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

5 authors, including:

Robin Berthier
University of Illinois, Urbana-Champaign
42 PUBLICATIONS 1,827 CITATIONS

Rakesh Bobba
Oregon State University
43 PUBLICATIONS 1,750 CITATIONS

All content following this page was uploaded by Rakesh Bobba on 22 May 2015.

The user has requested enhancement of the downloaded file.
State Estimation and Contingency Analysis of the Power Grid in a Cyber-Adversarial Environment

Robin Berthier\textsuperscript{1}, Rakesh Bobba\textsuperscript{1}, Matt Davis\textsuperscript{2}, Kate Rogers\textsuperscript{2}, and Saman Zonouz\textsuperscript{3}

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

\textsuperscript{2}PowerWorld Corporation
Champaign, IL, USA
\{matt, kate\}@powerworld.com

\textsuperscript{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…
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
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 a measure for frequency excursions as well as deviations from scheduled interchanges – ideally, it should be $\text{zero}$
  - Automatic generation control (AGC) implements proportional-integral-derivative (PID) control to keep $\text{ACE} = \text{zero}$
Power System Operations

Data flow in power system operations

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

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

EMS

State Estimation

Network Apps
- Network Apps include real time contingency analysis on the state estimated model
Power Grid Operation

State Estimation
Power Grid Observability

- Analog measurements
- Digital states

Control center housing EMS

Third party such as market operator

* 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

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

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

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

Power Applications

Actuators/Apps/Operators

Control Center

MEASUREMENTS

Attack Surfaces
Network Exploits

**Firewall Rules:**
- Allowed
- Forbidden
- Restricted (VPN only)

**Internet**

**Attack Scenario:**
1. Buffer overflow against public web server
2. Social engineering attack against employee
3. Remote control backdoor on desktop
4. Password brute force on desktop
5. VPN access from desktop to control system
6. Remote PHP attack against control system

**Network Diagram:**
- **DMZ Network**
  - Web server (compromised)
- **Corporate Network**
  - Employee Desktop (compromised)
- **Control Network**
  - VPN Gateway (compromised)
- **Control System**

**Case Hourly Cost:** 16887 $/hr

**Top Area Cost:** 7989 $/hr

**Left Area Cost:** 2798 $/hr

**Right Area Cost:** 2983 $/hr
False Data Injection on State Estimation

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

All states at non-malicious buses are preserved!

**Values**
- $|V|$ (pu)
- $\theta$ (deg)
- $P$ load (MW)
- $Q$ load (MVAr)

| $|V|$ (pu) | $\theta$ (deg) | $P$ load (MW) | $Q$ load (MVAr) |
|-----------|----------------|---------------|-----------------|
| 1.03 pu   | 9.35°          | -1 MW         | 34 MVAr         |
| 1.03 pu   | 3.79°          | 90 MW         | -70 MVAr        |
| 1.03 pu   | -2.22°         | 0 MW          | 64 MVAr         |
| 1.03 pu   | 1.34°          | 1.03 pu       | 1.03 pu         |
| 1.03 pu   | 2.44°          | 1.07 pu       | 1.07 pu         |
| 1.03 pu   | -2.22°         | 1.04 pu       | 0.00°           |
| 1.03 pu   | 1.04 pu        | 1.04 pu       | 1.04 pu         |
| 1.03 pu   | -1.297°        | 1.07 pu       | 1.07 pu         |
| 1.04 pu   | 1.34°          | 1.03 pu       | 1.03 pu         |
| 1.04 pu   | 2.44°          | 1.03 pu       | 1.03 pu         |
| 1.04 pu   | -2.22°         | 1.04 pu       | 0.00°           |
| 1.04 pu   | 1.34°          | 1.03 pu       | 1.03 pu         |
| 1.04 pu   | 2.44°          | 1.03 pu       | 1.03 pu         |
| 1.04 pu   | -2.22°         | 1.04 pu       | 0.00°           |
Defense Solutions

Cyber Network Intrusion Detection
Intrusion Detection Techniques

**Legitimate Actions/Protocol Specification**
- **Anomaly-based**
  + detect unknown attacks
  + high scalability
  - no root cause
  - high false positive rate

**Malicious Actions**
- **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:

Situational Awareness

Deploy config. on sensors in the field

Tune policy to system

Formal Verification of C12.22 protocol

- Validation through state machine:
Formal Verification (cont.)

Subgoal "1/1"
(IMPLIES (AND (NOT (CONSP FLOWLIST))
  (FLOWLISTP FLOWLIST)
  (PROCESS_FLOWS FLOWLIST))
  (VALID_PROTOCOL FLOWLIST)).

But simplification reduces this to T, using the :definitions FLOWLISTP,
PROCESS_FLOWS and VALID_PROTOCOL.

That completes the proof of *1.

Q.E.D.

The storage of RULE_1 depends upon the :type-prescription rule
VALID_PROTOCOL.

Summary
Form: (DEFTHM RULE_1 ...)
Rules: (DEFINITION ENDP)
  (DEFINITION FLOWLISTP)
  (DEFINITION NOT)
  (DEFINITION PROCESS_FLOW)
  (DEFINITION PROCESS_FLOWS)
  (DEFINITION VALID_PROTOCOL)
  (DEFINITION VALID_PROTOCOL_CHECK)
  (EXECUTABLE-COUNTERPART EQUAL)
  (EXECUTABLE-COUNTERPART NOT)
  (INDUCTION FLOWLISTP)
  (INDUCTION PROCESS_FLOWS)
  (INDUCTION VALID_PROTOCOL)
  (TYPE-PRESCRIPTION FLOW-P)
  (TYPE-PRESCRIPTION VALID_PROTOCOL)

Time: 0.04 seconds (prove: 0.02, print: 0.01, proof tree: 0.00, other: 0.00)
RULE_1
ACL2 >
Attack Detection

• Violations at the network level

<table>
<thead>
<tr>
<th>Type</th>
<th>Feature</th>
<th>Extracted automatically</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access</td>
<td>Origin/Dest.</td>
<td>From CE to meter</td>
</tr>
<tr>
<td>Data</td>
<td>Protocol</td>
<td>C12.22 over TCP/IP</td>
</tr>
<tr>
<td>Temporal</td>
<td>Frequency</td>
<td>1-2 per 1000 meters per day</td>
</tr>
<tr>
<td>Resource</td>
<td>Session size</td>
<td>&lt; 100 bytes</td>
</tr>
</tbody>
</table>

• Violations at the application level

<table>
<thead>
<tr>
<th>Type</th>
<th>Feature</th>
<th>Extracted automatically</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access</td>
<td>C12.19 tables</td>
<td>Table 0 (read), Table 3 (write)</td>
</tr>
<tr>
<td>Data</td>
<td>C12.19 values</td>
<td>Table 3, data: 0x01, offset: 0x00</td>
</tr>
<tr>
<td>Temporal</td>
<td>Session duration</td>
<td>&lt; 1 minute</td>
</tr>
<tr>
<td>Resource</td>
<td>Services used</td>
<td>Logon, Full read, Partial write, Logoff</td>
</tr>
</tbody>
</table>
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
  \[ | | \mathbf{z} - H\mathbf{x} | | \leq T \quad \text{no bad measurements} \]

- Goal of residual approaches: detect corrupted power measurements

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

<table>
<thead>
<tr>
<th>i</th>
<th>j</th>
</tr>
</thead>
<tbody>
<tr>
<td>P_{ij}</td>
<td>4</td>
</tr>
<tr>
<td>P_{ij}</td>
<td>2</td>
</tr>
<tr>
<td>P_{ij}</td>
<td>9</td>
</tr>
<tr>
<td>P_{ij}</td>
<td>5</td>
</tr>
<tr>
<td>P_{ij}</td>
<td>6</td>
</tr>
<tr>
<td>P_{ij}</td>
<td>7</td>
</tr>
<tr>
<td>P_{ij}</td>
<td>7</td>
</tr>
<tr>
<td>Q_{ij}</td>
<td>4</td>
</tr>
<tr>
<td>Q_{ij}</td>
<td>8</td>
</tr>
<tr>
<td>Q_{ij}</td>
<td>7</td>
</tr>
<tr>
<td>Q_{ij}</td>
<td>3</td>
</tr>
<tr>
<td>Q_{ij}</td>
<td>4</td>
</tr>
<tr>
<td>Q_{ij}</td>
<td>4</td>
</tr>
<tr>
<td>Q_{ij}</td>
<td>5</td>
</tr>
<tr>
<td>Q_{ij}</td>
<td>8</td>
</tr>
</tbody>
</table>

Measurement Types

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

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

$$
\begin{bmatrix}
0 \\
a_k
\end{bmatrix} = \begin{bmatrix}
H'' & H'_k \\
H'_k & H_{kk}
\end{bmatrix}
\begin{bmatrix}
0 \\
c_k
\end{bmatrix}
\Rightarrow
0 = H'_k c_k
\quad a_k = H_{kk} c_k
$$

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

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

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

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

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

- Compute \( ||z - H(\hat{x})|| \), 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

- List of contingencies
- Violation summary
- Violations caused by contingency
- What happens during contingency
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