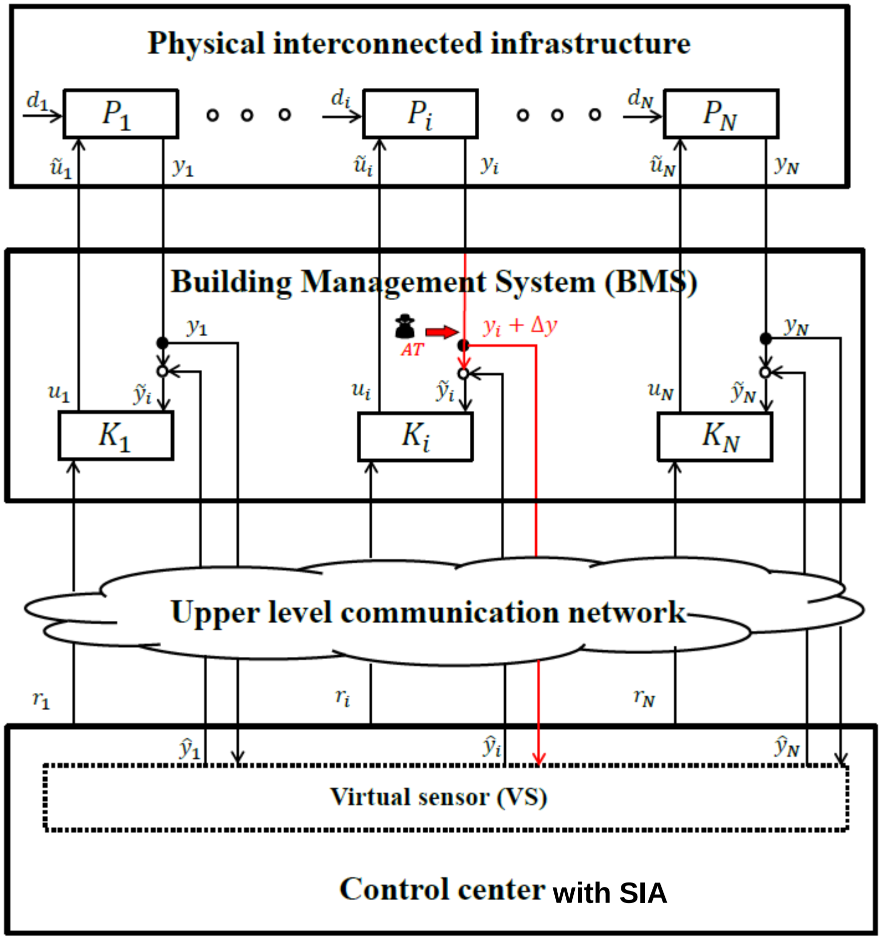


Adversary: Infect some field devices with malware (á la Stuxnet) corrupting measurements sent to PLCs (Here: AT_1 and AT_2)

Defender: Access to remote correlated measurements and a physical model (here temp. measurements and modeling by system identification)

PLC = Programmable Logic Controller (Local Controller)

Resilient Monitoring and Control



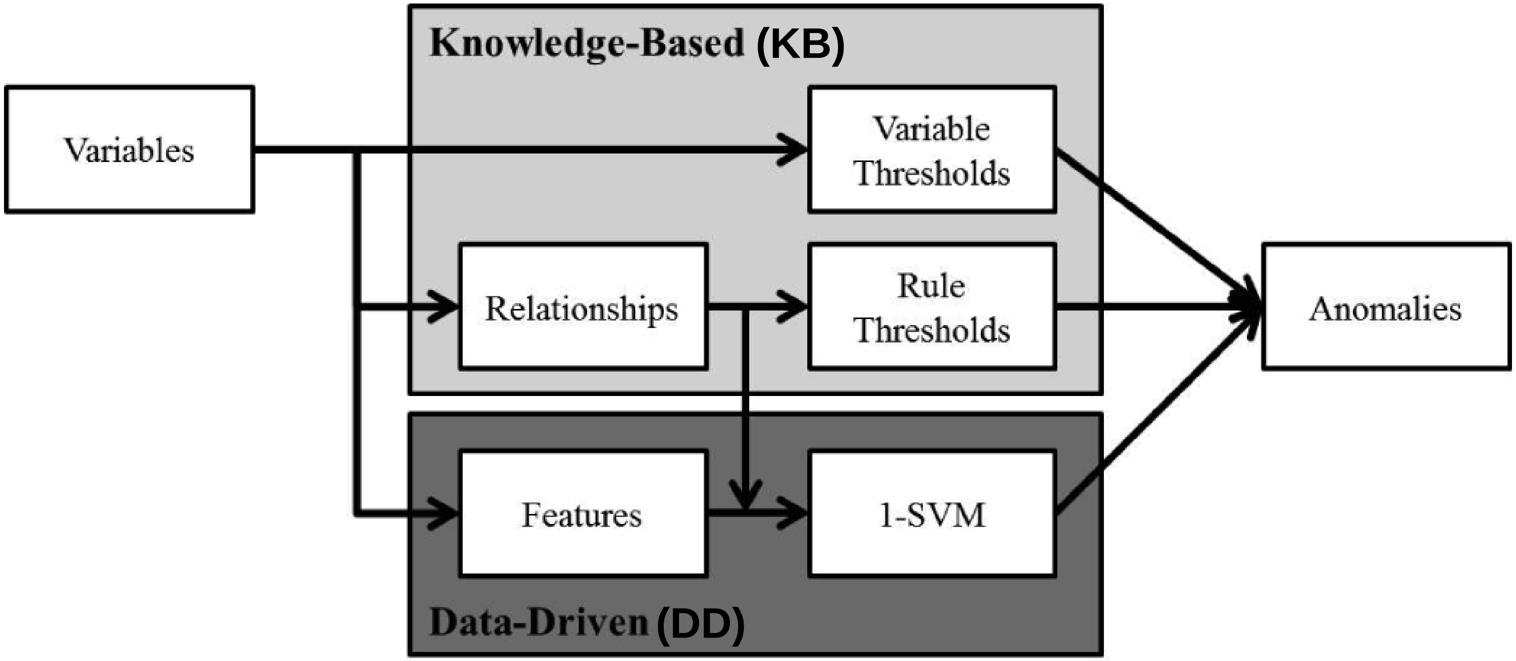
1) Anomaly detector (SIA) in control center detects attacked measurement $y_i + \Delta y$

2) Optimal physics-based prediction $\hat{y}_i$ from **un-attacked** measurements $y_1, \ldots, y_N$ (VS)

3) Feed $\hat{y}_i$ back to PLCs

SIA = Security Information Analytics

1) Anomaly Detector (SIA)

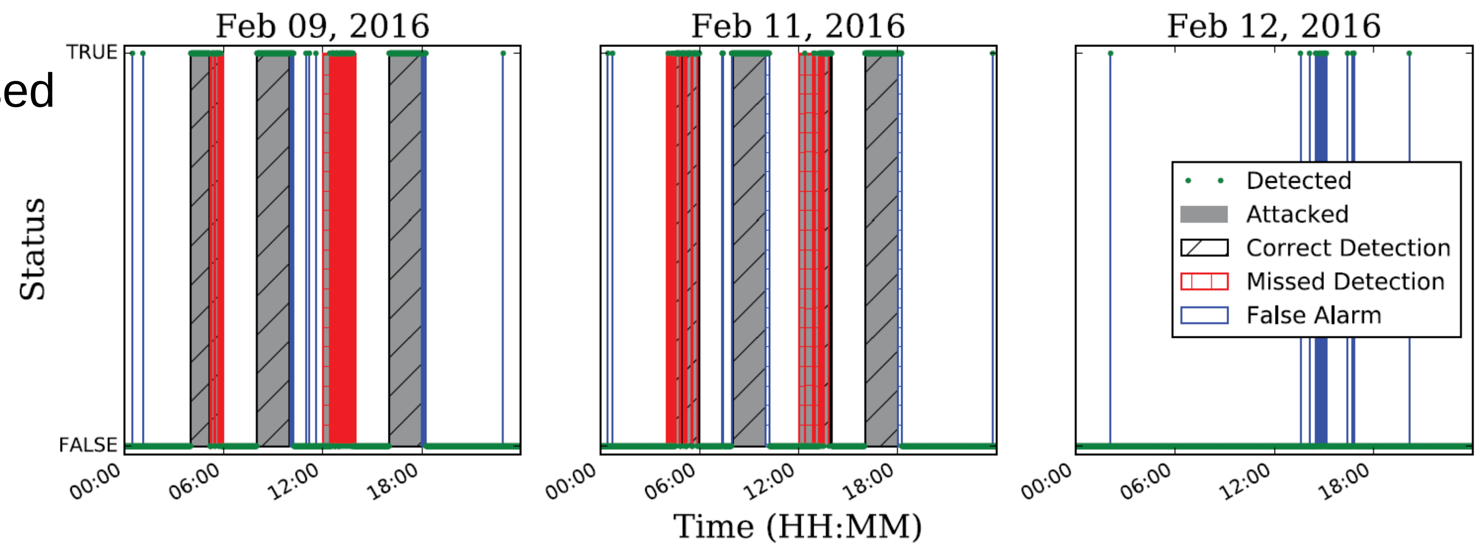


- KB Relationships: Physics-based model predictions
- DD Features: 1) Raw data, 2) KB residues, 3) Windowed mean and standard deviations
- Healthy data used to train 1-SVM

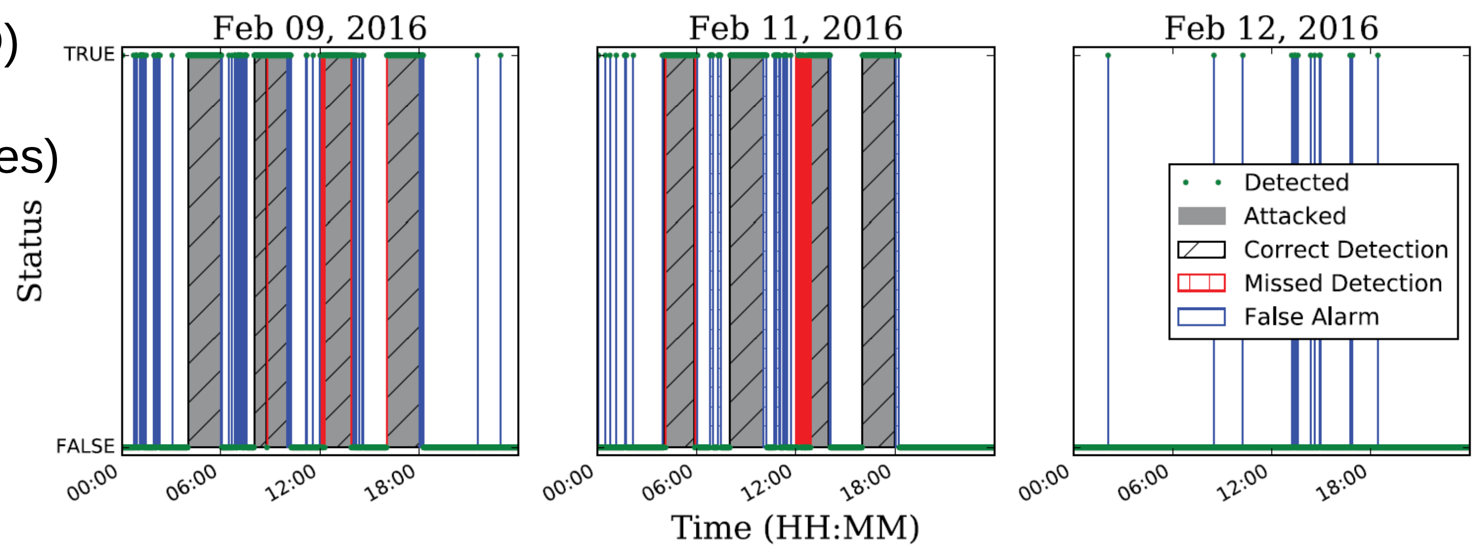
1-SVM = one-class Support Vector Machine

Test Results: Attack Detection

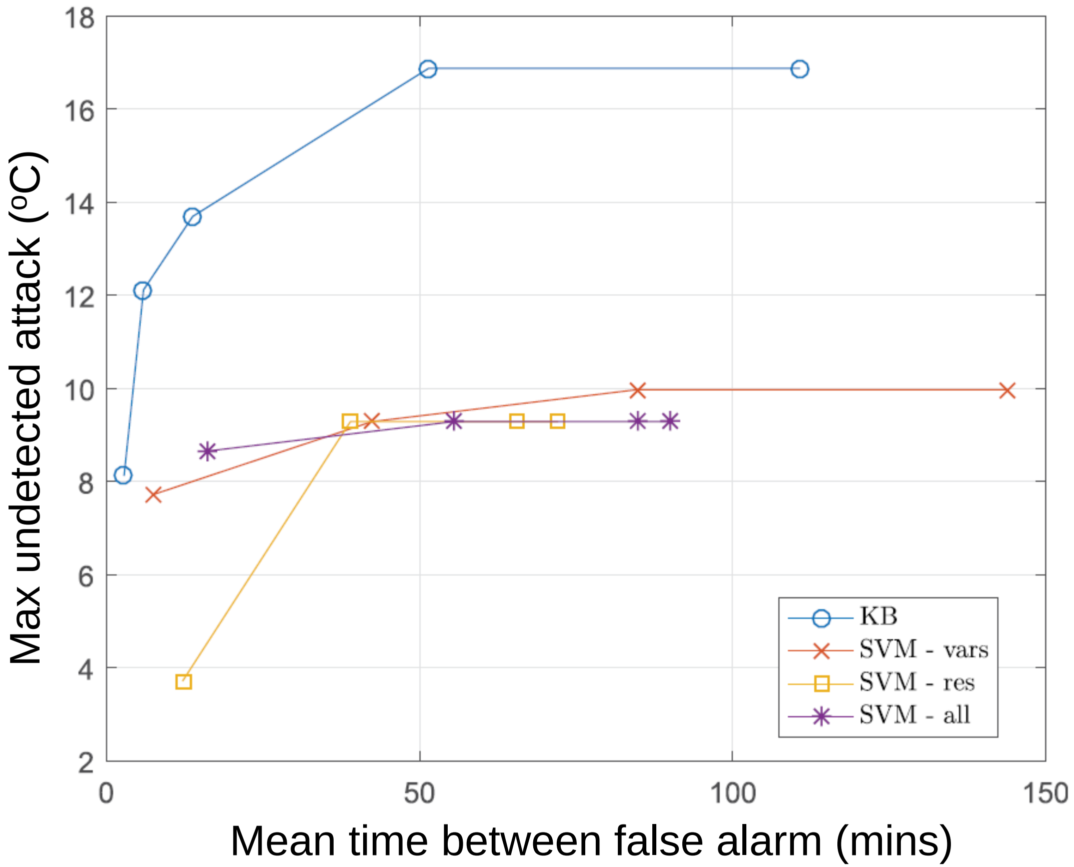
Knowledge-based (KB) detector:



Data-driven (DD) detector:
(raw data features)



Test Results: Attack Detection

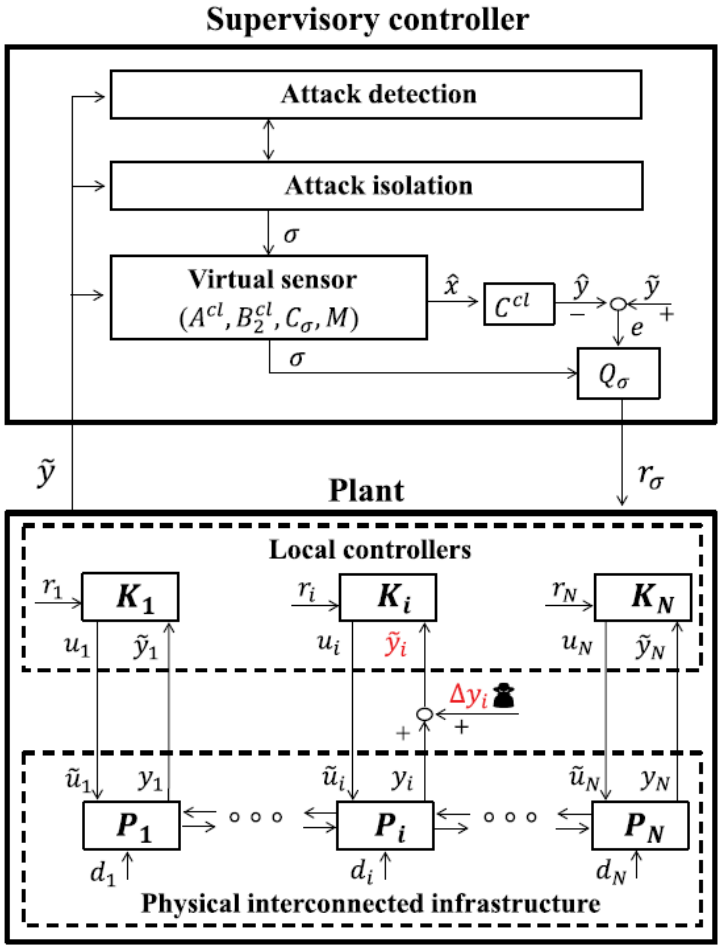


Attack start time	KB delay (mins)	DD delay (mins)
09-Feb-2016 04:00	0	0
09-Feb-2016 08:00	0	0
09-Feb-2016 12:00	24	2
09-Feb-2016 16:00	0	1
11-Feb-2016 04:00	6	0
11-Feb-2016 08:00	0	0
11-Feb-2016 12:00	22	7
11-Feb-2016 16:00	0	0

- DD detector restricts attacker more
- KB detector only checks “physicality” of time series
- DD detector also checks for unusual operation

Metric proposed in [Urbina *et al*., ACM CCS, 2016]

2-3) Reconfigured Control System



- Virtual sensor: KB switched Kalman filter

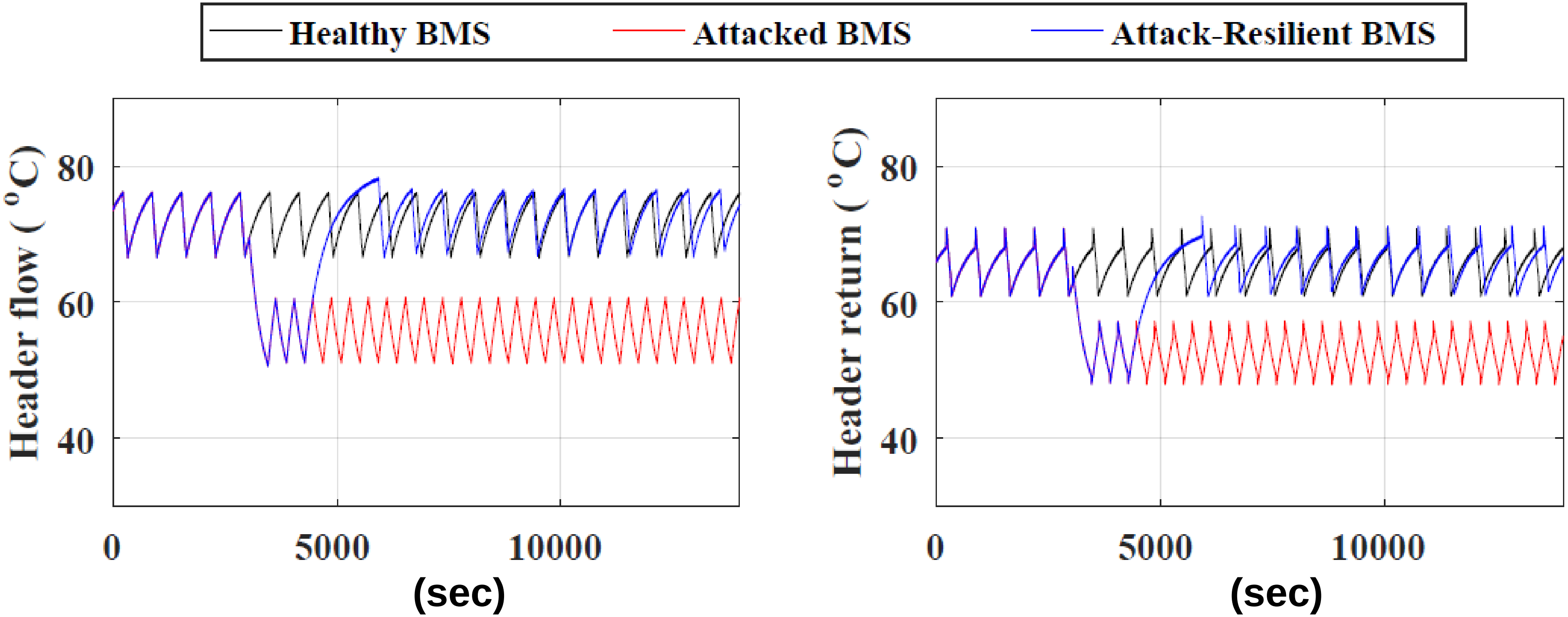
$$\hat{x}(k+1|k) = A^{cl}\hat{x}(k|k-1) + K_\sigma(k)\underbrace{\left[y_\sigma(k) - C_\sigma\hat{x}(k|k-1)\right]}_{\varepsilon(k)}$$

- Attack isolation chooses system mode $\sigma(k) \in \{1,2,\dots,M\}$
  - $\sigma = 1$: All sensors OK
  - $\sigma = 2$: Sensor 1 malfunction
  - $\vdots$
  - $\sigma = M$: Only trusted Sensor(s) OK

- Healthy sensors used to optimally correct unhealthy sensors, and signal the correction $r_\sigma(k)$ to affected Local Controllers

Test Results: Control Performance

24 min delay in anomaly detector (“attacker free time”):



BMS = Building Management System

Theoretical Analysis

Suppose closed-loop system is

- Linear
- Asymptotically stable when $\sigma = 1$ (all sensors healthy)
- Observable using only trusted sensor(s)
- Noise is i.i.d. Gaussian.

Theorem 1: For arbitrary switching sequences $\sigma(k)$, the switched Kalman filter yields an unbiased minimum error variance state estimate $\hat{x}(k)$.

Theorem 2: For arbitrary switching sequences $\sigma(k)$, the closed-loop system is asymptotically stable.

Case Study 1: Summary

DD and KB models, and trusted sensor used for

- Attack/fault detection and correction in untrusted low-level controllers
- Gracefully degraded real-time control performance under identified fault/attack conditions → Resilience
- Degraded performance due to increased time-delay and noise in feedback loops

Requirements

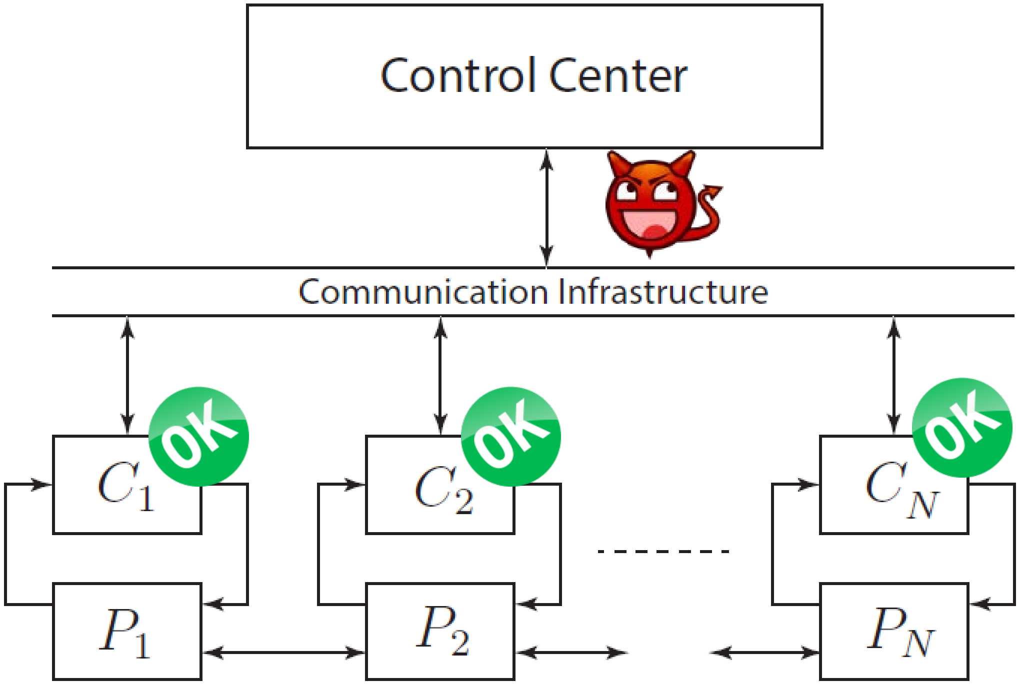
- Trusted control center and communication system
- Control center has authority to overwrite local actuation commands

How to allocate trusted sensor?

Session III: Jezdimir Milosevic *et al.*, "Security Measure Allocation for Industrial Control Systems"

[*A Framework for Attack-resilient Industrial Control Systems,*" Proc. IEEE, 2017]

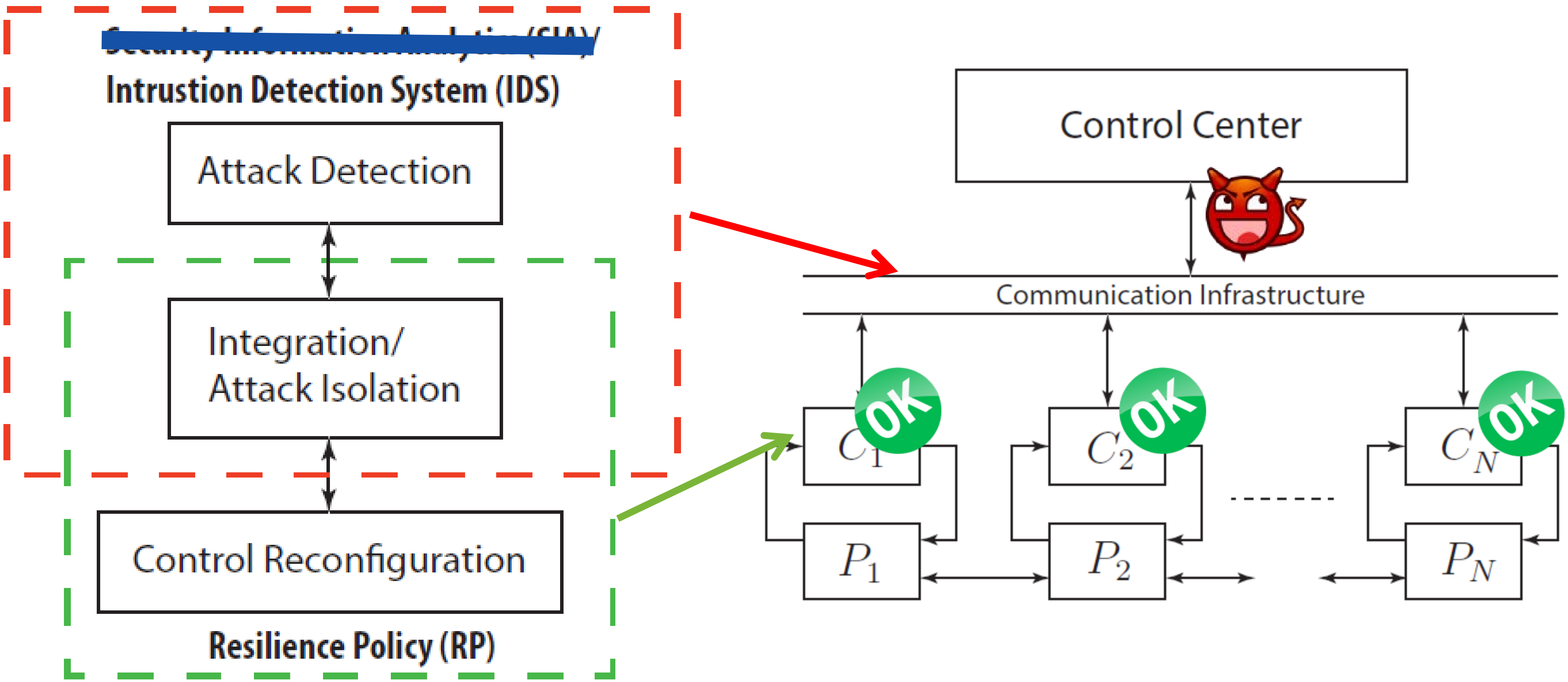
Case Study 2: Man-in-the-middle Attacks Against DERs



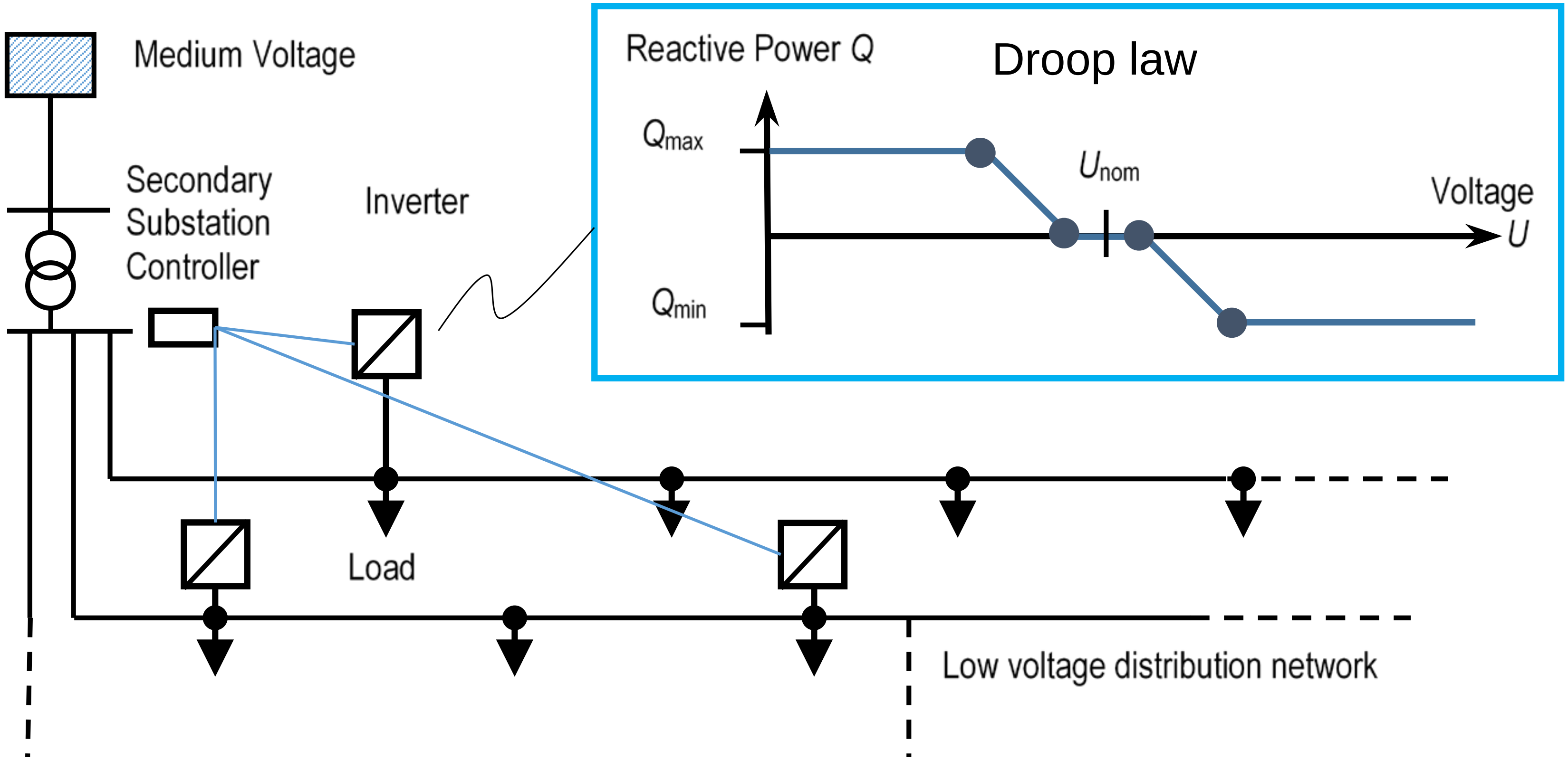
- Attacker corrupts some, or all, of the set-points from the Control Center to the local control loops (P_i, C_i)
- Local controllers C_i are trusted

[*SPARKS Cyber Security Demonstration Outcomes,*" SPARKS D6.4, D2017]
In collaboration with AIT and CSIT

Proposed Defense Architecture

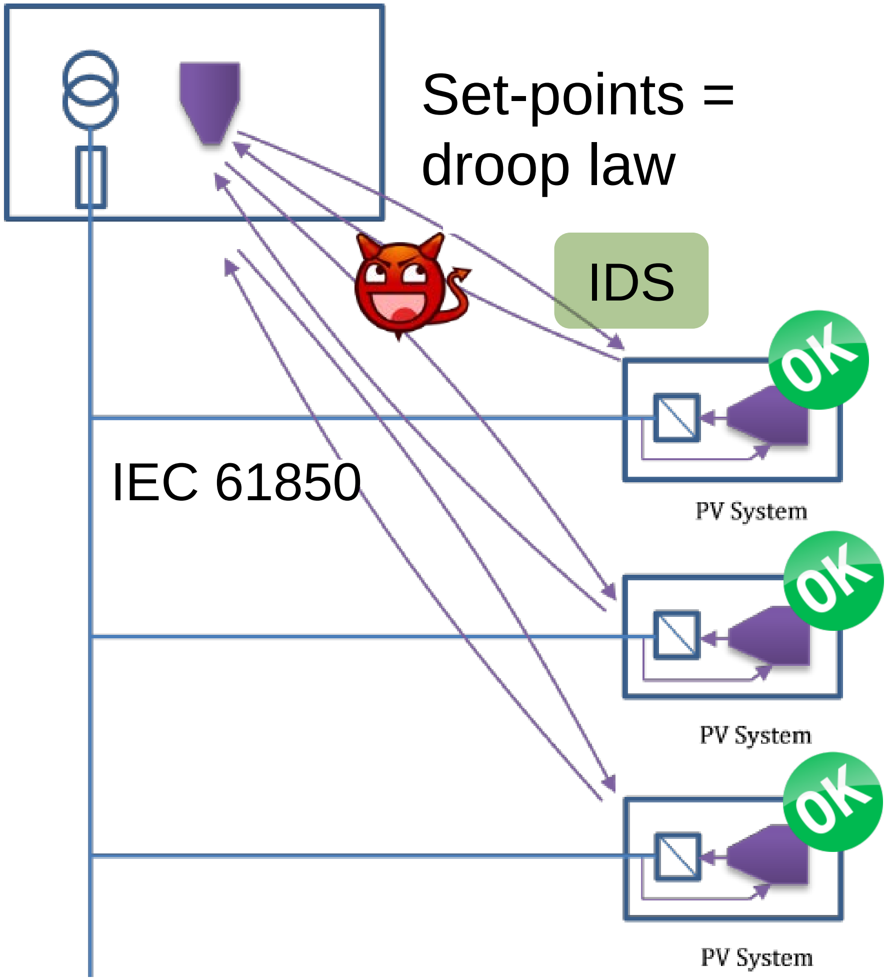


Use Case: Low-Voltage Grid Control with PV Inverters (AIT SmartEST Lab)



PVs = Photovoltaics

Concrete Scenario: Low-Voltage Grid Control with PV Inverters



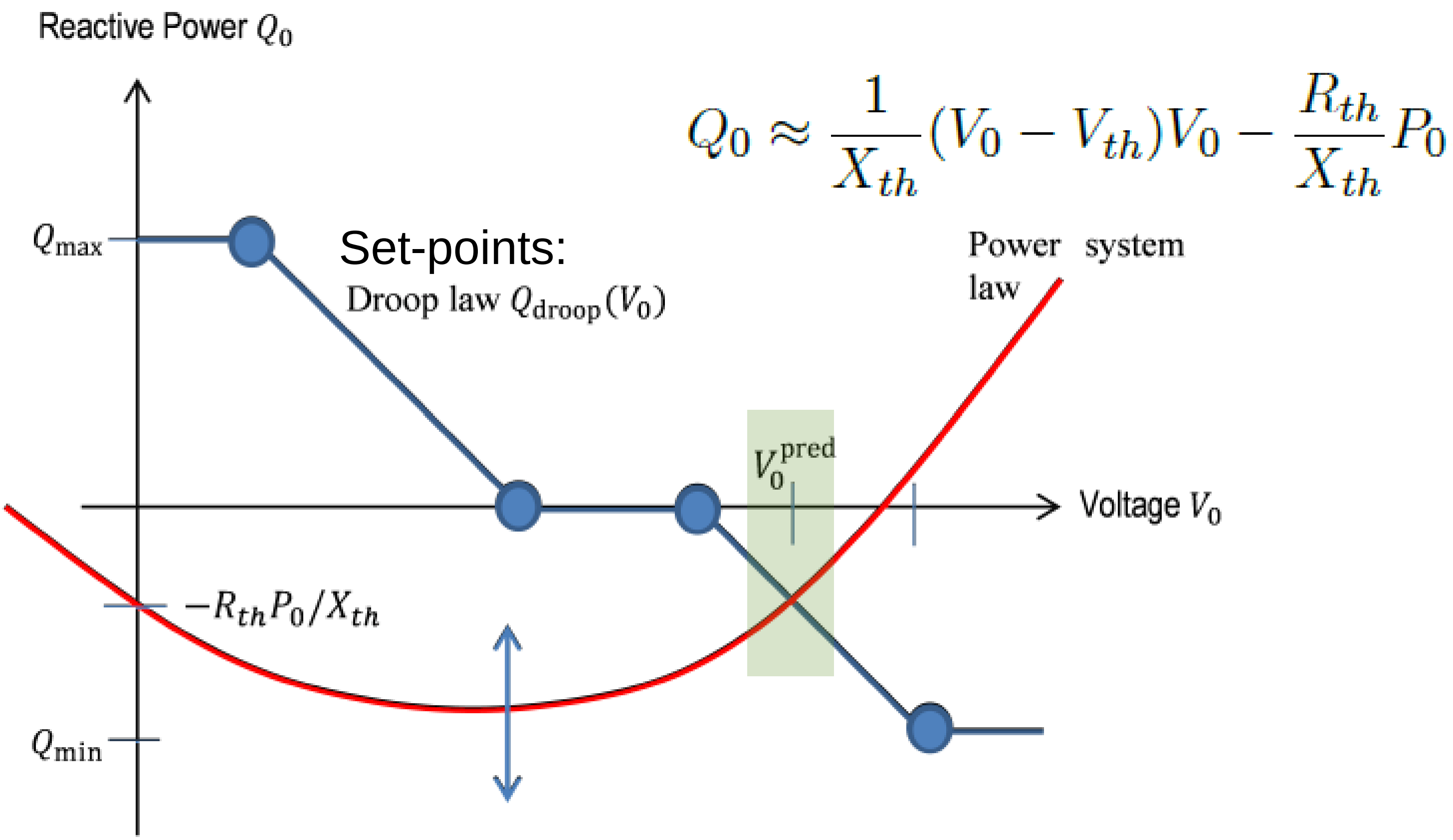
Resilience checks in PVs:

- Is new steady-state within safety limits?
- Is new droop law stabilizing?
- Communication with IDS:
  - Receive warnings
  - Report rule violations

IDS = Intrusion Detection System

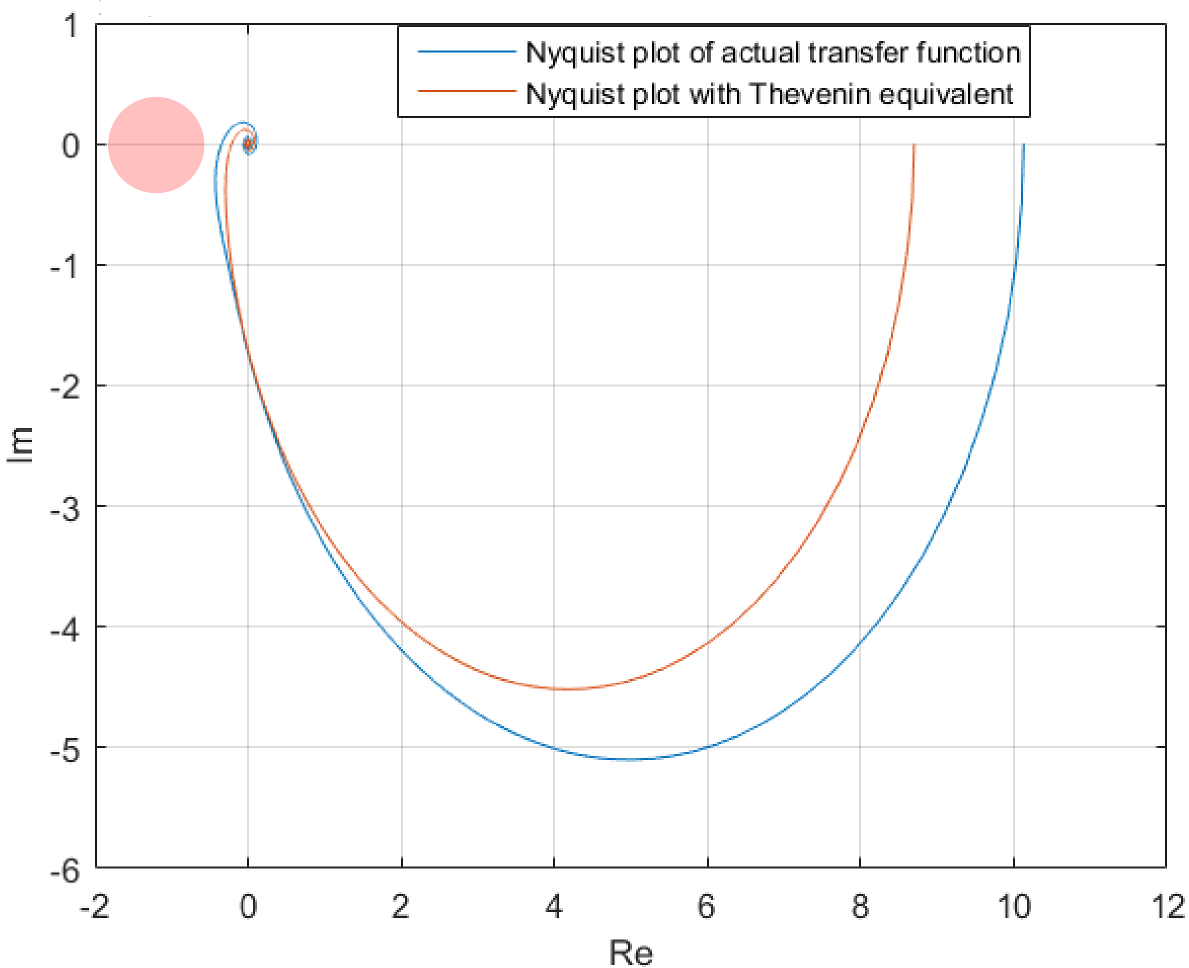
Decentralized Resilience Rule #1: New Predicted Steady-State Within Limits?

$$V_{\min}(t) \leq V_0^{\text{pred}}(t) \leq V_{\max}(t)$$



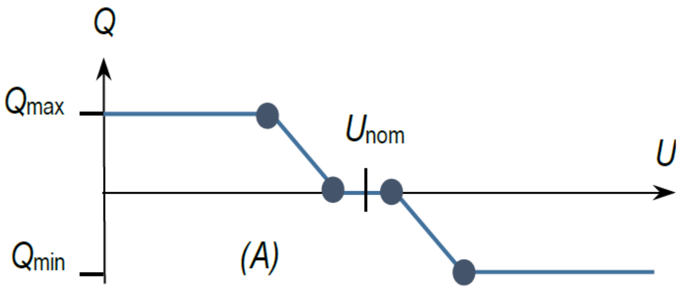
Decentralized Resilience Rule #2: New Feedback Gain Stabilizing?

$$K_{\text{droop}} < K_{\text{crit}} = \frac{V_{th} X_{th}}{R_{th}^2 + X_{th}^2} \quad \text{(circle criterion)}$$

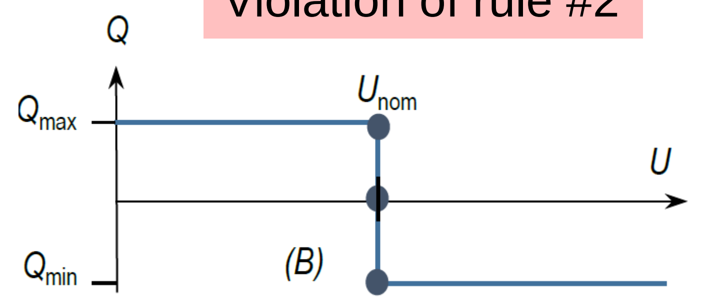


Experimental Verification

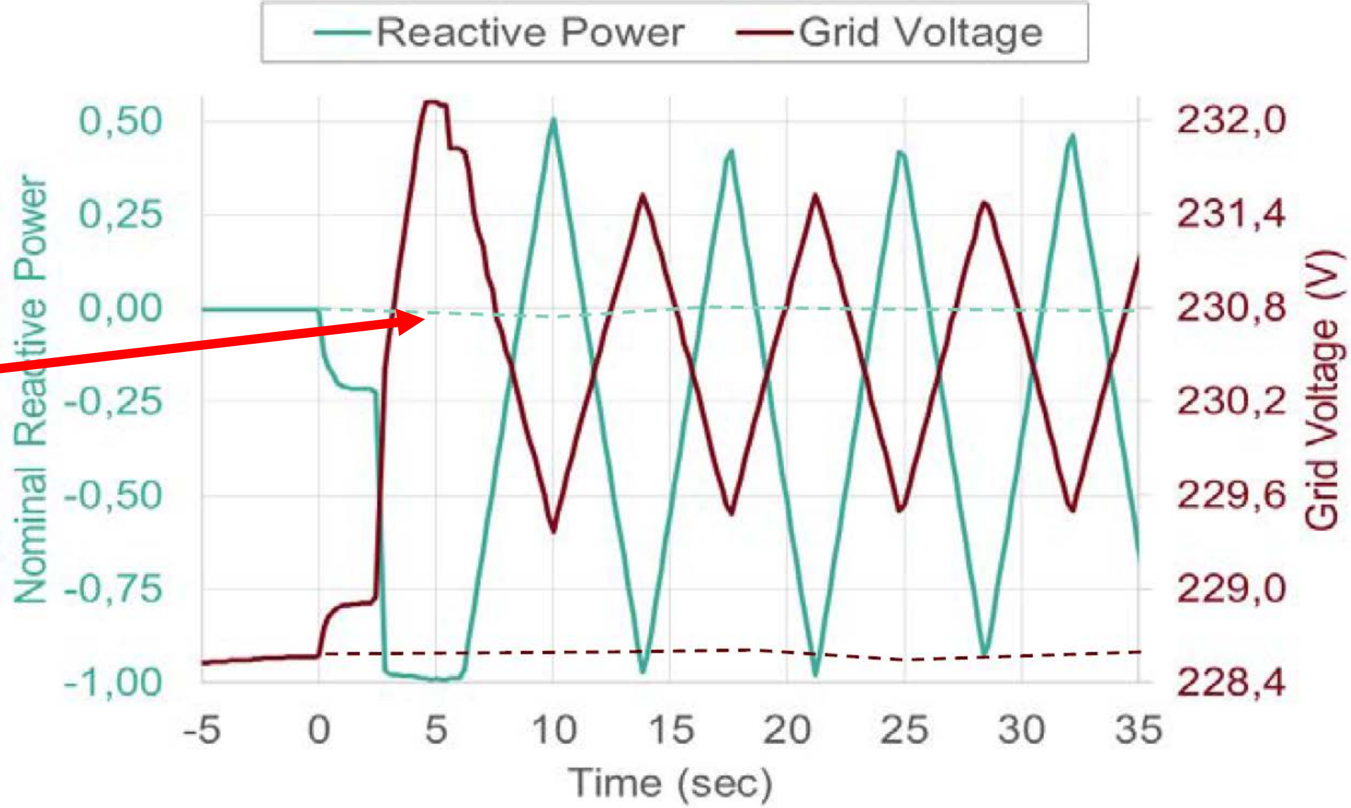
$K_{\text{droop}} > K_{\text{crit}}$
Violation of rule #2



Attack/Fault



Set-points (B) rejected by resilient controller



Case Study 2: Summary

Trusted local controller and network-based IDS used for

- Attack/fault detection in untrusted remote commands
- Possibly rejected/curtailed remote commands → Resilience
- Degraded performance due to reduced remote control authority

Requirement

- Local controller has authority to ignore/correct remote commands

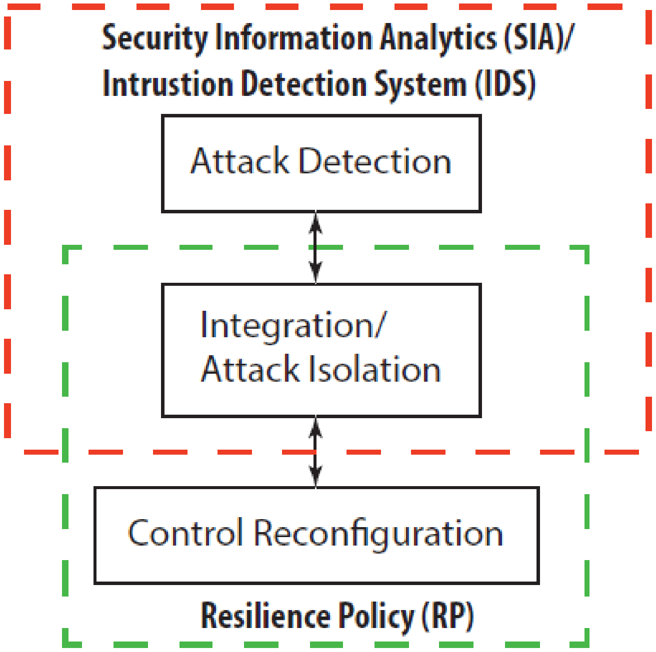
Challenges

- Interaction rules between local controller and networked-based IDS
- Adaptation of local resilience rules (not overly conservative)
- Trade-off performance, safety, and security

[*SPARKS Cyber Security Demonstration Outcomes,*" SPARKS D6.4, D2017]

Conclusions

- Two concrete attack scenarios considered
- Common high-level defense architecture, with different distributed implementations
- **Goal:** Increased resilience and possible to integrate with legacy systems



- **Future work:**
  - Combinations of attack/fault models (Case Study 1 and 2)
  - Trade-off analysis in resilient control: Decreased control authority/performance vs increased resilience

References

Case Study 1 (NIMBUS):

- "*Cyber-Physical-Security Framework for Building Energy Management System*," 2016 ACM/IEEE 7th International Conference on Cyber-Physical Systems (ICCPS)
- "*A Framework for Attack-resilient Industrial Control Systems: Attack Detection and Controller Reconfiguration,*" Proceedings of the IEEE, 2017

Case Study 2 (AIT SmartEST Lab):

- "*Voltage control for interconnected microgrids under adversarial actions*," 2015 IEEE 20th Conference of Emerging Technologies & Factory Automation (ETFA)
- "*SPARKS Cyber Security Demonstration Outcomes*," SPARKS Deliverable 6.4, 2017
- Demo movie: https://youtu.be/oLMKPVQv8yk

KTH ROYAL INSTITUTE OF TECHNOLOGY

Resilient Smart Grid Control: Two Case Studies

Henrik Sandberg

hsan@kth.se

Department of Automatic Control, School of Electrical Engineering

KTH, Stockholm, Sweden