Detecting Spoofing Attacks in Industrial Control Systems Using Machine Learning

Colman McGuan, Dr. Qin Lin, Dr. Chansu Yu
Department of Electrical Engineering and Computer Science
Cleveland State University

Motivation and Challenges

Motivation

- Industrial Control Systems (ICS) are vulnerable to attacks due to cyber components
- Spoofing attacks can cause damage to ICS
- Very few simulated testbeds available for launching attacks and performing detection

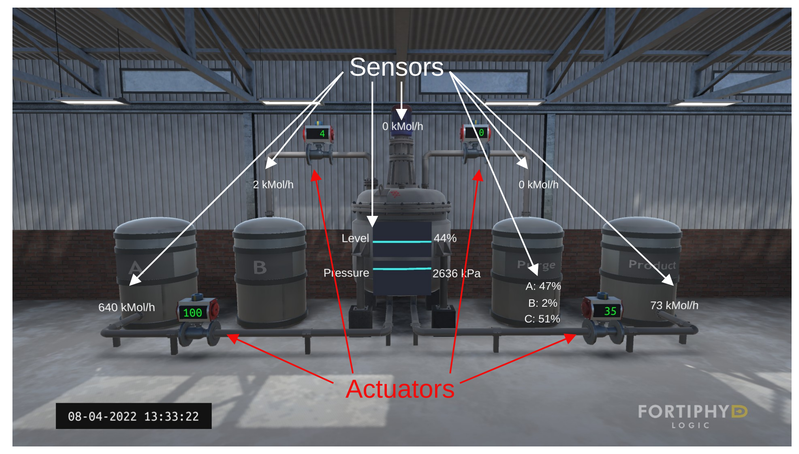
Challenges

- Conventional classification is not appropriate
  - Only benign, unlabled data is available
- Feature selection for training from multiple processes, sensors, and actuators
- Tradeoff between robustness and accuracy

Platform

GRFICS

- Chemical plant simulation created at Georgia Institute of Technology
  - Maximizing product of 3 reactants while minimizing purged excess
- Simulation consists of 5 sub-processes controlled by a single PLC
  - Modbus TCP/IP
- 9 sensors and 4 actuators
- Enables launching man-in-the-middle (MITM) attacks on communication between the PLC and sub-processes
- Rendered using a game engine



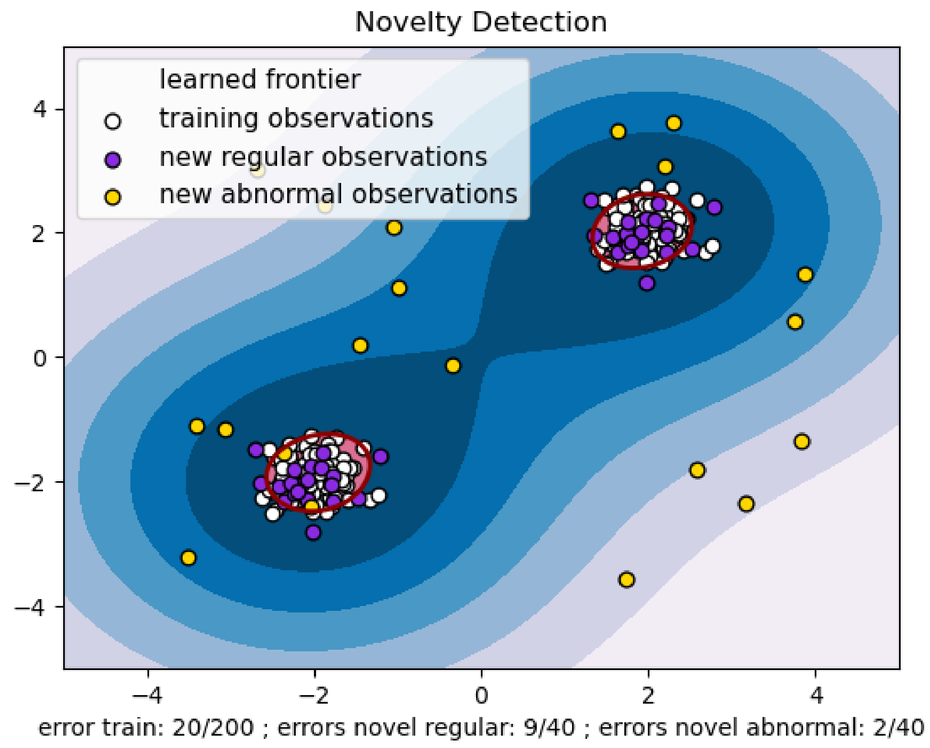
Approach

Attacks

- We create 54 attacks of 5 different categories
  - Single-sensor, multi-sensor, single-actuator, multi-actuator, and complex
- Attacks are strictly limited to manipulating sensor readings from each sub-processes in transit to the PLC and manipulating actuator write commands in transit from the PLC to each sub-process
- A number of the attacks are stealthy attacks – i.e., difficult to detect

Machine Learning Model

- We test several ML algorithms and find One Class Support Vector Machine (SVM) to be the most accurate
- Trained on benign, unlabled data
  - Impossible to know every attack, so we train the model on normal behavior
- Feature set
  - Collected from PLC
  - Sensor and actuator readings
  - Calculated difference between the current sensor reading and the previous
  - The new actuator value calculated by the PLC
- We use a sliding window with a 20 second interval for classification
- Binary classification with a threshold of 60%
  - When the classifier identifies 60% or greater of datapoints in the window as abnormal, the system is determined to be under attack



Results and Future Work

Results

- Our model correctly identifies 47 of 54 attacks (87%)
- Furthermore, 6 of the unidentified attacks are stealthy attacks
- Average detection time of 54.6 seconds
- Median detection time of 18.5 seconds
- We experience no misclassifications of benign data, with over 5 hours worth of testing

Future Work

- In the future, we plan to test our model against different ICS scenarios using the FactoryIO platform
- We plan to run an experiment with a real PLC, allowing us to conduct a hardware-in-the-loop experiment