Distributed Agent-Based Intrusion Detection for the Smart Grid

Presenter: Esther M. Amullen

January 19, 2018
Introduction

- The smart-grid can be viewed as a Large-Scale Networked Control System (LSNCS).
- LSNCS components such as controllers, plants, sensors and actuators are connected through communication links.
- Typically the computational and physical infrastructure operate side by side in a highly integrated manner.

- The next generation power system is envisioned to integrate advanced control, communication and computational technology improving resilience, reliability and efficiency.
Motivation

- Control of LSNCS is mostly centralized.
- Challenges associated with centralized management:
  - Computational burden
  - Reliance on telemetered data
  - Sensitivity to failure and modeling errors
  - Dynamic topology, configuration not always known
- Distributed operations, monitoring and control architectures solve some problems associated with centralized management.
- Computational advancements support such distributed algorithms.
- Multi-agent systems and robust control algorithms such as consensus are some desirable distributed paradigms.
  - Consensus algorithms are robust and scalable
  - Agents are autonomous, reactive, sociable and proactive.
- Facilitate distributed intrusion detection and mitigation in a time-bound and computationally efficient manner.
Our approach

- Study the impact of cyber attacks on the power grid control system
  - False data injection attacks (FDIA)
- Adapt well studied control systems algorithms to address cyber related problems.
  - Multi-agent systems
  - State Estimation algorithms
  - Consensus algorithms
- We propose a multi-agent system comprising multiple interacting autonomous agents that can:
  - Breakdown a complex power system into smaller logical partitions
  - Poll RTUs and IEDs for measurement data
  - Process data in parallel
  - Exchange data and state information in a time-bound fashion.
- RTU and IED data collected can be used by agents for state estimation, intrusion detection and resilient control.
- Consensus algorithms can be used by agents to rapidly and interactively share information to coordinate results.
Overview—False data injection attacks

False data injection attacks affect:
- Control commands originating from the control center.
- Measurement data sent to the control center from remote field devices.

Attacks on control commands alter the topology of the power grid.
Attacks on measurement data affect state estimation.
Attack Model

- Adversaries can gain access to control traffic by penetrating the control center’s local area network (LAN).
- Within the substations, IEDs can be penetrated by attackers.
- We assume that the only data that can be trusted is data obtained directly from sensors and actuators within substations.
Proposed approach—Distributed agent-based framework

- Deploying software-based agents at substations.
- We assume there’s some form communication among adjacent substations (Specified under the IEEE substation automation standards).
- Agents leverage this communication infrastructure to interact with adjacent agents and substation IEDs.

![Diagram of Distributed Agent-Based Intrusion Detection for the Smart Grid](image-url)
Software agent architecture

Inputs:
- Data from the RTU and PMUs
- Data from other agents

Outputs:
- State Estimates
- Measurements
- Intrusion Detection results

Algorithm suite (Knowledge base)
- Attack detection
- State estimation
- Consensus
Using MAS to detect FDIA

FDIA against state estimation

- Consider a power network with $n$ substations and $n$ agents each deployed at a substation.
- For substation $i$, the corresponding agent determines the measurement vector $z_i$ and corresponding state $x_i$ from

$$z_i = H_i x_i + e$$  \hspace{1cm} (1)

- For an FDIA vector $a$, to evade detection the attack must satisfy the condition

$$a_i = H_i c_i$$  \hspace{1cm} (2)

- The attack is detected if for any agent $i$ the condition (2) is not satisfied.
- The condition is not satisfied if $a_i \in \text{image}(H_i)$. For a subsystem created around a substation, $H_i$ is sufficiently small.
Using MAS to detect FDIA

FDIA against control commands

- Let $x_i$ be the correct state estimate and $z_i$ be the vector of measurements for subsystem $i$.

$$x_i = (H_i^T R_i H_i)^{-1} H_i^T R_i z_i$$  \hspace{1cm} (3)

- For a command with semantics $s_i$, agents can simulate the impact of $s_i$ by computing

$$\hat{x}_i = (H_i^T R_i H_i)^{-1} H_i^T R_i (z_i + s_i)$$ \hspace{1cm} (4)

- The resulting power flows can then be simulated by computing

$$z_{si} = H_i \hat{x}_i$$  \hspace{1cm} (5)
The Consensus problem

- Agents converge to desired state values using local information and that from neighboring agents.
- Let the undirected graph \( G = (\mathcal{V}, \mathcal{E}) \) represent the multi-agent system where the nodes \( \mathcal{V} = \{1, 2, \ldots, n\} \) represent agents and edges \( \mathcal{E} \subset \mathcal{V} \times \mathcal{V} = (\mathcal{V}, \mathcal{E}) \) represent communication links between agents.

Information Sharing

- Agent \( i \) uses state information from its neighbors to update its state according to the law

\[
\psi_i(k + 1) = - \sum_{j=1}^{n} a_{ij}(\psi_i(k) - \psi_j(k))
\]  

(6)

- The information at each agent asymptotically converges to

\[
\psi_i := \lim_{k \to \infty}(k) = \frac{1}{n} \sum_{j=1}^{n} \psi_i(0)
\]

(7)
Detection Algorithm

Algorithm 1 Distributed FDIA detection at agent

Require: Sampling time $k$, Subsystem $i$, where $i = \{1, \ldots, n\}$,
1: Initialize $k = 0$, $z_i(0)$, $x_i(0)$, $\psi_i(0)$

Ensure: $z_i(0)$, $x_i(0)$, $\psi_i(0)$, $\psi_j(0)$, $A_i$, $H_i$, $\tau_i$

2: for Each iteration $k \geq 0$ do

3: $\psi_i(k + 1) = \psi_i(k) + \sum_{j=1}^{n} a_{ij}(\psi_j(k) - \psi_i(k))$

4: $z_i(k + 1) \leftarrow f(\psi_i(k + 1), z_i(k))$

5: $\hat{x}_i(k + 1) = (H_i^T R_i H_i)^{-1} H_i^T R_i (z_i(k + 1))$

6: $z_{si}(k + 1) = H_i \hat{x}_i$

7: for $z_{si}(k + 1) \geq \tau_i$ do

8: Generate alert

9: end for

10: repeat for $k = k + 1$

11: end for
Experimental evaluation

Attacks against measurement data

- MATPOWER is used to simulated power flow for the IEEE 9, IEEE 14 and IEEE 30 bus systems.
- Attack scenario: 1000 random attack vectors are simulated
- Each agent performs a distributed state estimation with a tighter bound on the threshold of bad data
- For the attack cases simulated, probability for a succesfull FDIA against state estimation was \( \leq 0.01 \)
Experimental evaluation on detecting FDIA against commands

- Using the IEEE 118 and IEEE 38 power systems simulated using MATPOWER
- Agents continuously run state estimation and consensus to update neighbors.
- To demonstrate how agents detect malicious commands, we simulate commands that disconnect transmission lines and vary loads and generation
  - 1000 random attacks
  - 1000 targeted attacks
- The agent based architecture successfully detects random and targeted attacks with a success rate of over 96%
Experimental evaluation on detecting FDIA against commands

Random attacks

Targeted attacks
Experimental Evaluation on consensus algorithm

- The consensus algorithm described in (6) enables agents rapidly communicate their results to adjacent neighbors.

\[ \text{Time} = \frac{n_i(3n_b)|\psi_i|}{n_t} = 0.001498 \]  
(8)

\[ \text{Time} = \frac{n_i(3n_b)|\psi_i|}{n_t} = 0.0101952 \]  
(9)
Conclusion

Recap

- Introduced a distributed false data injection attack framework based on multi-agent systems.
- Demonstrated how agents use a limited amount of information to detect attacks and coordinate detection results by a consensus-based rapid information exchange algorithm.

Future Work

- Evaluate the MAS systems in a realistic power grid environment.
Thank you!! Questions??