State Estimation and Contingency Analysis of the Power Grid in a Cyber-Adversarial Environment

Robin Berthier
University of Illinois, Urbana-Champaign

Rakesh Bobba
Oregon State University

State Estimation and Contingency Analysis of the Power Grid in a Cyber-Adversarial Environment

Robin Berthier[1], Rakesh Bobba[1], Matt Davis[2], Kate Rogers[2], and Saman Zonouz[3]

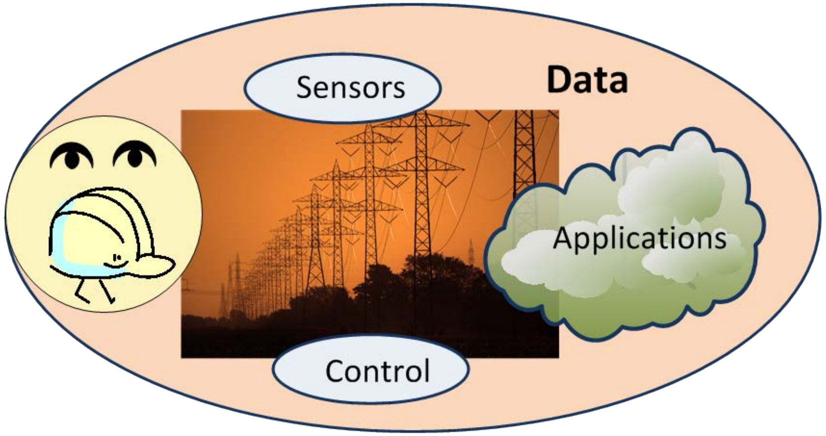
[1]Information Trust Institute
University of Illinois at Urbana-Champaign
Urbana, IL, USA
{rgb, rbobba}@illinois.edu

[2]PowerWorld Corporation
Champaign, IL, USA
{matt, kate}@powerworld.com

[3]Department of Electrical and Computer Engineering
University of Miami
Miami, USA
s.zonouz@miami.edu

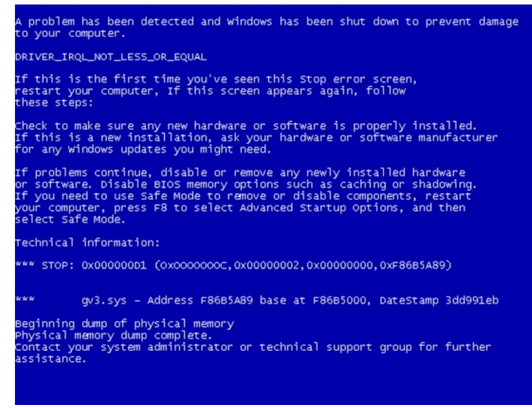
Motivation

- New technologies and new resources
- Extensive data integration
  - Sensory data
  - Control data
- Complex dependencies
- Stringent requirements



Security vs. Dependability

- Dependability and fault tolerance
  - ***Accidental** failures*
  - *Second party is the (unintentional) nature*
    - *Future action set can (probabilistically) be predicted*
  - *Traditional probabilistic analysis/modeling*

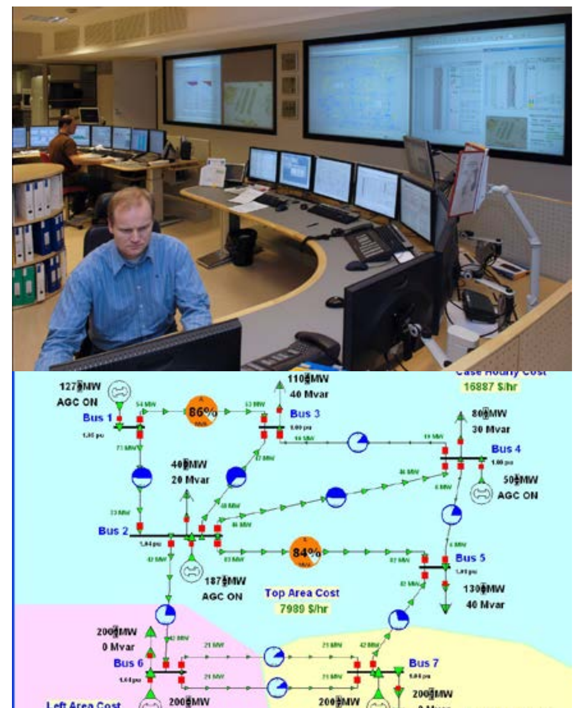


- Security and intrusion tolerance
  - ***Malicious** failures*
  - *Second party are (intentional) attackers*
    - *If predicted, they can exploit the prior information to damage further*
  - New solutions are needed…

Cyber-Physical System Security

- Systems in which cyber & physical systems are tightly integrated
  - Power systems
  - Process control networks
  - …
- (Potentially) more catastrophic security incidents…

Targeting nuclear plants



Power Control Network

Outline

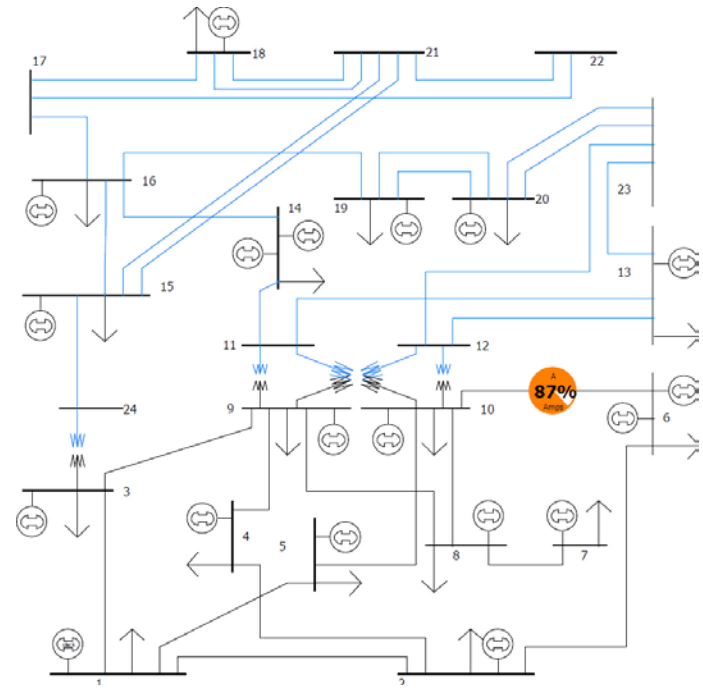
- Power Grid Operation
  - Cyber-physical relationships
  - State estimation
- Cyber-Physical Threat Model
  - Step-1: Cyber network exploits
  - Step-2: Physical system-aware attacks
- Defense Solutions
  - Cyber network intrusion detection
  - System-aware detection and protection
    - Measurement protection and bad-data detection
  - System contingency analysis

Power Grid Operation

Cyber-physical relationships

Power System Structure

- Major components:
  - Generators: produce electricity
  - Loads: consume electricity
  - Lines (T&D): transport energy from generators to loads



- Key Features
  - Absence of large-scale storage capabilities
  - Constraints: power balance, Kirchhoff's laws
  - Power flows through paths of "least resistance"
  - "Just-in-time" type manufacturing system

Operation and Control

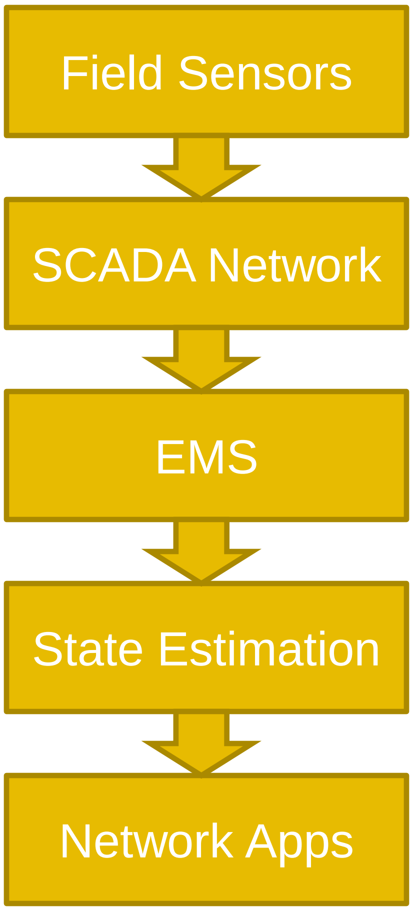
- ***Economics*** and ***reliability*** are the key drivers in power system operations and control
- Economics leads to large optimization problems for
  - Resource scheduling via unit commitment
  - Least-cost dispatch of available generation
- Reliability requirements typically entail no violations of physical limits and voltages and frequencies within prescribed bounds
  - Continuous monitoring
  - Hierarchical control architecture

Monitoring and Control

- Large and complex hardware-software systems are used for real-time operations and control
  - Energy management system (EMS)
  - Supervisory control and data acquisition (SCADA)
- Frequency is closely monitored and maintained around 60 Hz
  - Area control error (ACE) is measure for frequency excursions as well as deviations from scheduled interchanges – ideally, it should be ***zero***
  - Automatic generation control (AGC) implements proportional-integral-derivative (PID) control to keep ACE = ***zero***

Power System Operations

Data flow in power system operations



Sensors are becoming faster and more intelligent (e.g., PMUs)

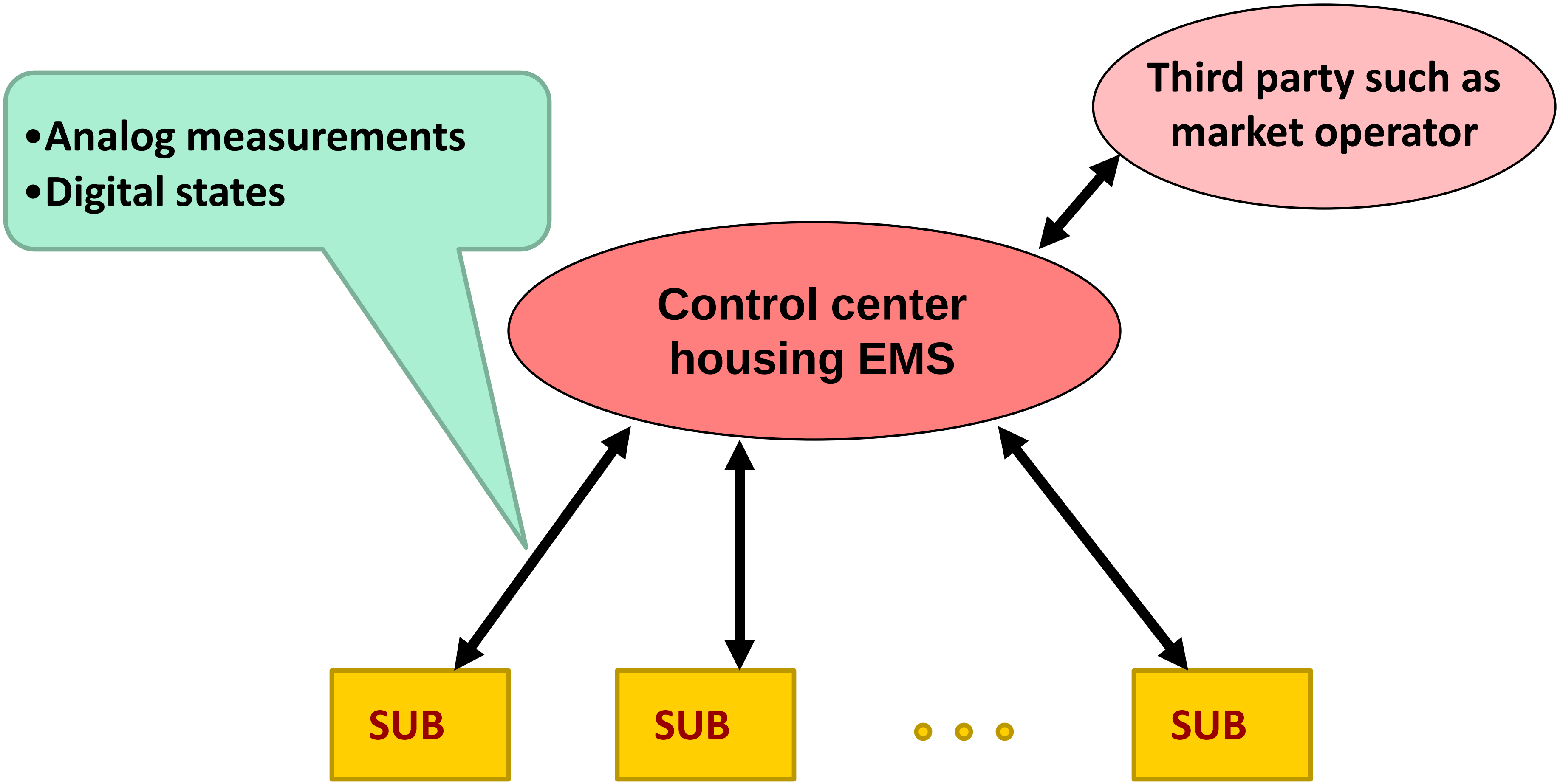
SCADA networks that have traditionally been serial or microwave links are becoming network based

Network Apps include real time contingency analysis on the state estimated model

Power Grid Operation

State Estimation

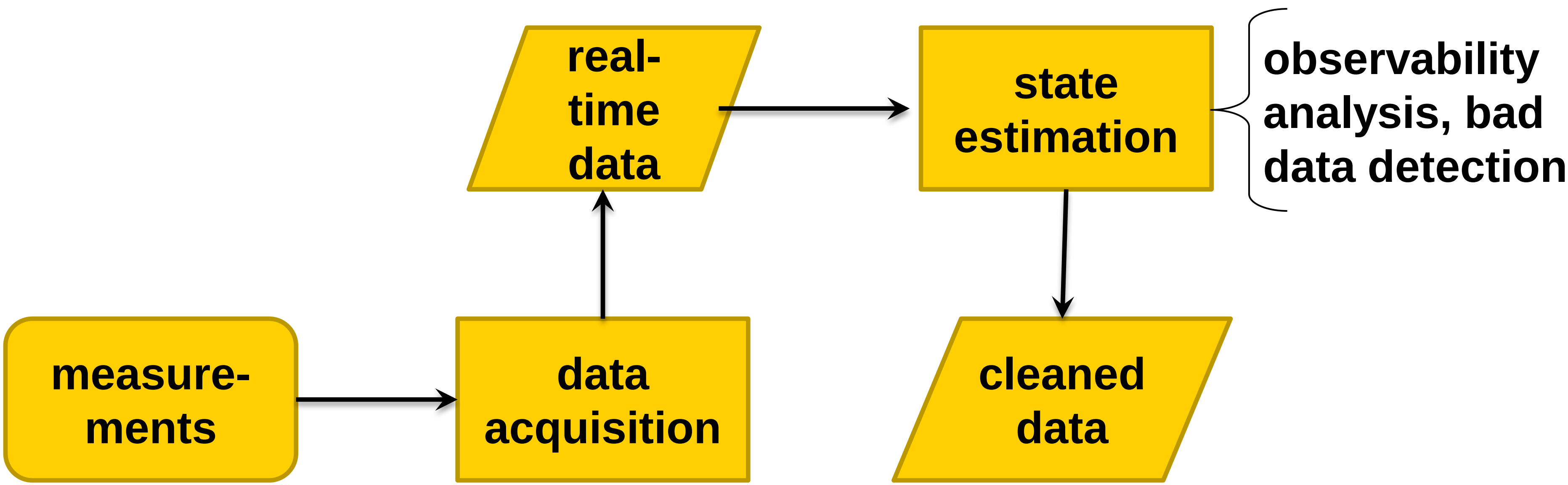
Power Grid Observability



* Figure source: Anupama Kowli and Anjan Bose

State Estimation

- Key process in power system operation and control
- Problem statement: given certain measurements, find the ***states*** (voltages and angles) of the system



* Figure source: Anupama Kowli

State Estimation

- The power flow is the central tool of power system planners and operators

Inputs:
System topology
Generation output
Load values

Outputs:
Voltage magnitude and angle
Line flows

$$\mathbf{P_{ij}} = \mathbf{V_i^2}[-\mathbf{G_{ij}}] + \mathbf{V_i}\mathbf{V_j}[\mathbf{G_{ij}}\cos(\theta_i - \theta_j) + \mathbf{B_{ij}}\sin(\theta_i - \theta_j)]$$

$$\mathbf{Q_{ij}} = \mathbf{V_i^2}[-\mathbf{G_{ij}}] + \mathbf{V_i}\mathbf{V_j}[\mathbf{G_{ij}}\sin(\theta_i - \theta_j) + \mathbf{B_{ij}}\cos(\theta_i - \theta_j)]$$

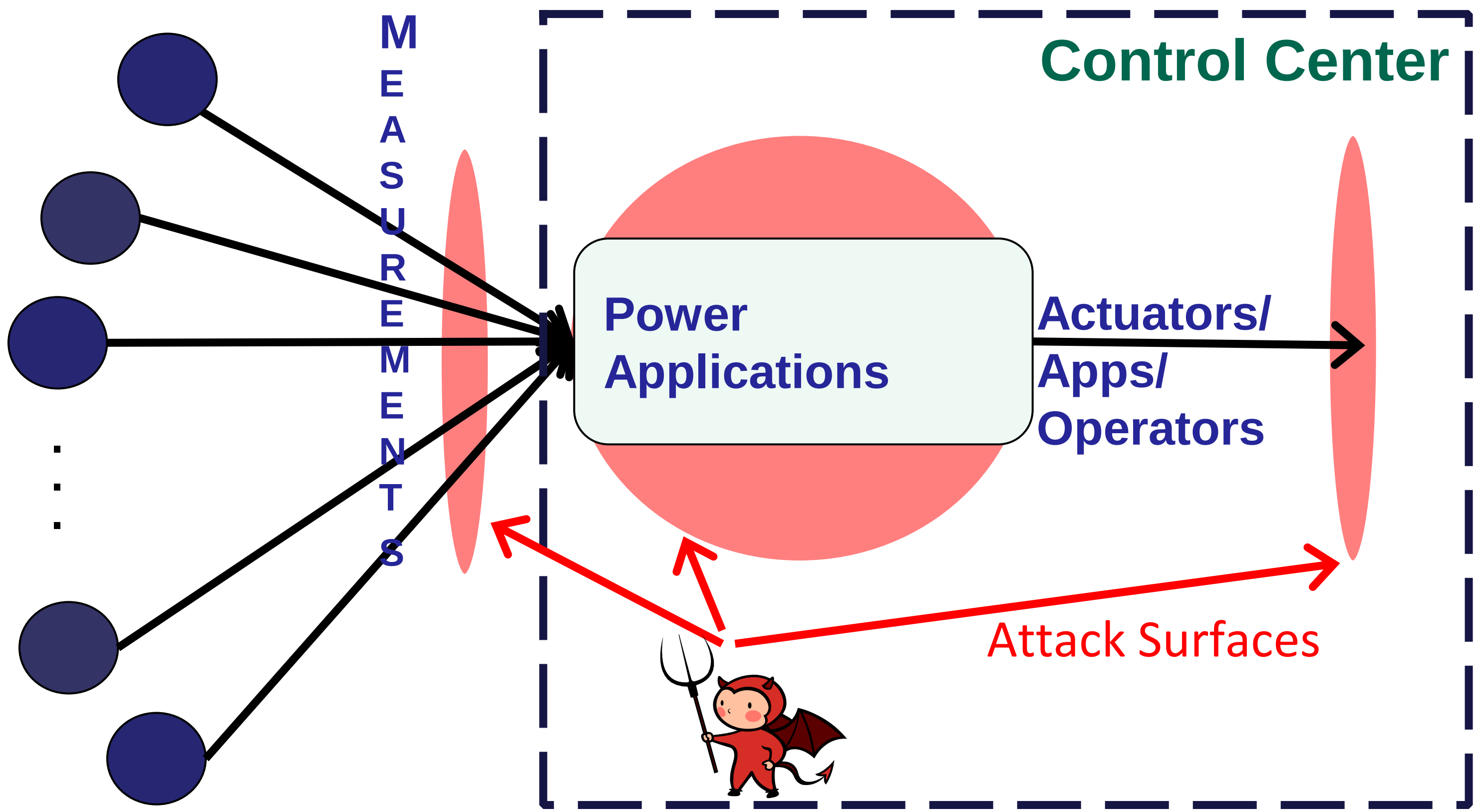
- Fundamentally, the power flow enforces the conservation of power at every Kirchoff’s voltage law node in the system

Cyber-Physical Threat Model

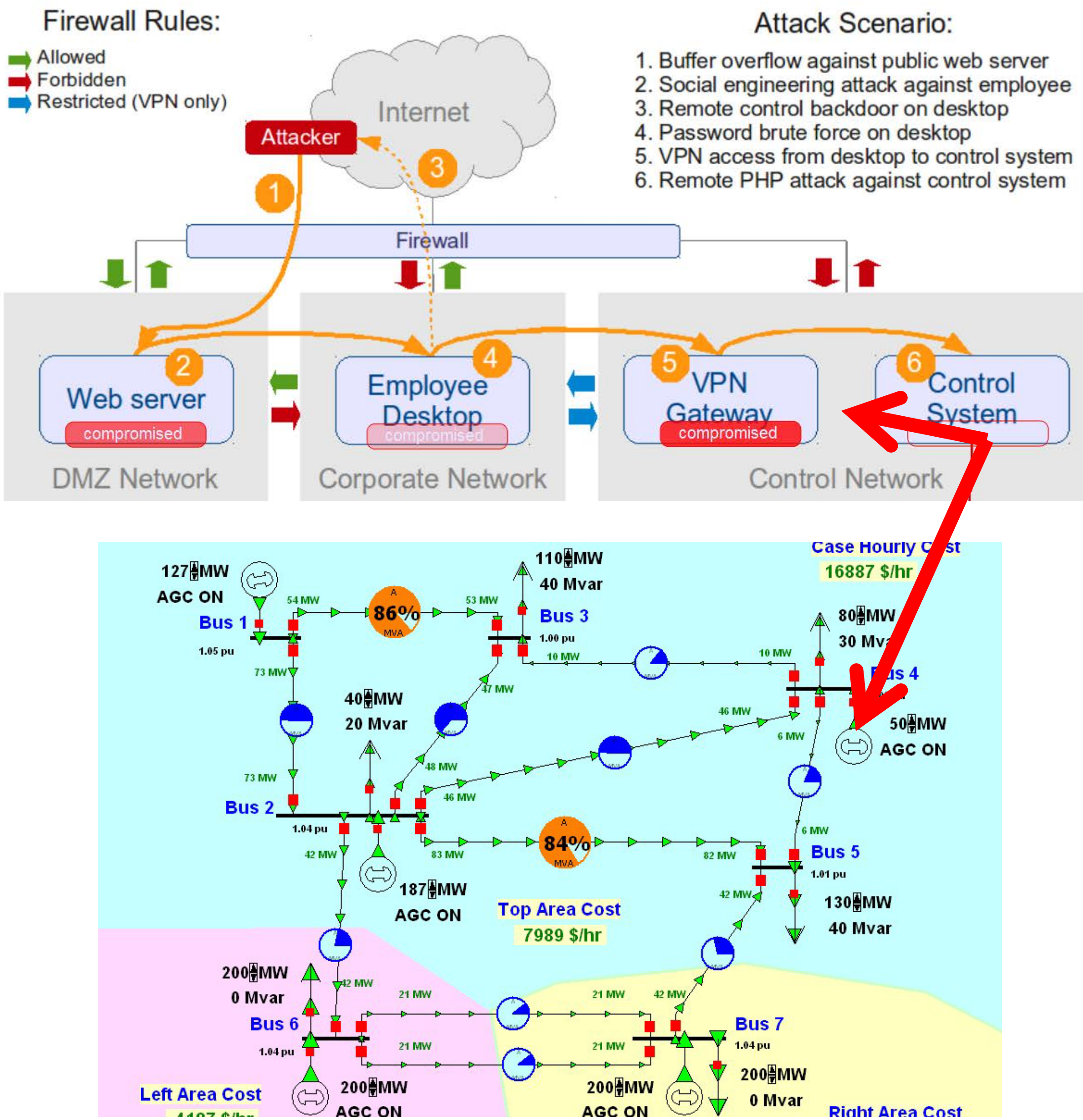
Step-1: Cyber network exploits

Step-2: Physical system-aware attacks

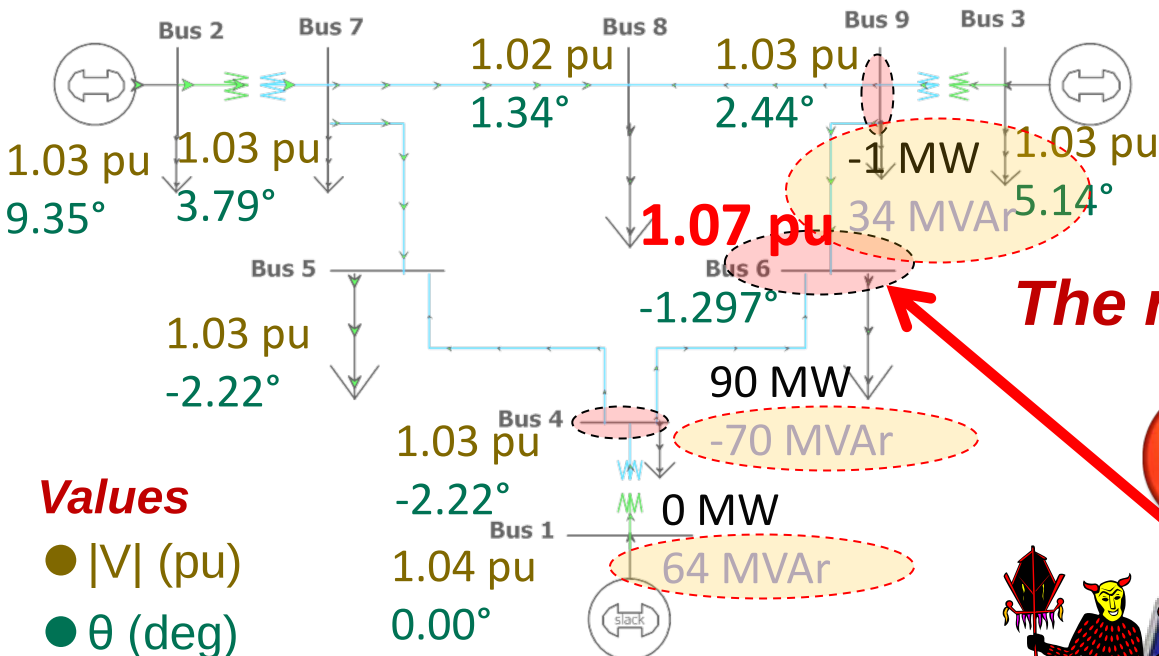
Cyber-Physical Threat



Network Exploits



False Data Injection on State Estimation



Values

- |V| (pu)
- θ (deg)
- P load (MW)
- Q load (MVAr)

Attack design:
Specifically chosen to satisfy the AC power flow solution equations

All states at non-malicious buses are preserved!

Defense Solutions

Cyber Network Intrusion Detection

Intrusion Detection Techniques

Legitimate Actions/Protocol Specification	Malicious Actions

Anomaly-based
+ detect unknown attacks
+ high scalability
- no root cause
- high false positive rate

Signature-based
+ low false positive rate
+ attack root cause
- require frequent update
- limited to known attacks

Specification-based
+ detect unknown attacks
+ high accuracy
- poor scalability
- high development cost

Specification-based Intrusion Detection

- Opportunities:
  - Leverage tight control over communication protocols and system behavior
  - Specification-based:
    - Little requirements about existing attacks
    - Ability to detect unknown attacks
    - No frequent update required
  - Enable the use of mathematical proof (formal methods)

- Challenges:
  - Scalability: stateful protocol analysis is resource intensive
  - Development costs: every protocol/application has to be specified

Solution Overview*

Offline development process:

Protocol

Network

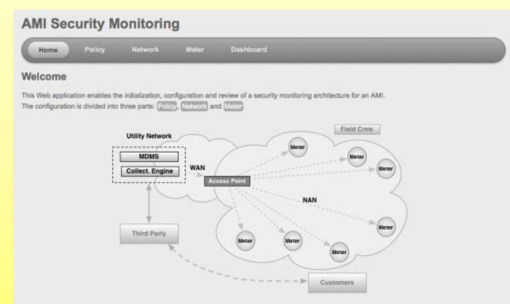
Use cases

Build specification-based checkers

Mathematically prove coverage of security policy

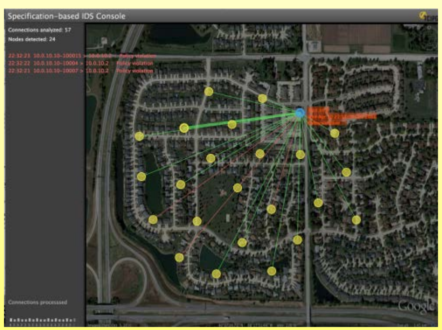
Online operation process:

Tune policy to system



Deploy config. on sensors in the field

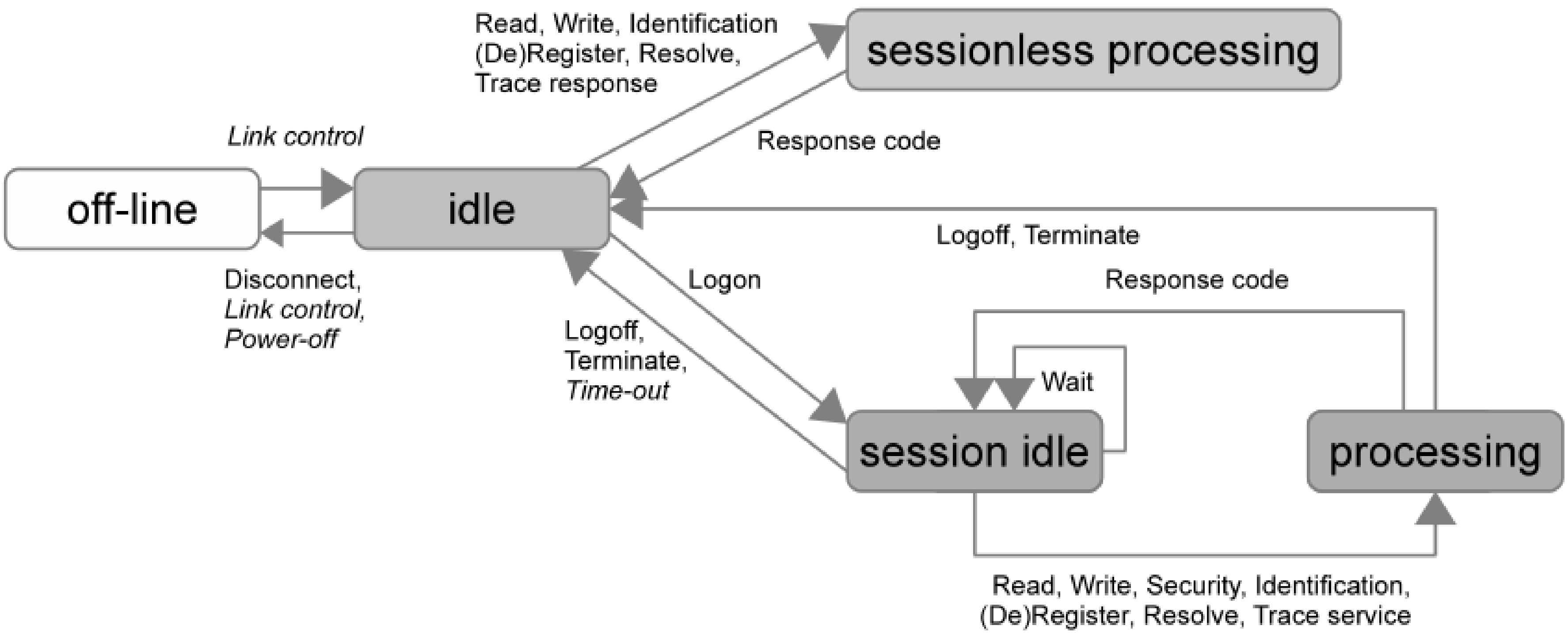
Situational Awareness



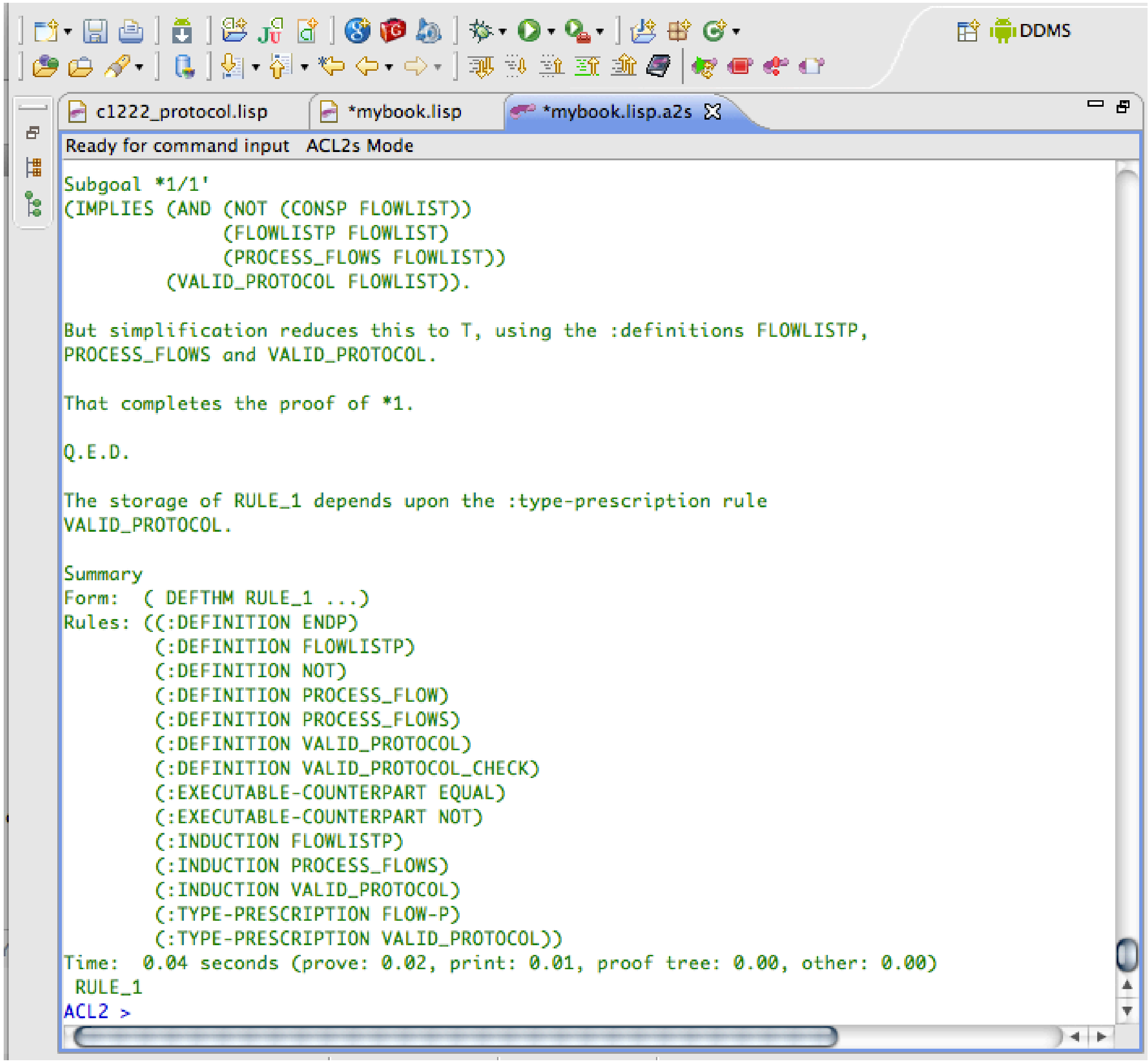
*Robin Berthier, William Sanders: Specification-Based Intrusion Detection for Advanced Metering Infrastructures. PRDC 2011: 184-193

Formal Verification of C12.22 protocol

- Validation through state machine:



Formal Verification (cont.)



Attack Detection

- Violations at the network level

Type	*Feature*	*Extracted automatically*
Access	Origin/Dest.	From CE to meter
Data	Protocol	C12.22 over TCP/IP
Temporal	Frequency	1-2 per 1000 meters per day
Resource	Session size	< 100 bytes

- Violations at the application level

Type	*Feature*	*Extracted automatically*
Access	C12.19 tables	Table 0 (read), Table 3 (write)
Data	C12.19 values	Table 3, data: 0x01, offset: 0x00
Temporal	Session duration	< 1 minute
Resource	Services used	Logon, Full read, Partial write, Logoff

Defense Solutions (cont.)

System-aware detection and protection

Power-System Measurement Protection and Bad-data Detection

Current Bad Data Detection Solutions: Residual-Based Approaches

- Need to account for possibility of bad data
  - ***Bad data*** definition from (*): “measurements that are grossly in error”
  - Bad data can potentially result in incorrect power-state estimates
- Measurement residuals – typical bad data detection for state estimation

  if $||\mathbf{z} - \mathbf{H}\underline{\mathbf{x}}|| \le \tau$ *no bad measurements*
- Goal of residual approaches: detect corrupted power measurements

* A. Monticelli, State estimation in electric power systems: a generalized approach. Kluwer Academic Publishers, 1999.

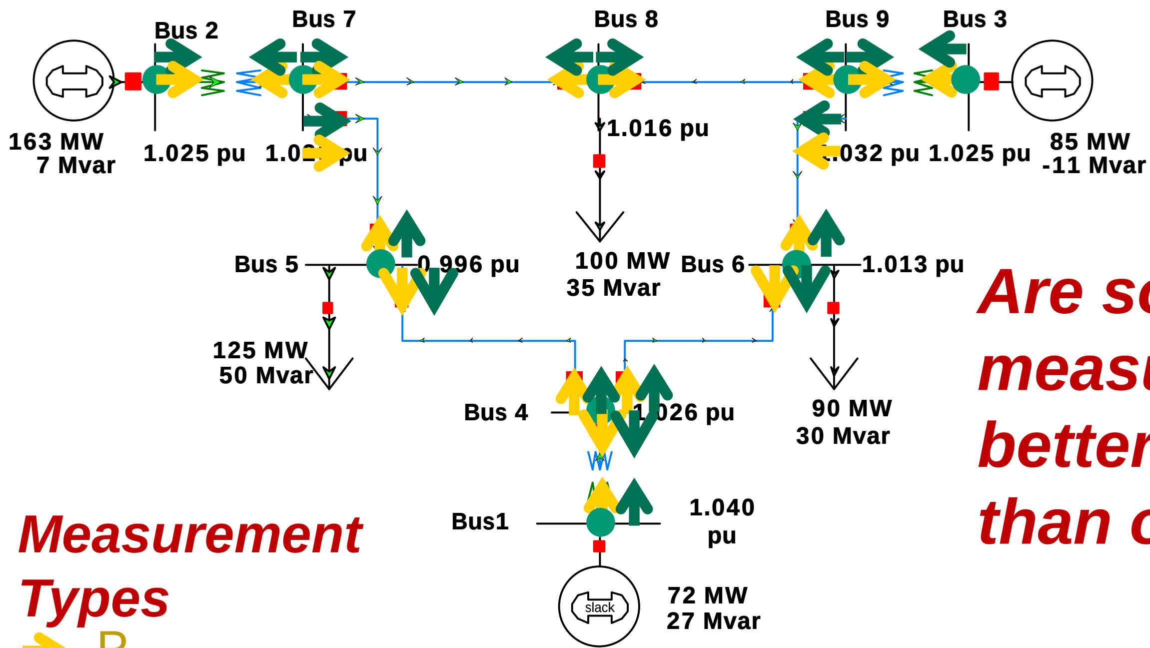
Bad Data Detection: Residual -Based Approaches

- Coordinated attacks can work by creating “interacting bad-measurements” that satisfy the power flow solution equations, making them difficult or impossible to detect using conventional means

- ***Residual-based approaches may be fundamentally insufficient against coordinated security compromises***

- One obvious approach:
  - Protect all measurements from compromises

System-Aware Measurement Protection



Are some measurements better to protect than others?

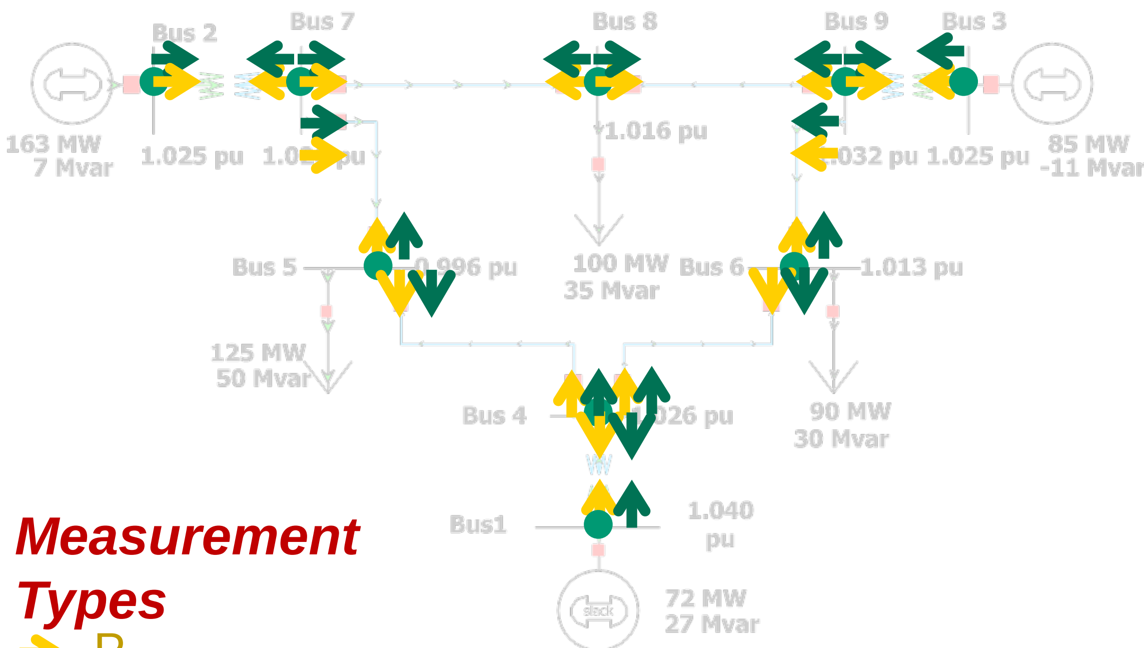
Measurement Types

$P_{i,j}$

$Q_{i,j}$

V_i

System-Aware Measurement Protection



Example: Basic Measurements

	i	j
P_{ij}	4	1
P_{ij}	2	7
P_{ij}	9	3
P_{ij}	5	4
P_{ij}	6	4
P_{ij}	7	5
P_{ij}	7	8
P_{ij}	8	9
Q_{ij}	4	1
Q_{ij}	8	9
Q_{ij}	7	2
Q_{ij}	3	9
Q_{ij}	4	5
Q_{ij}	4	6
Q_{ij}	5	7
Q_{ij}	8	7

Measurement Types

- $P_{i,j}$
- $Q_{i,j}$
- V_i

We show that no attacks are possible if H'_k has full rank

$$\begin{bmatrix} \mathbf{0} \\ \mathbf{a}_k \end{bmatrix} = \left[\begin{array}{c|c} \mathbf{H}'' & \mathbf{H}'_k \\ \mathbf{H}_k' & \mathbf{H}_{kk} \end{array}\right] \begin{bmatrix} \mathbf{0} \\ \mathbf{c}_k \end{bmatrix}$$

$$\mathbf{0} = \mathbf{H}'_k \mathbf{c}_k$$

$$\mathbf{a}_k = \mathbf{H}_{kk} \mathbf{c}_k$$

Accomplished by protecting ***basic measurements***

Cost-Optimal Measurement Protection

- Protect a set of *Basic Measurements**
  - it is necessary but not sufficient to protect *n* measurements, to detect stealthy false data injection attacks
  - it is necessary and sufficient to protect a set of **basic measurements (BM)** to detect stealthy false data injection attacks
  - approaches to identify **BM** already exist and well-studied
  - choices are available – the set of **BM** is not unique
  - each verifiable state variable (e.g., PMU) reduces number of measurements to be protected by one
  - approach validated on the IEEE 9,14,30,118, and 300 bus test systems

*R. B. Bobba, K. M. Rogers, Q. Wang, H. Khurana, K. Nahrstedt, T. J. Overbye, “Detecting False Data Injection Attacks on DC State Estimation,” *First Workshop on Secure Control Systems (SCS 2010)*, April 2010.

Defense Solutions (cont.)

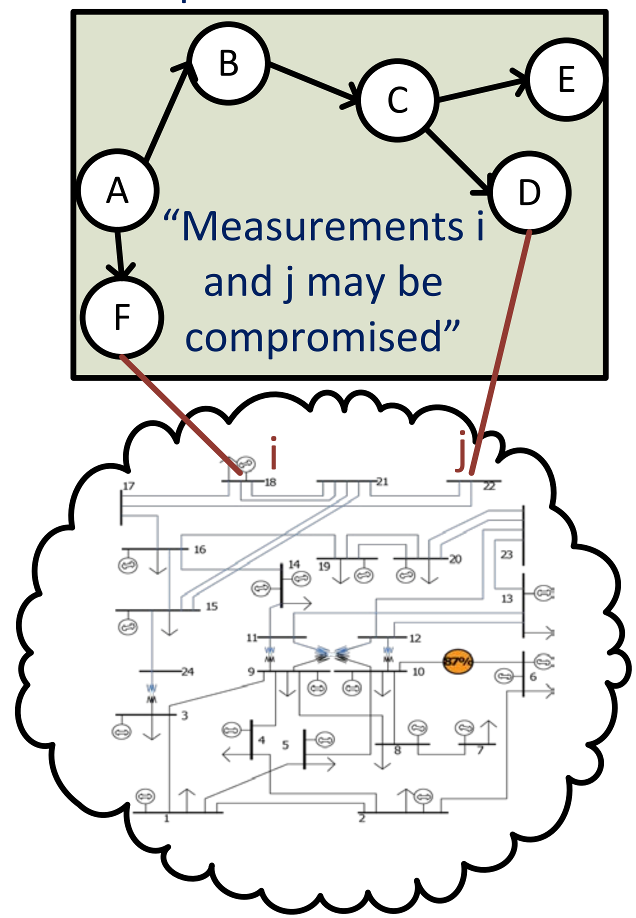
Integrated Cyber-Physical State Estimation

Cyber-Physical State Estimation (CPSE)*

- Co-utilize information from ***cyber*** and ***power*** network to (more precisely) determine the ***state*** of the ***cyber-physical*** system

- Use combined ***information state*** to provide a scalable approach to detecting bad data caused by a cyber event

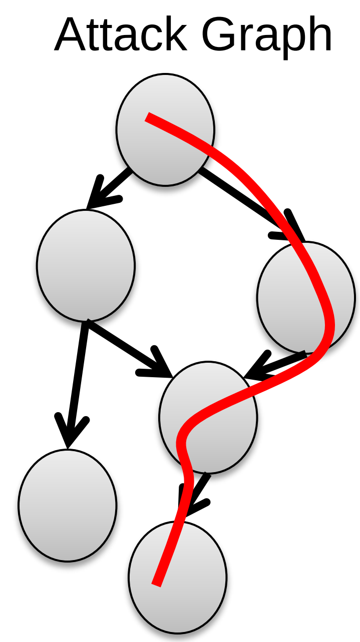
Example



*S. A. Zonouz, K. M. Rogers, R. Berthier, R. B. Bobba, W. H. Sanders, T. J. Overbye, "CPIDS: A Cyber-Physical Intrusion Detection System for Power-Grid Critical Infrastructures," in review for *IEEE Transactions on Smart Grid*.

Algorithm Step 1: Potentially-bad Data Identification

- From IDS reports, we (probabilistically) know attacker's current privileges
  → From power network's topology, we know which measurements could/might have been modified by the adversary



- Example:
  - network's topology
    - i-th measurement (by PMU_i): real power of the bus B2
  - IDS alerts
    - PMU_i is compromised
      → i-th measurement might have been corrupted!

Algorithm Step 2: Power State Estimation & Verification

- Throw the potentially-bad data away, and run a power state estimation using the remaining power measurements

$$\mathbf{P_{ij}} = \mathbf{V_i^2}[-\mathbf{G_{ij}}] + \mathbf{V_i}\mathbf{V_j}[\mathbf{G_{ij}}\cos(\theta_i - \theta_j) + \mathbf{B_{ij}}\sin(\theta_i - \theta_j)]$$

$$\mathbf{Q_{ij}} = \mathbf{V_i^2}[-\mathbf{G_{ij}}] + \mathbf{V_i}\mathbf{V_j}[\mathbf{G_{ij}}\sin(\theta_i - \theta_j) + \mathbf{B_{ij}}\cos(\theta_i - \theta_j)]$$

- Compute $\| \mathbf{z} - \mathbf{H}(\hat{\mathbf{x}}) \|$, and identify the corrupted measurements
  - based on how much they differ from their estimates

CPSE Benefits

- Improved Bad-data Detection
  - Accuracy and Scalability
- Quick State Estimation Convergence
- Improved State Estimates

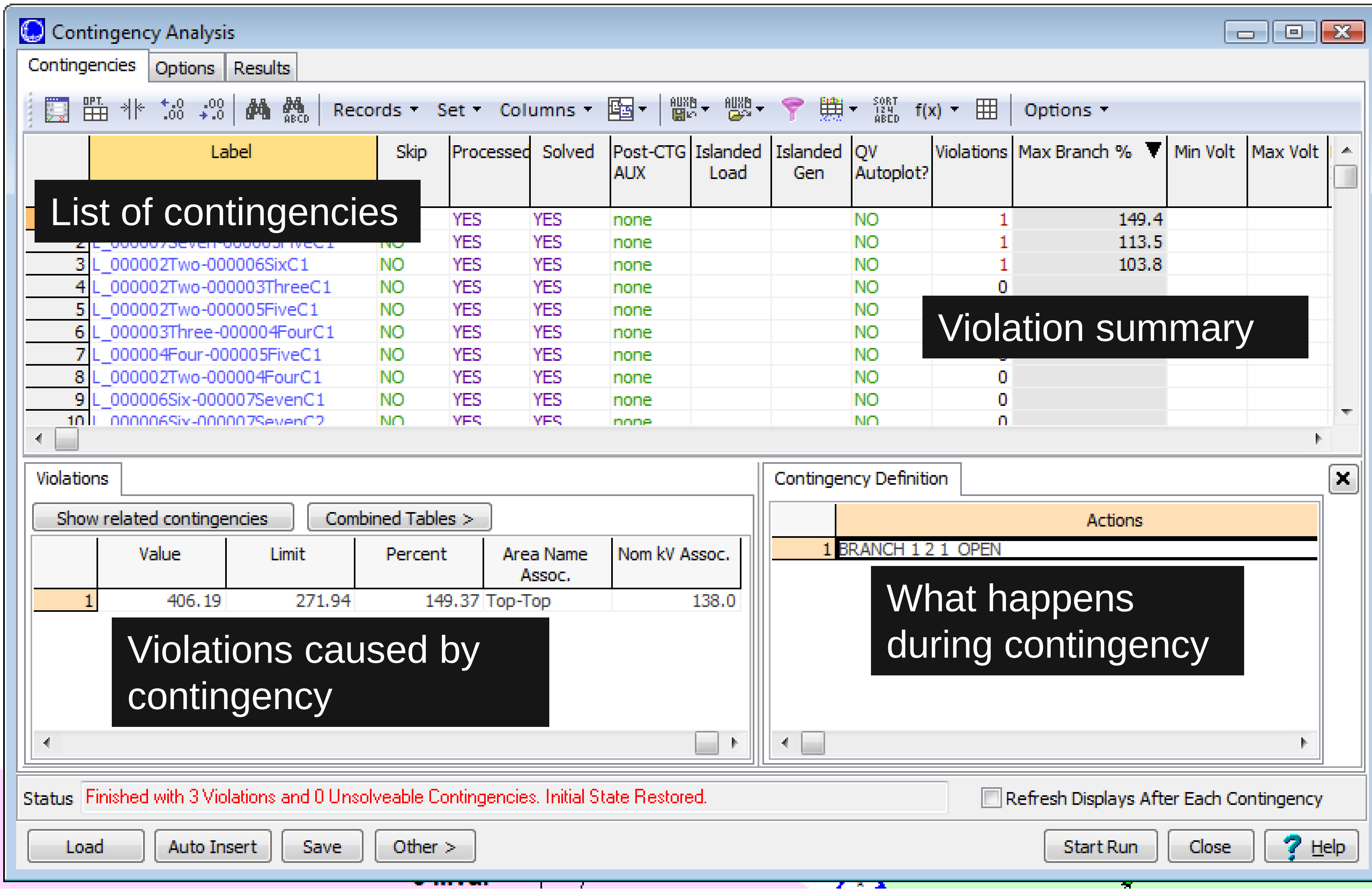
Defense Solutions (cont.)

System Contingency Analysis

Contingency Analysis (CA)

- Contingency analysis is a fundamental tool of power systems analysis
- Typically, a contingency analysis works with a power system model (power flow case) to determine potential problems
  - Full topology (node breaker) vs. planning models (bus branch)
- Answers the question: "*What happens when X goes out of service?*"

Contingency Analysis Results



CA in Power System Operations

- State estimator runs every 2min or so
- After getting the state estimate real time contingency analysis (RTCA) runs on the estimated model
  - The list of contingencies must be picked carefully before being added to the RTCA contingency list
  - The RTCA list needs to include important contingencies, but it is time constrained

CA Solution Methods

- There are several ways of solving the contingency analysis
  - Full AC power flow (Slowest, Most accurate)
  - DC power flow (Fast, no voltage/var information)
  - Linear sensitivities (Fast, less sensitive to topology)
- There is the traditional engineering tradeoff between accuracy and speed
- All solution methods are used in practice

CA Solution Details

- Modeling a contingency accurately can be an intricate process
- The devil is in the details
- A few of the things that must be accounted for
  - Voltage controller and phase shifter response
  - AGC response
  - Special protection schemes / Breaker actions
  - Contingency modeling (full topology vs planning model)
- There is a lot that happens when a contingency is solved or even solving a power flow case

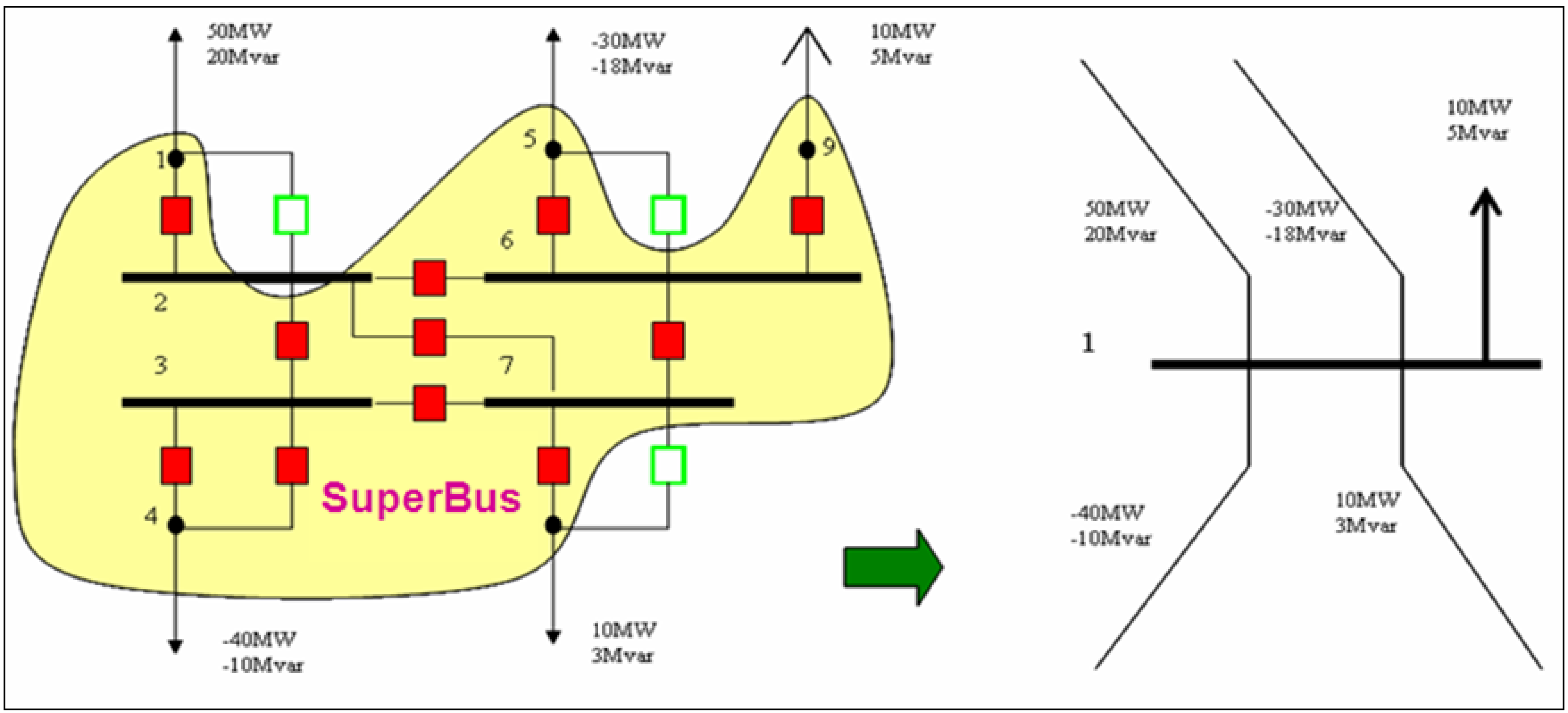
EMS and Planning Models

EMS Model

- Used for real-time operations
- Call this ***Full-Topology*** model
- Has node/breaker detail

Planning Model

- Used for off-line analysis
- We call this ***Consolidated*** model



Traditional Contingency Analysis (CA)

- The “N-1” criteria is used to operate the system so that there will be no violations when any one element is taken offline
- Future requirements are strengthening the security criteria (“N-1-1”) meaning many more contingencies need to be solved*
  - Once multiple outages begin to be considered, the size of the contingency list can grow very large
  - For 1000 lines
    - N-1 means solving 1000 line outages
    - N-2 means solving 499500 line outages (1000 choose 2)

*Charles Davis, Thomas Overbye: Linear Analysis of Multiple Outage Interaction. HICSS 2009: 1-8

Proposed System Contingency Analysis

- Question: "*What happens when X goes out of service?*"
  - X could be either a critical power component or cyber asset.
- Unlike traditional scenarios, cyber asset outages may be due to cyber adversaries
- Ongoing Research Topic!

Conclusions

- Criticality of cyber-physical infrastructure security:
  - Complex relationship between cyber and physical components
  - Importance of accurate state estimation → target of interest for adversaries:
    - Step-1: Cyber network exploits
    - Step-2: Physical system-aware attacks
- Requirements for advanced defense solutions:
  - Specification-based network intrusion detection tailored for cyber-physical system characteristics
  - System-aware measurement protection and bad-data detection
  - System-wide contingency analysis
- Contingency analysis as potential solution for a unified cyber-physical state estimation

Questions?

Robin Berthier rgb@illinois.edu

Saman Zonouz s.zonouz@miami.edu