Cyber Risk Assessment of Traditional and Automated Machine Learning in Autonomous Control Systems

Patience Yockey
Nuclear and Radiological Engineering Program, Georgia Institute of Technology

Abstract

One key goal for advanced reactor design is the implementation of semi or fully autonomous control systems (ACS) to reduce costs associated with operations and maintenance costs. Machine learning (ML) based digital twins (DTs) reduce cost by monitoring plant operations, predicting control actions, as well as diagnosing and providing recommendations for abnormal operational conditions such as malfunctions and component degradation. As advanced reactor vendors develop ACS, techniques to address cyber risk must be developed to ensure safe operation using ML-based DT technologies. The full-scope advanced nuclear cybersecurity (FANCY) testbed has been developed at Georgia Institute of Technology to simulate digital controls within an advanced reactor system. FANCY is a hardware-in-the-loop (HIL) system of the Generic Pressurized Water Reactor (GPWR) simulator coupled with a programmable logic controller (PLC). By connecting an adversarial machine to hardware, real time communications, and ACS, cyber-attacks against advanced reactor systems were simulated.

Traditional and automated ML (AutoML) algorithms were used to develop ACS capabilities using the FANCY testbed. Each type included a plant-level DT to monitor plant-health as well as two device-level DTs to monitor steam generator behavior. The plant-level DT monitored plant-health using autoencoders and 70 variables across subsystems. One device-level DT identified steam generator transients using classification models, while the other device-level DT forecasted the steam generator flow rate using recurrent neural networks.

To test the cyber-resilience of both ACS designs, false data injection against real-time data, SQL injections against training data, and adversarial hyperparameter/weight tuning were conducted to determine which architecture imposes less cyber-risk. Cyber-risk assessment results indicate that AutoML algorithms were more resilient to false data injections, but vulnerable to SQL injections. On the other hand, traditional ML algorithms were more vulnerable to false data injections but resilient to SQL injections. In conclusion, the inherent cyber-risk of both architectures was equal and can be mitigated through proper safeguards and controls.