Enhancing Cybersecurity of Nuclear Systems using Machine Learning/Artificial Intelligence

- Dr. Fan Zhang, Assistant Professor, Georgia Tech, fan@gatech.edu

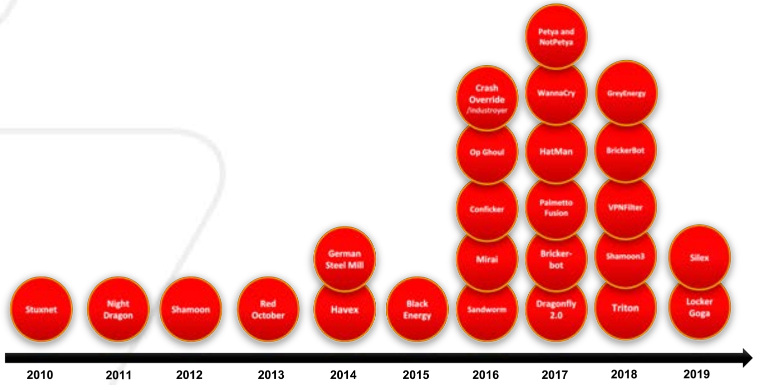
iFAN Lab

Intelligence for Advanced Nuclear

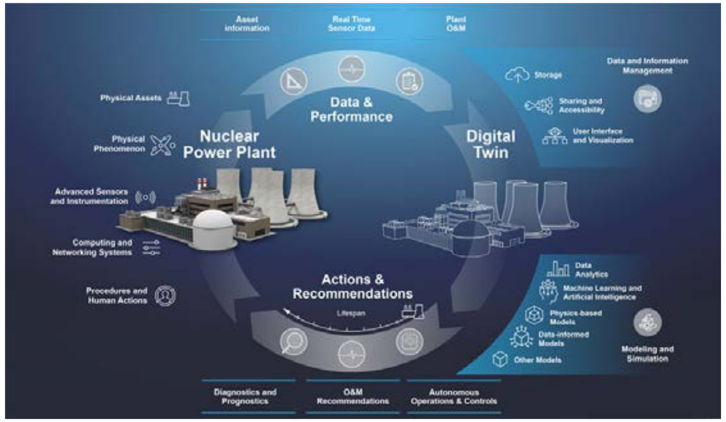
Georgia Tech

Cybersecurity Challenges Posed by Digital Transition and AI Technologies

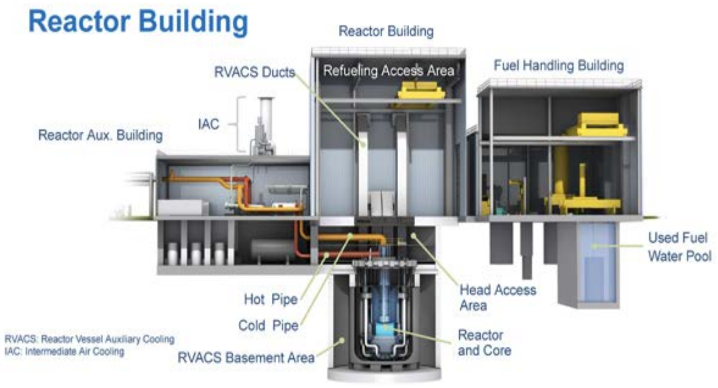
Cyberattacks – growing in number and sophistication



Digital instrumentation and control (I&C) systems

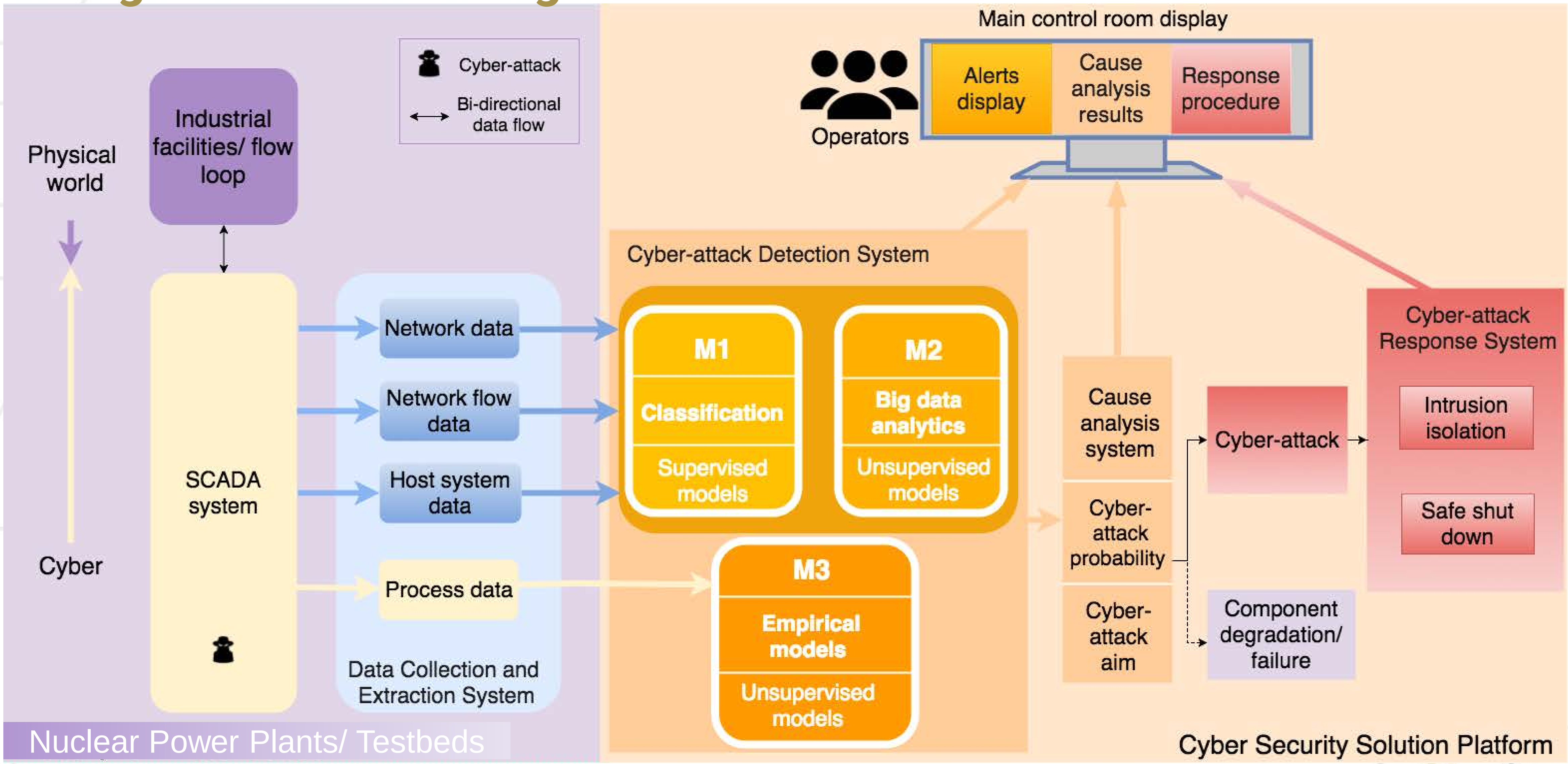


Advanced reactors



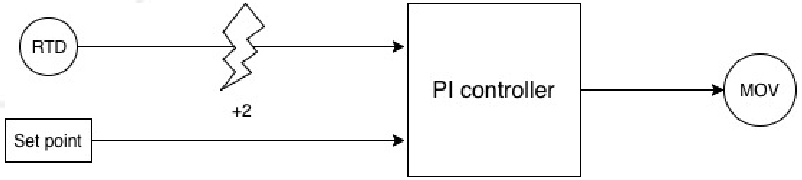
https://www.nrc.gov/reactors/power/digital-twins.html
https://www.nrc.gov/reactors/new-reactors/advanced/licensing-activities/pre-application-activities/natrium.html

Multi-layer Cyber-attack Detection System Using Machine Learning



DOE, Office of Nuclear Energy funded Research

Machine Learning Provides Additional Detection Layer



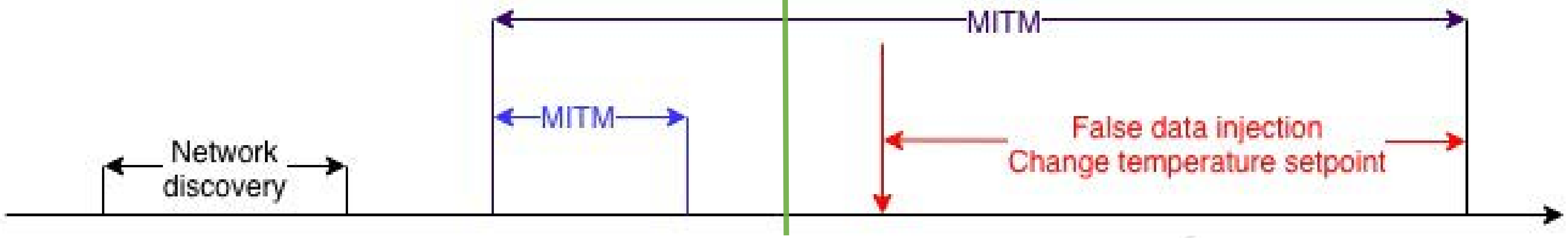
Start time (Obs Index)	End time (Obs Index)	Attack description
600s (150)	630s (158)	Network discovery
840s (210)	1020s (255)	MITM by Ettercap
1020s (255)	1020s (255)	Malicious code injection
1200s (300)	2400s (600)	LabVIEW model run with the malicious code

Cyber data-based IDSs may detect the attacks

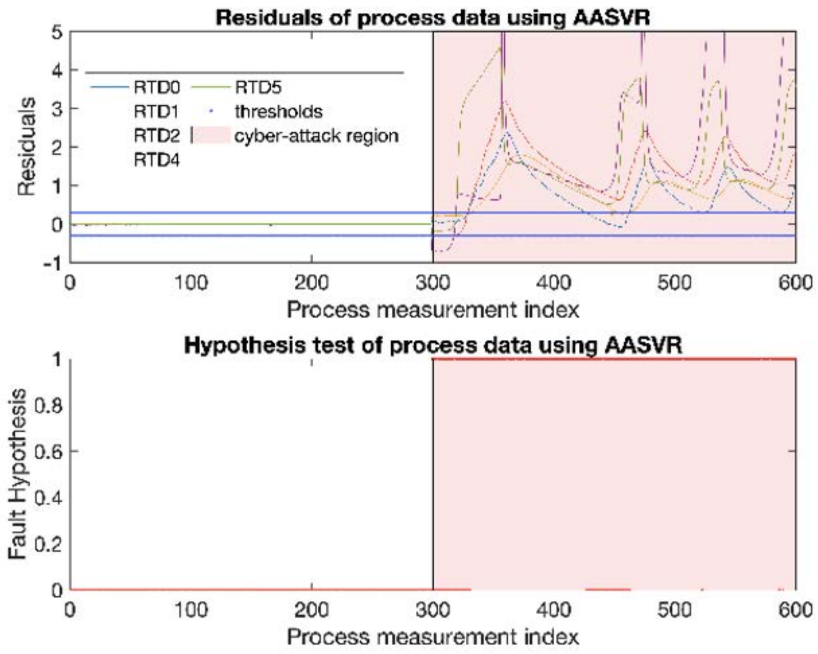
Malicious IPs can be removed

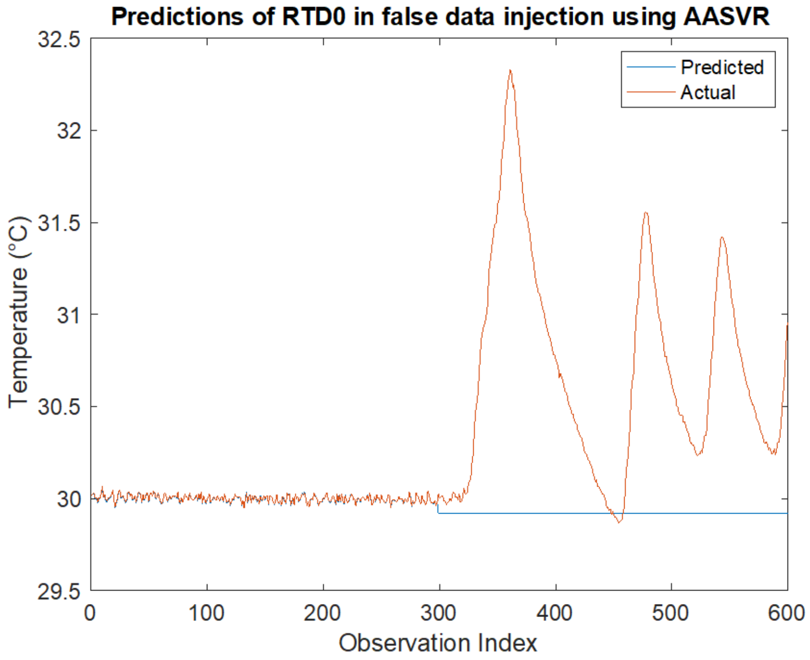
Only process data can indicate process changes

Safe?



Machine Learning Model Detection Results





- Auto-Associative Support Vector Regression (AASVR)
- Observation 301, the malicious code is executed
- Short time to detection
- High true positive

Sensitivity measures how well a model is able to make correct predictions of the variables when the faulty variables are included in the input of the model.

Explainability and Trustworthiness

Explainability

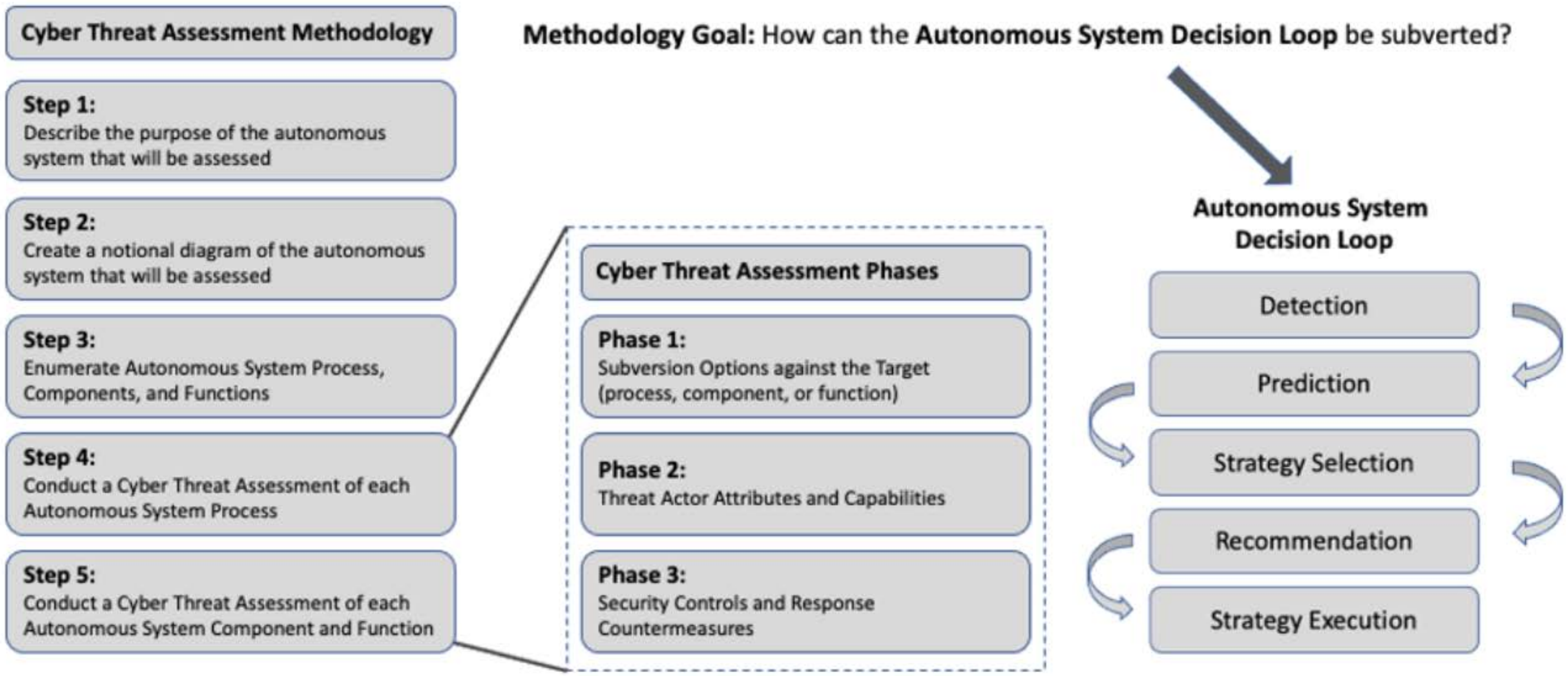
- Machine learning (ML) models can be explainable
- ML-based detection and decisions presented with **evidence** to support decision
- Evidence for detection of new zero-day exploits

Trustworthiness

- **Confidence** in ML-based detection and decisions
- Real-time **decision reliability** assessment
- **Verification and validation (V&V)** in realistic scenarios
- Continual V&V for new and zero-day exploits

Georgia Tech

Cybersecurity of Autonomous Systems



Full-scope Advanced Nuclear CYbersecurity (FANCY) Hardware-in-the-loop testbed

DOE, Office of Nuclear Energy funded Research

Report CT-22IN110402
https://gain.inl.gov/NCRCybersecurityByDesign/Document_INL-RPT-22-68871.pdf

Georgia Tech

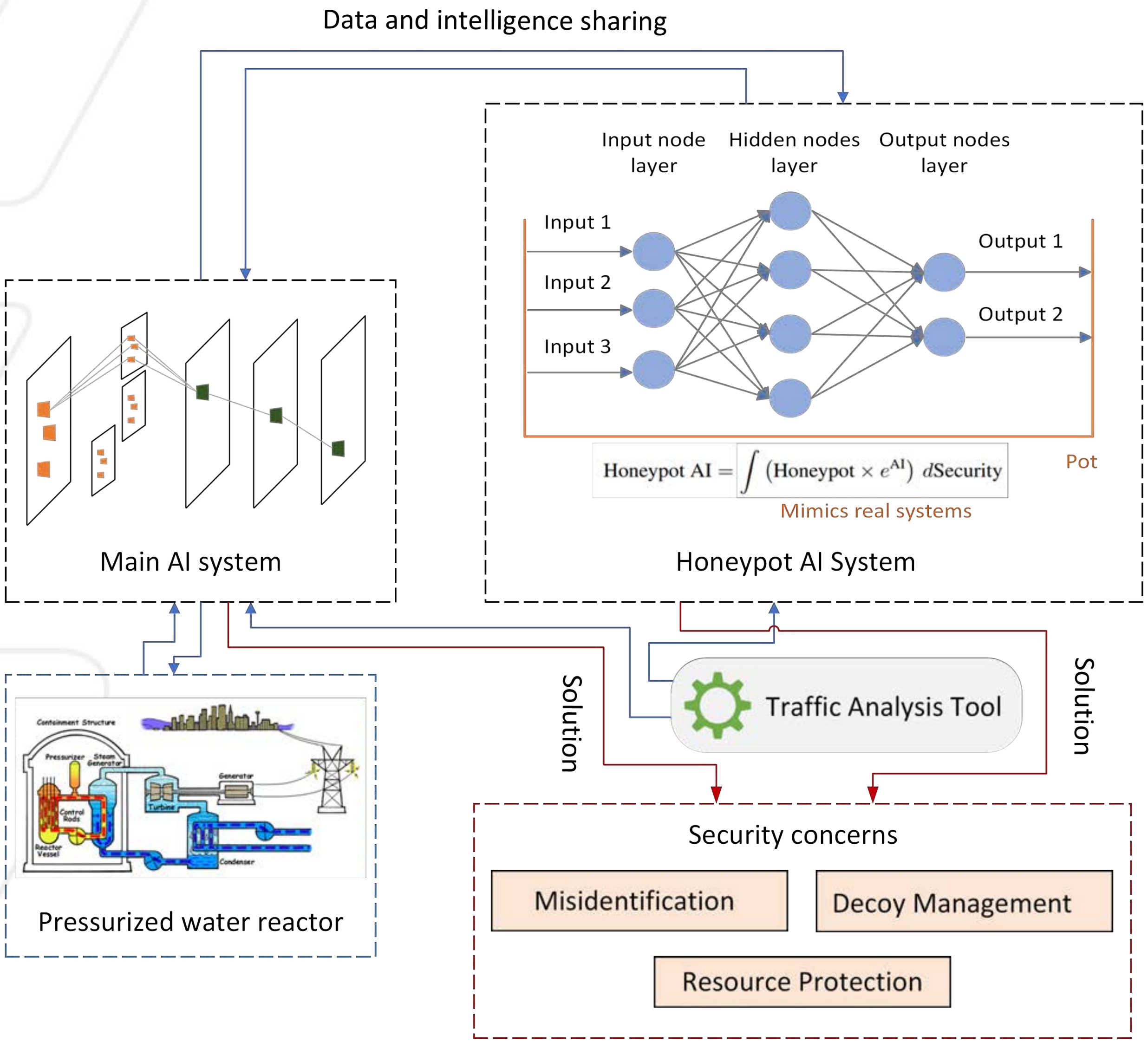
AI/ML – A Double-edged Sword

- AI/ML gives us the ability to carry out complex actions and activities very quickly – faster than was previously possible
- We can achieve this automation faster than ever before – and in a more data-driven way
- Tedious human effort can be kept to a minimum – improving overall performance from a human factor perspective
- Automating away processes can leave us open to new kinds of attacks and vulnerabilities
- AI/ML can introduce new security concerns
- We need strong failsafe(s) in case AI/ML automation fails - and the workforce needs to be prepared to use these

Bad Actors Are Using AI/ML, Shouldn’t We?

- AI/ML technologies are being developed so rapidly that it’s impossible to put a “fence” around them
- Bad actors using AI/ML are not just learning how to use these technologies – they’re learning how to exploit them.
- If we don’t keep pace, bad actors will be 10 steps ahead of us by the time we decide we want to
- If defenders try to stay away from AI/ML, we risk not being on the same playing field as bad actors using these technologies
- Even amateurs are using AI/ML to conduct attacks – and advanced attackers have even more powerful capabilities
- We need to embrace AI/ML to develop best practices and evolve new ways to deal with new attacker capabilities

Potential Solutions with Advanced ML/AI



AI-based Honeypot

- Isolate the honeypot AI from the real control systems
- Monitor for malicious behaviors and attacks
- Continuous training
- Provide data for security improvement

Georgia Tech

Summary

- Constant monitoring: provide fast attack-detection, allowing for a risk-informed regulatory
- High efficiency and effectiveness
- Explainability: many transparent algorithms, allowing for inspection prior to implementation
- Use in an assistant role: no decision or control privileges
- Defense-in-depth: adding another layer of safety and/or security
- Potential detection: ML based security approaches can detect cyber-attacks that have never been seen before
- Easily digestible: once a high-level of confidence is achieved, a broader audience can easily digest risk status information
- Different requirements for different applications
- Embracing AI/ML is needed

Thank you!

Georgia Tech