Cyber Security for Smart Grid Devices

Annarita Giani
Electrical Engineering & Computer Sciences
University of California at Berkeley
agiani@eecs.berkeley.edu

Trustworthy Cyber Infrastructure for the Power Grid center here at Illinois
February 4, 2011
50 Years Ago

[Images of various technological advancements]
Outline

- Background
- Power Systems Background
- Phase Measurement Units
- State Estimation & PMU Data
- Our Approach to Integrity Attack Detection
Outline

- Background
- Power Systems Background
- Phase Measurement Units
- State Estimation & PMU Data
- Our Approach to Integrity Attack Detection
My Background

- PhD Dartmouth 2007
  - Detection of attacks on cognitive channels
  - [G. Cybenko]

- Post-doc TRUST Center [2007-2009]
  - Trustworthy information systems
  - [S. Sastry]

- Post-doc Berkeley [2009-]
  - Renewable integration, Cyber-security in power systems
  - [K. Poolla]
Security Objectives

- **Confidentiality**: information disclosure only to authorized users
  - Eavesdropping, Phishing
  - Access Control, Authentication, Authorization, Encryption

- **Integrity**: trustworthiness of information resources
  - Replay, Man in the Middle, Data Injection, Data Jam, Data Corruption
  - Encryption, Redundancy

- **Availability**: Availability of data whenever need it
  - Denial-of-Service
  - Traffic Anomaly Detection

- **Authorization**
- **Authentication**
- **Non Repudiation**
Security Objectives

- Misuse of user data (confidentiality)
- Grid resilience (availability)
- Trustworthiness of devices (integrity)
- Metrics
Current Work Summary

- Testbed for Secure and Robust SCADA Systems with Vanderbilt (Karsai) and CMU (Sinopoli) [IEEE Real-Time and Embedded Technology and Applications Symposium 2008]
- Optimal Contracts for Wind Power Producers in Electricity Markets (Poolla) [CDC 2010]
- Renewable integration and smart grid
- Integrity Attack Detection of PMU data [This talk] (Poolla, Khargonekar, Bitar)
Outline

- Background
- **Power Systems Background**
- Phase Measurement Units
- State Estimation & PMU Data
- Our Approach to Integrity Attack Detection
Context and Notation

- Considering AC synchronous power systems
- **Assume quasi steady-state analysis**
  Voltages and currents are well approximated as fixed frequency sinusoids with *slowly* changing phases

  \[
  v(t) = V \sin(\omega t + \phi) \\
  \mathbb{V} = V \exp(j\phi)
  \]

- **Notation**

  - \( M^* \): complex-conjugate transpose
  - \( \| \cdot \| \): standard euclidean norm
  - \( \sigma^2 \): noise variance
  - \( \mathbb{V}, \mathbb{I} \): phasors
  - \( Y = G + jB \): bus admittance matrix
  - \( G \): bus conductance matrix
  - \( B \): bus susceptance matrix
  - \( E \): expectation operator
Static State of a Power System

- **What is it?**
  The set of *voltage magnitudes and angles* at all network buses

- **Why is it important?**
  Bus voltages and angles are the key variables
  These determine
  - static flows on transmission lines
  - locational marginal prices
  - current stress state of system
  - future generation that should be scheduled
Measurements

- **Bus powers** [real, reactive] are commonly measured
  - Used for settlement of contract, compensation, etc
- **Bus voltages magnitudes** are easy to measure
  - Used for voltage regulation, system protection, etc
- **Bus voltage phases** are much **harder** to sense
  - Power flows depend on the phase difference between buses
  - Need global clock to determine times of voltage maxima
  - So, voltage phases are estimated
- **Dynamic state estimation**
  - Not commonly used
  - Computationally prohibitive
- **Static state estimation**
Static State Estimation

- **What is it?**
  
  Find the phase angles given:
  
  - measured real power $P$ and reactive power $Q$ at load buses
  - measured real power $P$ and voltage $V$ at generator buses

- **Current practice**
  
  - Data available every 1-15 minutes thru SCADA system

- **Load flow equations**
  
  - Over-determined set of algebraic nonlinear equations
  - Nonlinear programming to estimate states $V, \delta$
  - Takes **5-15 minutes** depending on problem size
  - Can have > 5000 buses
WAMS = wide area monitoring systems

Integral component of power system operation today
- Telemetry
- Data storage
- Alarming and status

Application
- Situational awareness
- Alarming and status (early warning)
- Root cause analysis of events
- State estimation
Today: SCADA Data

- Supervisory control and data acquisition (SCADA) data since the 1960’s
  - Voltage & Current Magnitudes
  - Frequency
  - Every 2-4 seconds
- Believed to be secure (not part of the commodity internet)
- Limitation
  - Low speed data acquisition
  - Steady state observability of the system
Outline

- Background
- Power Systems Background
- Phase Measurement Units
- State Estimation & PMU Data
- Our Approach to Integrity Attack Detection
Synchro Phasors

- Synchronized sampling with 1 microsecond accuracy using GPS
- Protocol: IEEE C37.118-2005 standard
- Cost: 2-3000$ each

Advantages of PMU Data

- PMUs collect location, time, frequency, current, voltage and phase angle (>40 Hz sampling)

- Why are they important?
  - Grid-scale renewable energy systems [ex: photovoltaic and wind]
  - Large unexpected variability
  - Can produce phase instability
  - Results in poor decision making [ex: scheduling]
  - Which can lead to big problems [ex: voltage instability, islanding, cascading failures]

- Directly provides the phase angles [from State Estimation to State Measurement]
PMU Architecture

- **Measurement Layer**
  - PMUs

- **Data Collection Layer**
  - Phasor Data Concentrator (PDC)
    - A hardware/software device
    - Performs precise time alignment of data from multiple PMUs
    - Usually centrally located
  - Archives, processes and display PMU data (optional)

- **Communication Network**
  - NASPInet

http://www.naspi.org/
North American SynchroPhasor Initiative (NASPI)
High speed for fast data streaming
Secure exchange of data
The owner of a phasor gateway that publishes the data to naspinet has full control of its data distribution
Pilot phase by 2014
Fully operational by 2019
NaspiNET Software Components

NaspiNET SECURITY
- Authentication
- Authorization
- Access Control
- Confidentiality
- Non-Reputation
- Auditing
- Key Management
- Identity Management
- Trust Authorization Management
- Network Based Components
- Physical Component

http://www.naspi.org/
PMU Deployment Today

Currently 200+ PMUs Installed. Expected to exceed 800+ PMUs by 2013 (under SGIG Investments)

Currently 137 PMUs Installed

34 Gigabytes of data collected Daily from 100 PMUs (~ 1 Terabyte per Month).
PMU System Security

- Cyber-security is one of the main obstacles to widespread deployment of PMUs
- Availability & Confidentiality attacks are secondary
- Integrity attacks are most critical
  - Can initiate inappropriate generator scheduling
  - Can result in voltage collapse, and subsequent cascading failures

- Our initial approach
  Consistency checking between cyber network [PMU data received] and physical network [load flow equations] using static state estimation tools
Taxonomy of cyber attacks

Potential Attack points:
Sensors, Phasor Data Concentrator (PDC), comm infrastructure (NASPInet)

Related Projects

- **TCIP: Trustworthy Cyber Infrastructure for the Power Grid**
  [http://www.iti.illinois.edu/content/tcip-trustworthy-cyber-infrastructure-power-grid](http://www.iti.illinois.edu/content/tcip-trustworthy-cyber-infrastructure-power-grid)

- Roadmap to Secure Control Systems, [http://www.controlsystemsroadmap.net](http://www.controlsystemsroadmap.net)


- Smart Grid Recovery Act, [https://www.arrasmartgridcyber.net](https://www.arrasmartgridcyber.net)

---

Our approach and broader objective:

to bring the physics of load flow to cyber-security methods
Outline

- Background
- Power Systems Background
- Phase Measurement Units
- State Estimation & PMU Data
- Our Approach to Integrity Attack Detection
Static State Estimation with PMU Data

- **Recall: What is static state estimation?**
  
  Find the phase angles given:
  
  - measured real power $P$ and reactive power $Q$ at load buses
  - measured real power $P$ and voltage $V$ at generator buses

- **Ubiquitous placement of PMUs**
  
  - Will eliminate need to do state estimation
  - But this is too expensive
    
    - Must live with PMU data at limited number of buses

- **Recent results**
  
  - incorporate PMU data
  - retain standard-form static estimation
    
    - Phadke et al [2006]
State Estimation Equations

- Coupled algebraic nonlinear equations

**Power Flow Constraint:**
\[ I = \mathbb{Y} \mathbb{V} \]

- Bus admittance matrix \( \mathbb{Y} \)
- Injected bus current phasor \( \mathbb{I} \)
- Bus voltage phasor \( \mathbb{V} \)

**Measurement equations:**

- **At load bus:**
  \[ P_k + jQ_k = \mathbb{V}_k \mathbb{P}_k^* + e_k + jf_k \]

- **At generator bus:**
  \[ P_k = \text{Re}\{\mathbb{V}_k \mathbb{P}_k^*\} + e_k \]
  \[ V_k = |\mathbb{V}_k| + f_k \]

- **At PMU bus:**
  \[ y_k = \angle \mathbb{V}_k + g_k \]

**SCADA data:**
\[ P_k, Q_k, V_k \]

**PMU data:**
\[ y_k \]

**IID noises:**
\[ e_k, f_k, g_k \]
State Estimation Problem

- Minimum variance of bus voltage and phase
- Estimate is $\hat{V}$

minimize $E \sum_k \| \hat{V}_k - V_k \|^2$
subject to
load flow equations
measurement equations

exploit: $\sigma_q^2 \ll \sigma_e^2, \sigma_f^2$
“DC load flow”

- For better intuition
  - Assume:
    - Lossless lines: \( Y \approx jB \)
    - Voltage support: \( V \approx 1 \) per-unit
    - Small angles: \( \sin(\delta_k - \delta_l) \approx (\delta_k - \delta_l) \)

- Problem:
  - Estimate power angles \( \delta \) using
    - Real power data [at all buses, noisy, possibly stale]
    - PMU data [at select buses, clean]
“DC load flow” eqns

- Problem becomes weighted least-squares

DC power flow: \[ P = B\delta \]

measurement eqn:
\[
\begin{bmatrix}
R \\
y
\end{bmatrix} = \begin{bmatrix}
P + e \\
C\delta + f
\end{bmatrix} = \begin{bmatrix}
B \\
C
\end{bmatrix} \delta + \begin{bmatrix}
e \\
f
\end{bmatrix}
\]

\( C \) is a permutation matrix: selects buses at which we have PMU data

solution:
\[
\hat{\delta} = [B^*B + \gamma C^*C]^{-1}[B^*R + \gamma C^*y]
\]
\[
\hat{n} = \begin{bmatrix}
\hat{e} \\
\hat{f}
\end{bmatrix} = \Pi \begin{bmatrix}
R \\
y
\end{bmatrix}
\]

where \( \gamma^2 = \frac{\sigma_e^2}{\sigma_f^2} \), \( \Pi \) = standard projection matrix
Outline

- Background
- Power Systems Background
- Phase Measurement Units
- State Estimation & PMU Data
- Our Approach to Integrity Attack Detection
Integrity Attack Detection

- **Basic Idea:** Consistency checking between cyber network [PMU data] and physical network [power flow equations]

- **Assumptions:**
  - PV data at generator buses are known secure
  - PQ data at load buses are known secure
  - at most one compromise in PMU data

- **Comments:**
  - Realistic because of rarity of coordinated attacks
  - Methods can be extended to two or more simultaneous uncoordinated attacks
  - Doesn’t distinguish between faults and attacks
Problem Formulation

- **Given traditional static state estimation data set**
  - PV data at generator buses
  - PQ data at load buses
  - Assumed secure
  - Updated asynchronously at slow time scales [5-15 minutes]

- **Given data from $p$ PMUs**
  - Assume at most one PMU is compromised
  - Updated at fast time scales [60 Hz]

- **Find**
  - Which (if any) PMU data is compromised

- **Solution strategy – Hypothesis testing**
Digression: LS Hypothesis Testing

- **Observation Model**
  
  parameters: \( \delta \in \mathbb{R}^n \)
  
  noisy observations: \( y \in \mathbb{R}^m \)
  
  linear observation model: \( y = A\delta + n \)
  
  i.i.d. noise model \( E[n] = 0, \quad E[nn^*] = \sigma^2 I \)

- **Fault/attack Hypothesis**
  
  \( \mathcal{H}_0 \) all observations are clean
  
  \( \mathcal{H}_k \) observation \( y_k \) is compromised

- Problem: determine most likely hypothesis

- Easy under linear observation model
ML Approach

- For each hypothesis, calculate log-likelihood:

  assume: hypothesis $\mathcal{H}_k$
  
  compute: $J_k = - \min \| n \|^2$
  
  subject to: load flow, observation model

- Choose most-likely hypothesis:

  $k^{\text{ML}} = \arg \max_k J_k$
Problem formulation:

- model: \( y = A\delta + n \)
- noise: \( n \) is i.i.d. with variance \( \sigma^2 \)
- find: which one (if any) observation \( y_k \) is compromised

Theorem:

- define \( N = I - A (A^* A)^{-1} A^* \)
- compute for \( k = 1 : m \)
  \[ \alpha = e_k^* N y, \quad \beta = e_k^* N e_k, \quad J_k = \alpha / \beta \]
- find \( k^o = \arg \max_k J_k \)

then, the ML hypothesis is
\[
\begin{cases} 
\mathcal{H}_{k^o} & \text{if } J_{k^o} \geq \sigma^2 \\
\mathcal{H}_0 & \text{else}
\end{cases}
\]
Application to PMU data

- **Observation model**

  DC load flow: \[ P = B\delta \]

  Measurement eqn:

  \[
  \begin{bmatrix}
  R \\
  y
  \end{bmatrix}
  =
  \begin{bmatrix}
  P + e \\
  C\delta + f
  \end{bmatrix}
  =
  \begin{bmatrix}
  B \\
  C
  \end{bmatrix}\delta +
  \begin{bmatrix}
  e \\
  f
  \end{bmatrix}
  \]

  where \( C \) is a permutation matrix that selects PMU buses

- **Normalization [to make noise i.i.d.]**

  \[
  \begin{bmatrix}
  R \\
  \gamma y
  \end{bmatrix}
  =
  \begin{bmatrix}
  B \\
  \gamma C
  \end{bmatrix}\delta +
  \begin{bmatrix}
  e \\
  \gamma f
  \end{bmatrix}
  =
  A\delta + n
  \]

  where \( \gamma^2 = \frac{\sigma_e^2}{\sigma_f^2} \)
PMU Integrity Attack Detection Algorithm

\[ n \quad \# \text{ of buses} \quad R \quad \text{measured real powers} \]
\[ p \quad \# \text{ of PMU} \quad y \quad \text{PMU data} \]
\[ \sigma_e^2 \quad \text{standard bus noise covariance} \quad e_k \quad k^{th} \text{ unit vector} \]
\[ \sigma_f^2 \quad \text{PMU noise covariance} \quad B \quad \text{bus suscceptance matrix} \]
\[ \gamma \quad \frac{\sigma_e}{\sigma_f} \quad C \quad \text{matrix that selects PMU buses} \]

1. define
   \[ N = \begin{bmatrix} I_n & 0 \\ 0 & I_p \end{bmatrix} - \begin{bmatrix} \frac{B}{\gamma C} \end{bmatrix} \left( B^*B + \gamma^2 C^*C \right)^{-1} \begin{bmatrix} B^* \\ \gamma C^* \end{bmatrix} \]

2. compute
   for \( k = n + 1 : n + p \)
   \[ \alpha = e_k^*Nz, \quad \beta = e_k^*Ne_k, \quad J_k = \alpha/\beta, \quad z = \begin{bmatrix} R \\ \gamma y \end{bmatrix} \]
   end

3. find
   \( k^o = \text{arg max}_k J_k \)

4. assess
   if \( J_{k^o} \geq \sigma_e^2 \quad \text{PMU} \ k^o \text{ is compromised} \)
   else all PMU data are likely secure
Current work

- Experiments with MATPOWER and PowerWorld to test this detection algorithm.
- DC vs AC
- Integration of PMU and SCADA data
- Optimal PMU allocation in terms of attack detectability
- Other detection algorithms
Extensions

- Exploiting sparsity of bus susceptance matrix
  - Can be done using only matrix-vector products
- Extending from DC power flow to nonlinear power flow
  - This is difficult
- Explicitly accounting for stale bus data
  - Can use bus power variance for this
Open research

- Metrics of attack detectability
- Vigilance
  - How frequently must we conduct attack detection? At what fidelity?
- Distinguishing between faults and malicious attacks
- Security-aware PMU placement
  - Which buses? Maybe in pair?
  - Competing objectives
    - WAMS applications vs. Integrity attack detectability
- Large scale simulation study
Some Open Questions

- How to model attacks
- How to detect these attacks
- Is there any difference from plain fault detection?
- How to distinguish faults from attacks
- How to test detection algorithms
Conclusion

- Cyber security research for PMUs is critical and challenging
- Our approach:
  consistency checking between
cyber network [PMU data] & physical network [power flow] using static state estimation tools
- Questions, comments?

Annarita Giani <agiani@eecs.berkeley.edu>
Kameshwar Poolla <poolla@berkeley.edu>
Pramod Khargonekar <ppk@ece.ufl.edu>
Miles A McQueen <Miles.McQueen@inl.gov>

Thanks