DRAGON: Deep Reinforcement Learning for Autonomous Grid Operation and Attack Detection

Presented by:
Matthew Landen  
Georgia Tech  

Keywhan Chung  
LLNL  

Moses Ike  
Georgia Tech  

Sarah Mackay  
LLNL  

JP Watson  
LLNL  

Wenke Lee  
Georgia Tech  

Dec. 7, 2022
Cyber Physical Attacks

- Cyberattacks on industrial control systems are on the rise
  - >2x more vulnerabilities published in 2021 as 2020
  - 19% published without a mitigation
- Sophisticated attackers will gain access to power grid control systems
- Systems need to respond to both the cyber and physical effects of an attack

**Goal**

- A system that detects cyberattacks while maintaining reliable power

**Approach**

- Dragon trains reinforcement agents to:
  1. Detect cyberattacks
  2. Maintain reliable power grid operations

Ref: https://www.dragos.com/year-in-review/
Background

### Reinforcement learning
- Technique to learn decision making to accomplish a task, defined as a Partially observable Markov decision process (POMDP)
- Given an observation of the environment $O$, an agent learn a policy $P$ that predicts actions $A$, which maximizes the total reward, $R$
- Agents learn the value of actions, given observations and chose the action with maximum value

### Power Grids
- Structure where power lines connect substations that carry power from generators (green) to loads (yellow)
- Lines carrying more power than their thermal limits are disconnected
Overview of Dragon

Substation → Observation → Control Command → Control Center

- Operator Agent
- SCADA HMI
- Detector Agent
- Alert
Threat Model

- Attacker can disconnect power lines for a maximum of 4 hours
- The line to target is selected randomly, weighted by current power flows
- Duration of attack is either
  - fixed (weighted random attacker)
  - randomly sampled (geometric attacker)
- Attacker can also inject false measurements into grid observations (FDIA)
**Power Operator Agent**

**Observation**
- Load, generator, line attributes

**Actions (179 total)**
- Reconnect power lines
- Bus switching

**Rewards**
- All loads receive power
- Lines obey their thermal limits

**Bus Switch Action**

**Control Command Selection**
1. Reconnect any lines
2. Sample potential commands from estimated action values
3. Simulate commands and rank them
### Observation
- Previous and current grid observations
- Operator’s control command

### Actions
- Detect attack
- Continue normal operation

### Rewards
- Penalize false positives and false negatives
Training Process

Observation Sequence

- Fully Connected
- Long Short-Term Memory
- Fully Connected

Action Values

**Training**

- Agent uses recurrent neural networks as policies
- Agents are trained with the Deep Recurrent Q Network algorithm
- Trained for a set number of backpropagation passes
Evaluation
Evaluation Setup

Evaluation

- IEEE 14 bus power grid with 1,110 load/generation profiles spanning all months and different years
- Ran hyperparameter optimization with 2% of the profiles for validation
- Applied 3-fold cross validation for training/evaluation

Baselines

- Operator
  - Second place team (Oroas) from recent autonomous grid operation competition
- Detector
  - Cumulative sum of residuals algorithm
Operator Evaluation

Dragon vs. Baseline Oroas
Operator Steps Survived

Steps

No Attacker
Weighted Random
Geometric
No FDIA
Weighted Random
FDIA
Geometric
FDIA

Dragon
Oroas
Detector Evaluation

**Attack Dwell Time**

- Cumulative Sum
- Dragon

<table>
<thead>
<tr>
<th>Average Dwell Time (Simulated mins)</th>
<th>Weighted Random</th>
<th>Geometric</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>25</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>

**Detection Rates**

- Weighted Random TPR
- Geometric TPR
- Weighted Random FPR
- Geometric FPR
- Weighted Random FNR
- Geometric FNR
Limitations

- Operator relies on the ability to execute commands to maintain grid reliability
- Small grid used during evaluation

Future Work

- Investigate how the two agents can benefit from each other’s knowledge
Thank you

Q&A
DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.