Enhancing the Cybersecurity of Nuclear Facilities through Data Aggregation and Machine Learning

Dr. Fan Zhang, Assistant Professor

Georgia Tech

fan@gatech.edu

9/12/2024

Georgia Tech

Georgia Tech

The Georgia Institute of Technology is one of the nation's top public research universities with more than 45,000 students who study in person at the main campus in Atlanta, at Georgia Tech-Europe in France, at Georgia Tech-Shenzhen in China, as well as through distance and online learning.

Students represent 50 states and 149 countries.

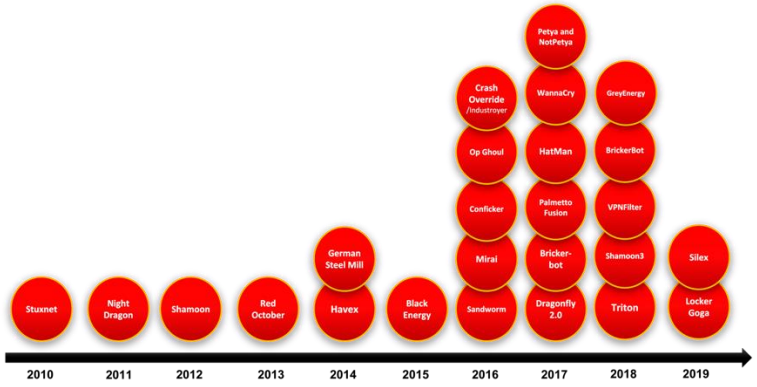
Tech's engineering and computing Colleges are the largest and among the highest-ranked in the nation. The Institute also offers outstanding programs in business, design, liberal arts, and sciences.

With nearly $1.3 billion annually in research awards across all six Colleges and the Georgia Tech Research Institute (GTRI), Georgia Tech is among the nation’s most research-intensive universities. It is an engine of economic development for the state of Georgia, the Southeast, and the nation.

Georgia Tech’s mission is to develop leaders who advance technology and improve the human condition. Its mission and strategic plan are focused on making a positive impact in the lives of people everywhere.

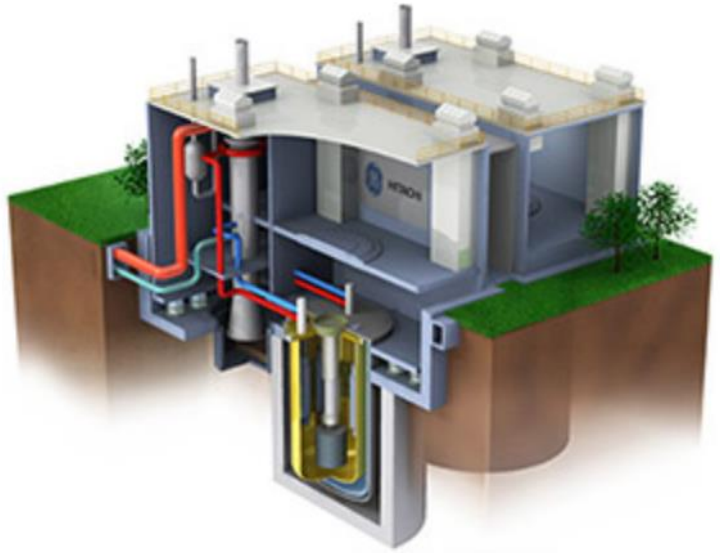
Cybersecurity challenges posed by digital transition and AI technologies

Cyberattacks – growing in number and sophistication



Digital instrumentation and control (I&C) systems

Advanced reactors



* https://www.nuscalepower.com (Photo courtesy of NuScale)
* https://www.power-eng.com/nuclear/doe-announces-230m-to-build-advanced-reactor-demonstration-project/

Network in NPP

Air gap is Not Sufficient:

Advanced reactor:

- Autonomous control
- Remote operation

Current fleet:

- Systems that support removable media
- Insiders

Davis-Besse NPP Slammer Worm

Time: 2003

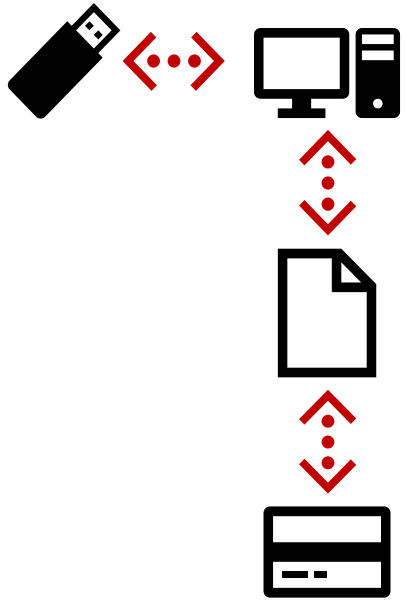
Reason: Slammer worm

Consequence: safety parameter display system inaccessible to operators for about 5 hours [1].

- The consultant created a connection behind the firewall to his office network
- Worm moved from consultant's network to the enterprise network and to the process control network

Stuxnet attacked Iranian nuclear enrichment centrifuges in 2010

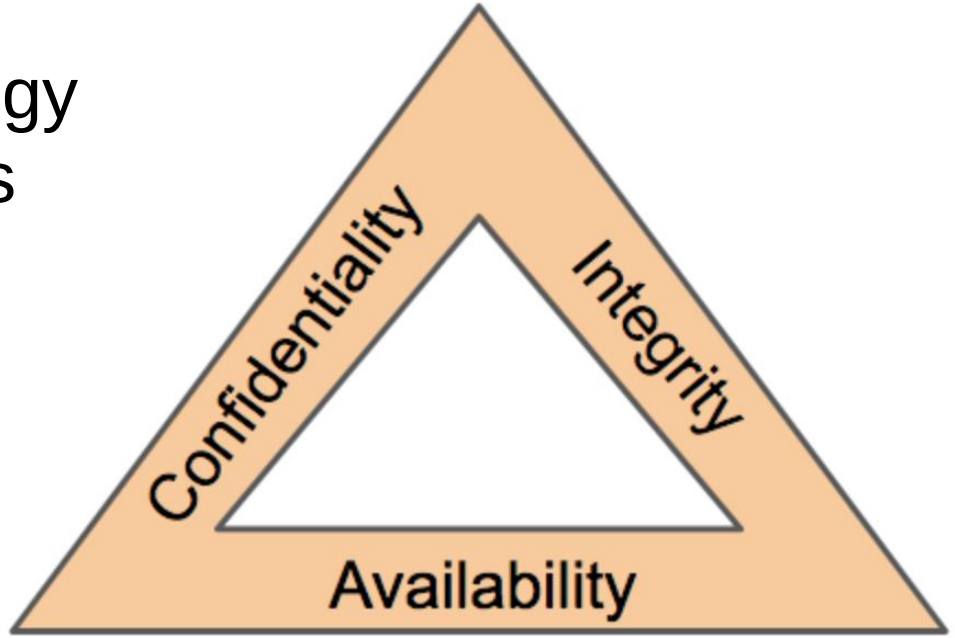
- Destroyed 984 centrifuges
- Malware on a USB drive
  - Propagate over the network
  - Find Siemens Step7 and injects malicious code: instructs PLCs to speed up and slow down the centrifuges
  - Mask the attack from operators by displaying normal data
- Does not require an Internet connection
- Masks malicious activity by faking normal data



Differences between ICS and IT in cybersecurity

Information Technology

- Security objectives
  - CIA
    - Confidentiality
    - Integrity
    - Availability
- High bandwidth
- Easy to update



security objectives

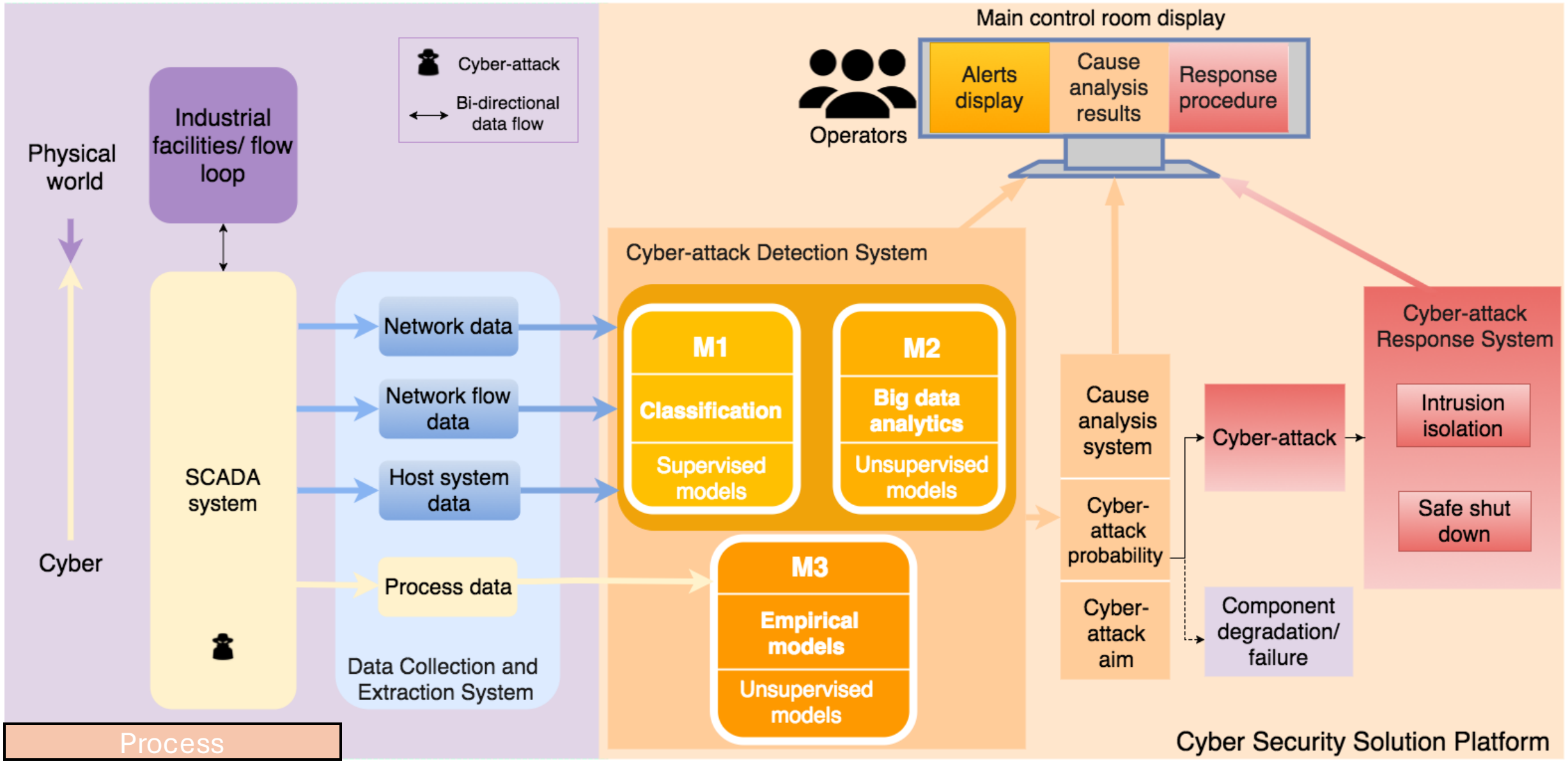
Industrial Control System

- Security objectives
  - AIC
    - Availability
    - Integrity
    - Confidentiality
- Low bandwidth
- Hard to update

What are Industry needs?

Combine IT and Operational technology (OT) team

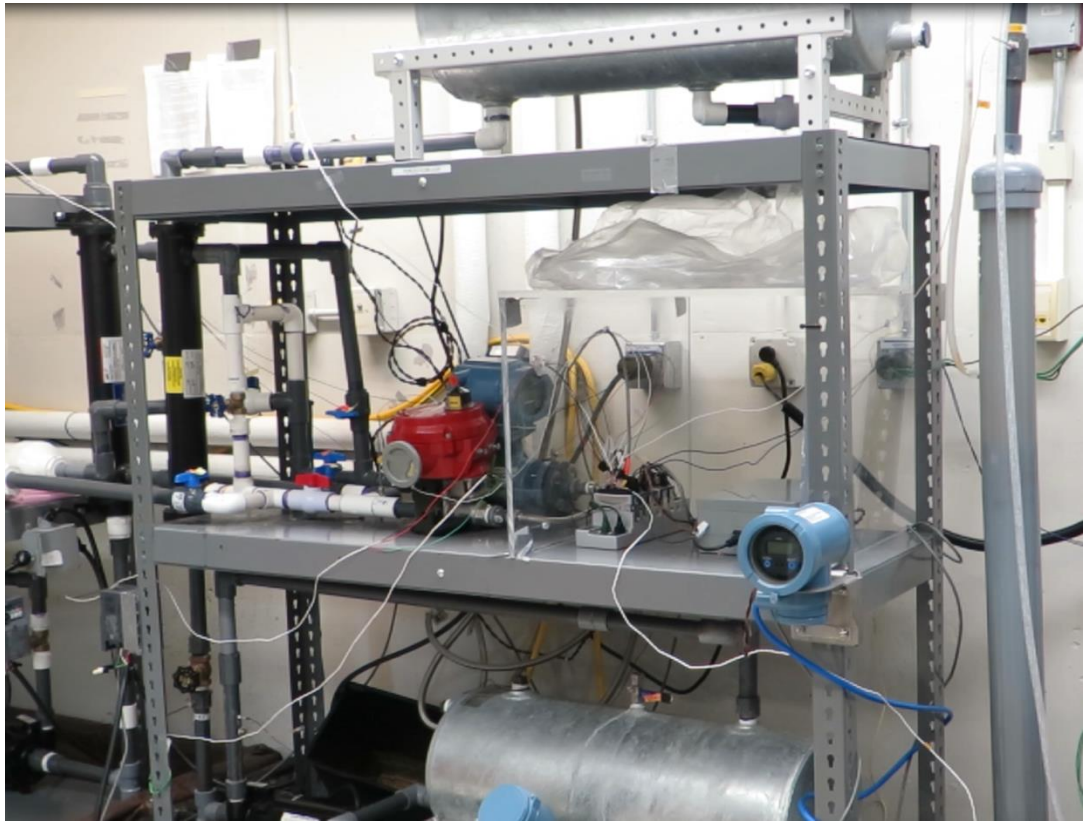
- Network-only based intrusion detection systems (IDSs) may not be sufficient for some intrusions.
- Aggregating **traditional data sources** (used by IT) and **process data** (used by OT) is valuable to improve situational awareness and provide a wide range of attack detection.
- A comprehensive cybersecurity platform to perform multiple functions.



*Zhang, F., Hines, J.W. and Coble, J.B., 2020. A robust cybersecurity solution platform architecture for digital instrumentation and control systems in nuclear power facilities. *Nuclear Technology*, *206*(7), pp.939-950

A real-time cybersecurity testbed with three components

- **A physical experimental facility** which simulates the thermal-hydraulic behavior of a typical two-loop nuclear power system
- A LabVIEW-based **control system**
- **A local area network** (LAN)

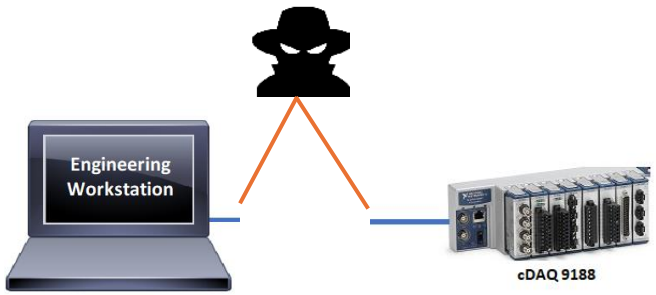


*Zhang, F., Hines, J. W., and Coble, J. (2018). Industrial control system testbed for cybersecurity research with industrial process data. In International Congress on Advances in Nuclear Power Plants (ICAPP 2018).

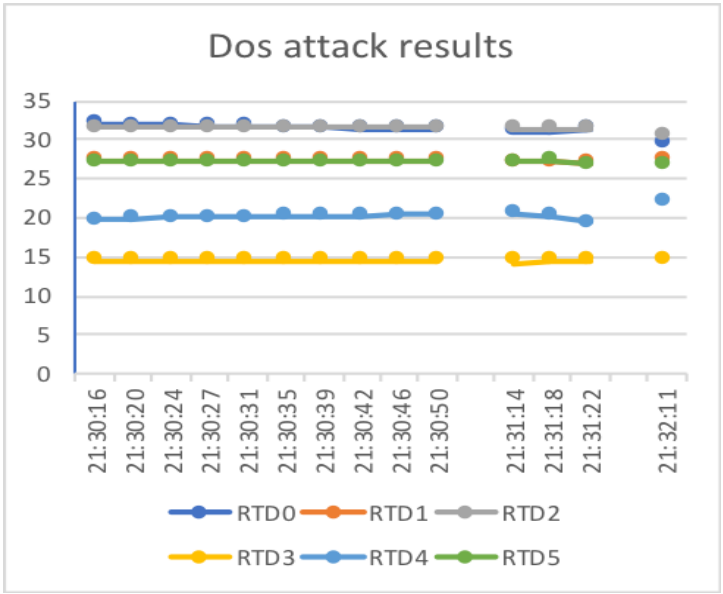
Packet sniffing and DoS attack

Packet sniffing with man-in-the-middle (MITM) attack

```
.....

Tue Mar  6 23:29:27 2018 [100632]
TCP  192.168.1.7:50754 --> 192.168.1.8:31415 | A (0)


Tue Mar  6 23:29:27 2018 [602950]
TCP  192.168.1.7:50754 --> 192.168.1.8:31415 | AP (12)
.........r..

Tue Mar  6 23:29:27 2018 [605454]
TCP  192.168.1.8:31415 --> 192.168.1.7:50754 | AP (12)
.....

Tue Mar  6 23:29:27 2018 [612698]
TCP  192.168.1.7:50754 --> 192.168.1.8:31415 | AP (12)
.........r..

Tue Mar  6 23:29:27 2018 [621359]
TCP  192.168.1.8:31415 --> 192.168.1.7:50754 | AP (12)
```



Denial of service (DoS) by spoofed IP and packet flooding



M1 – supervised models

- K nearest neighbors (KNN)
- Decision Tree
- Bootstrap aggregating (Bagging)
- Random Forest

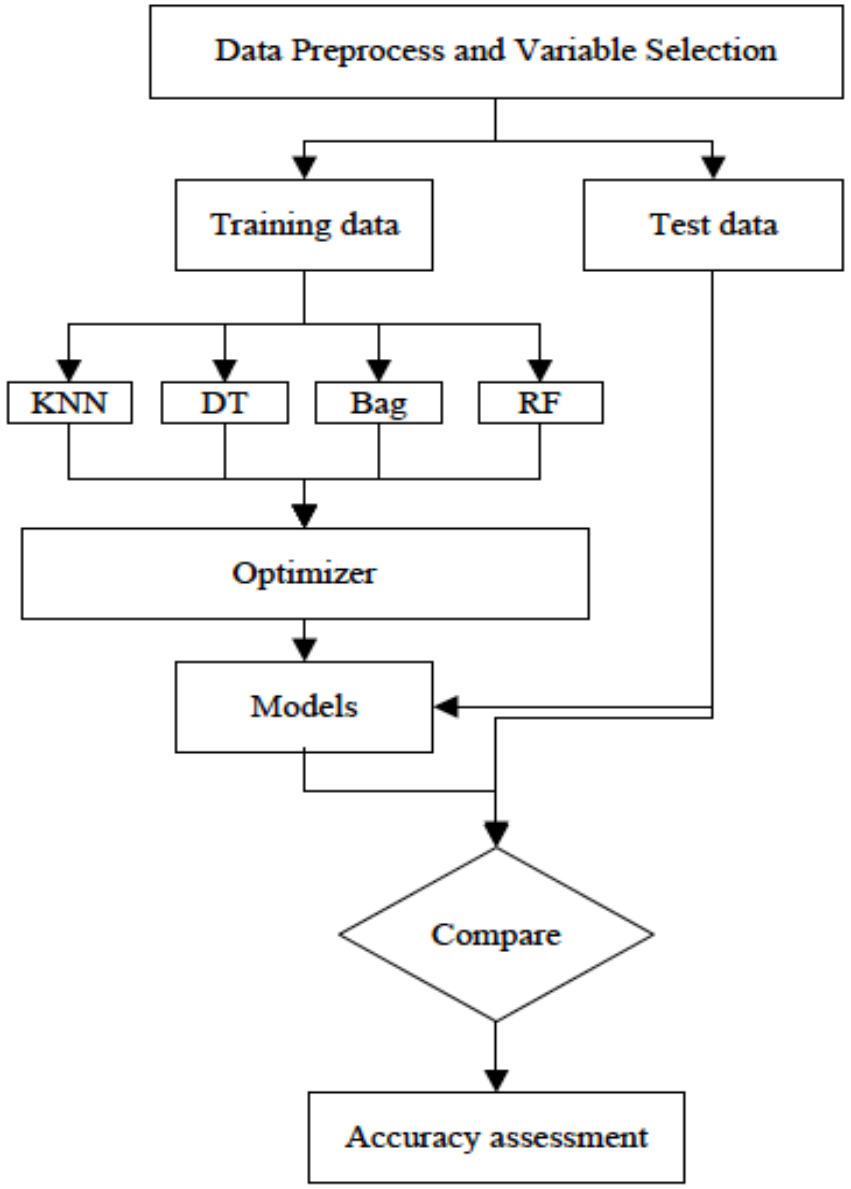
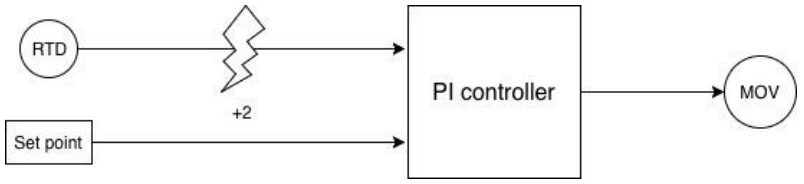


TABLE I
PERFORMANCE COMPARISON OF KNN, DECISION TREE, BAGGING AND RANDOM FOREST

Classification Methods	True positive	False negative	False positive	True negative
KNN	98.84%	1.16%	0.54%	99.46%
Decision Tree	94.8%	5.2%	1.25%	98.75%
Bagging	98.27%	1.73%	0	100%
Random Forest	97.69%	2.31%	0	100%

Process data is crucial in false data injection



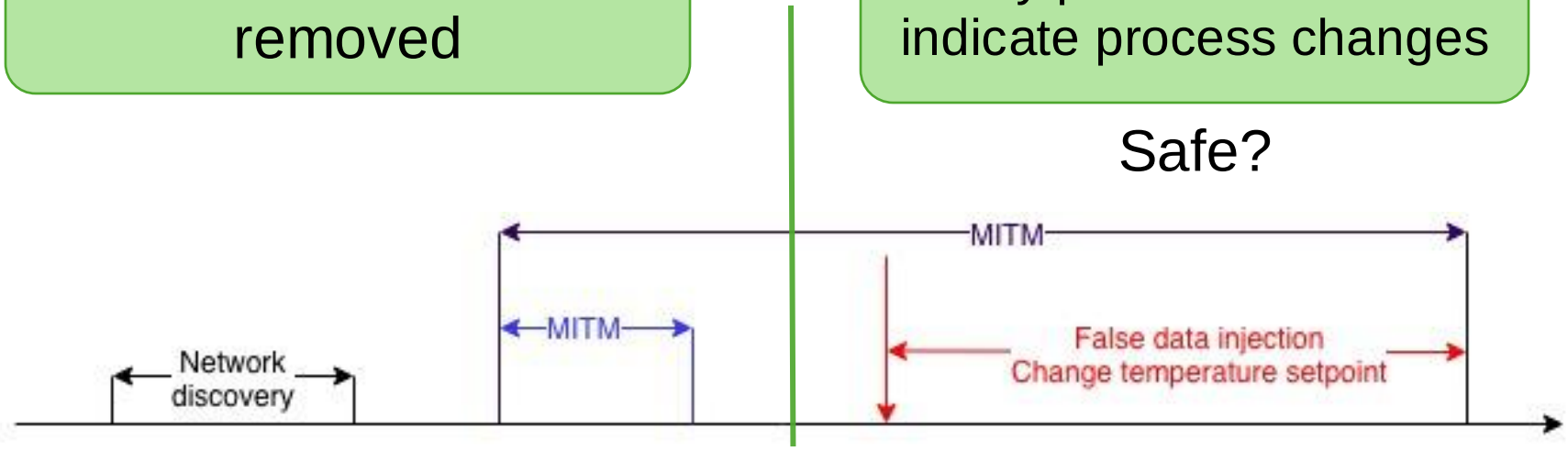
Start time (Obs Index)	End time (Obs Index)	Attack description
600s (150)	630s (158)	Network discovery
840s (210)	1020s (255)	MITM by Ettercap
1020s (255)	1020s (255)	Malicious code injection
1200s (300)	2400s (600)	LabVIEW model run with the malicious code

Cyber data-based IDSs may detect the attacks

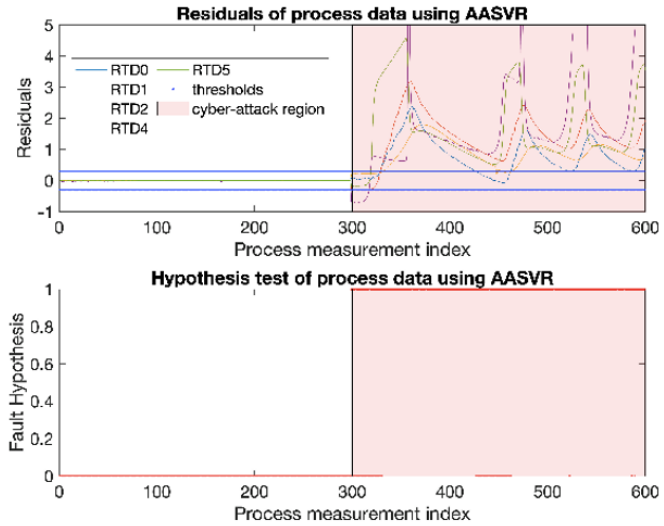
Malicious IPs can be removed

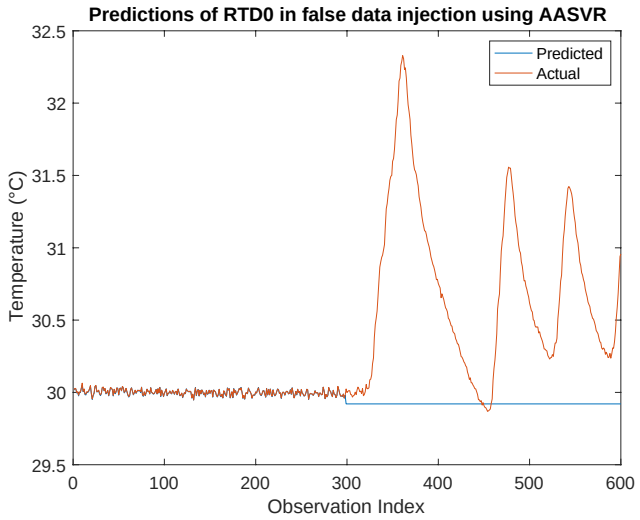
Only process data can indicate process changes

Safe?



M3 AASVR model detection results





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

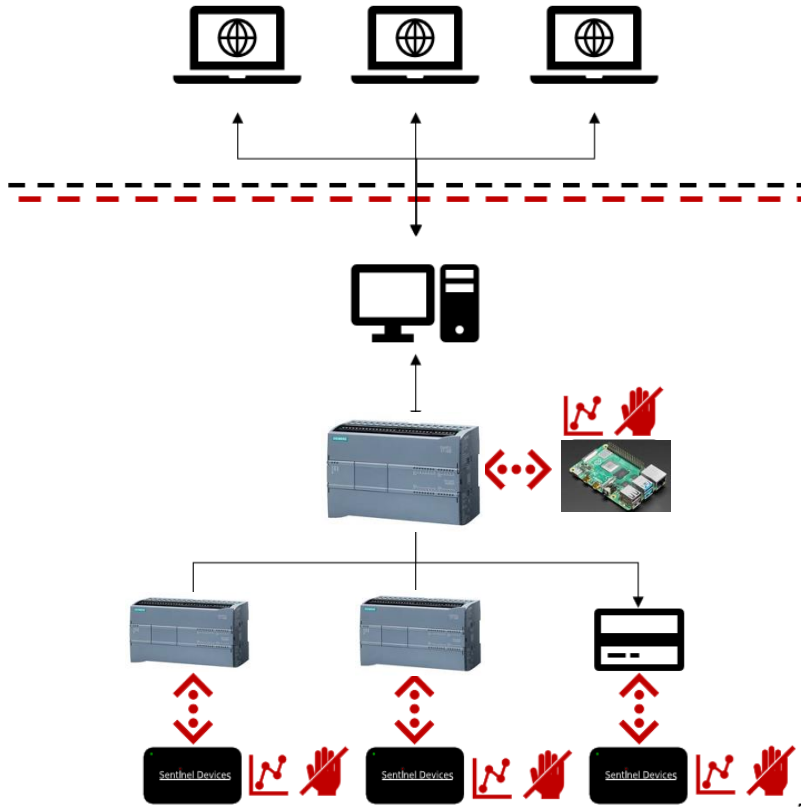
Sensitivity measures how well a model is able to make correct predictions of the variables when the faulty variables are included in the input of the model.

Looks great!

BUT
What about Stuxnet?

Localized strategy for countering Stuxnet-type attack

- Centralized monitoring :
  - Wide range of detection
  - Wide attack surface
  - Subject to data availability
- Localized monitoring for key equipment with embedded digital devices (e.g. PLC)
  - Minimal attack surface
  - Best data availability
  - Closest to ground truth of the sensor
  - Detects Stuxnet type of attack
  - a startup: Sentinel Devices

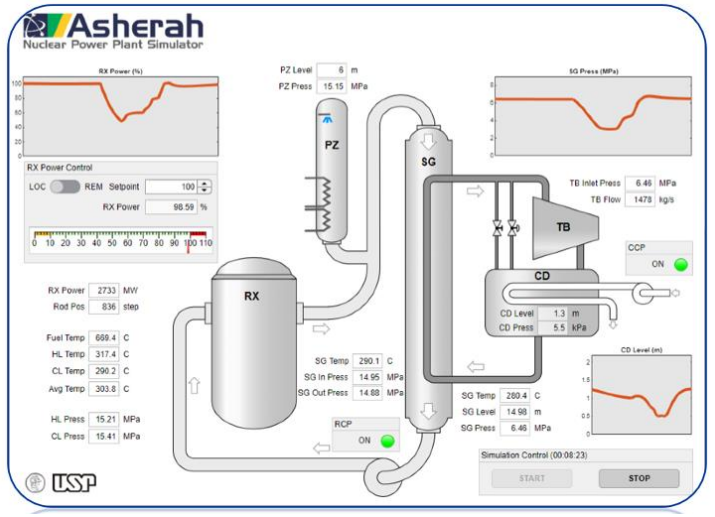


UT RESEARCH FOUNDATION
Technology Transfer & Licensing

HIL Testbed

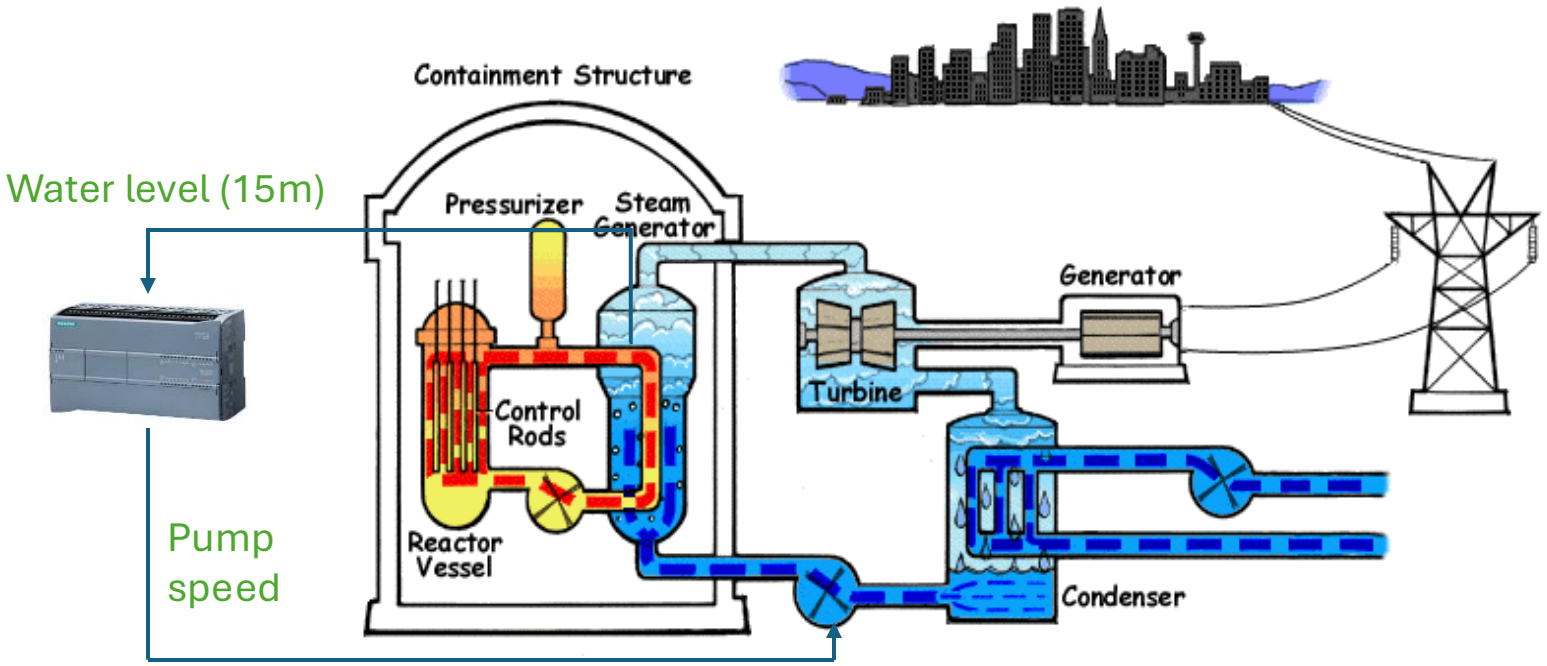
First achieved closed-loop hardware-in-the-loop (HIL) simulation for cybersecurity research in nuclear field

- Asherah simulator
- Programmable logic controller (PLC)
- Steam generator (SG) water control



IAEA International Atomic Energy Agency
Atoms for Peace and Development

20 institutes

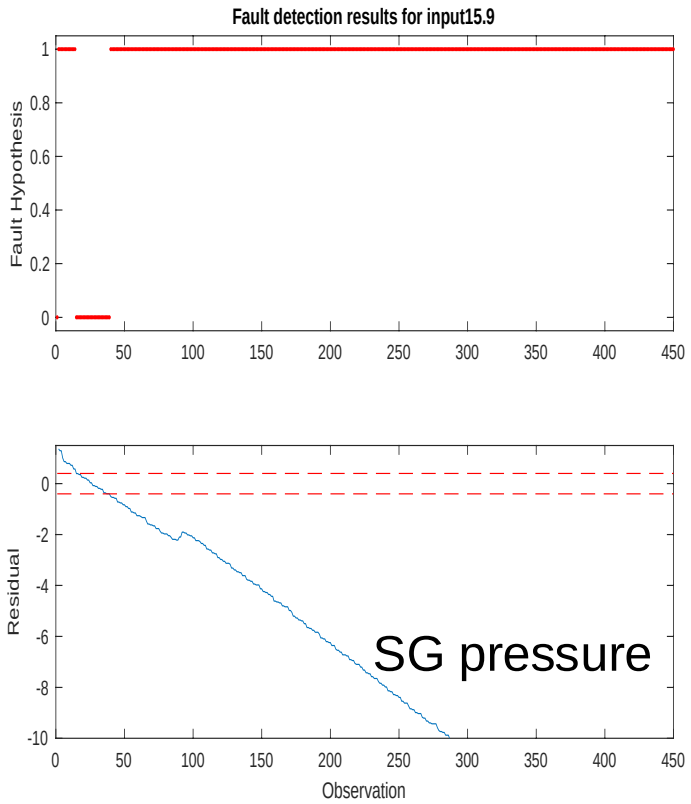


Three false data injection scenarios

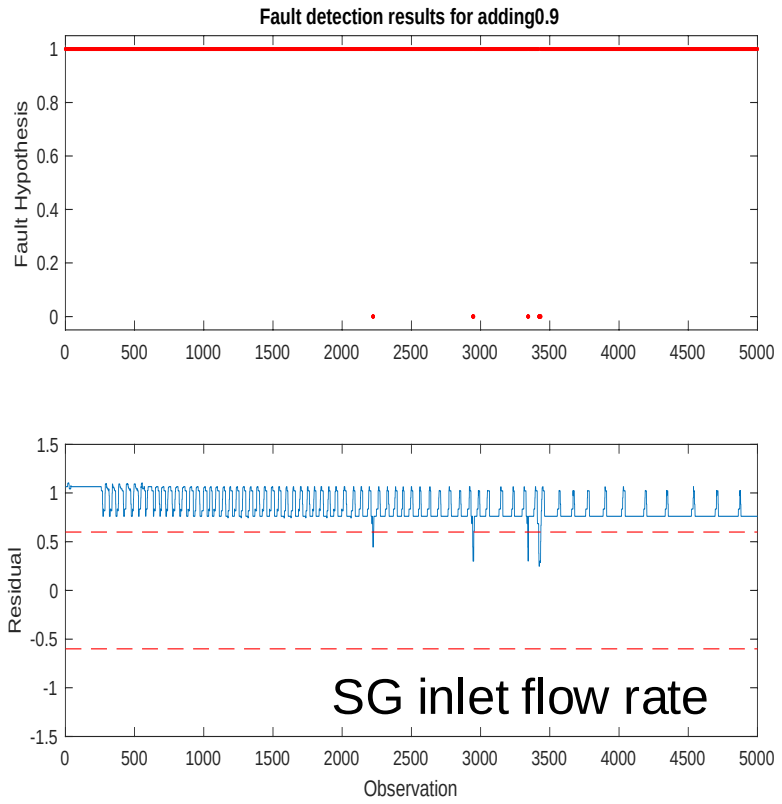
Scenario No.	Attack description	PLC input	PLC action and final SG water level
scenario I	Overwrite the SG level measurement to 15.9m	15.9m	reduces pump speed constantly until SG dry out with 0m
scenario II	Inject a malicious code to add a 0.9m to the SG water level measurement	X+0.9m	reduces pump speed until the SG level reaches 14.1m
scenario III	Simultaneous attack: alter the water level set point to 14m; mask the SG water level with 15m	14m eventually	reduces pump speed until the SG level reaches 14m

* The normal reference water level is 15m

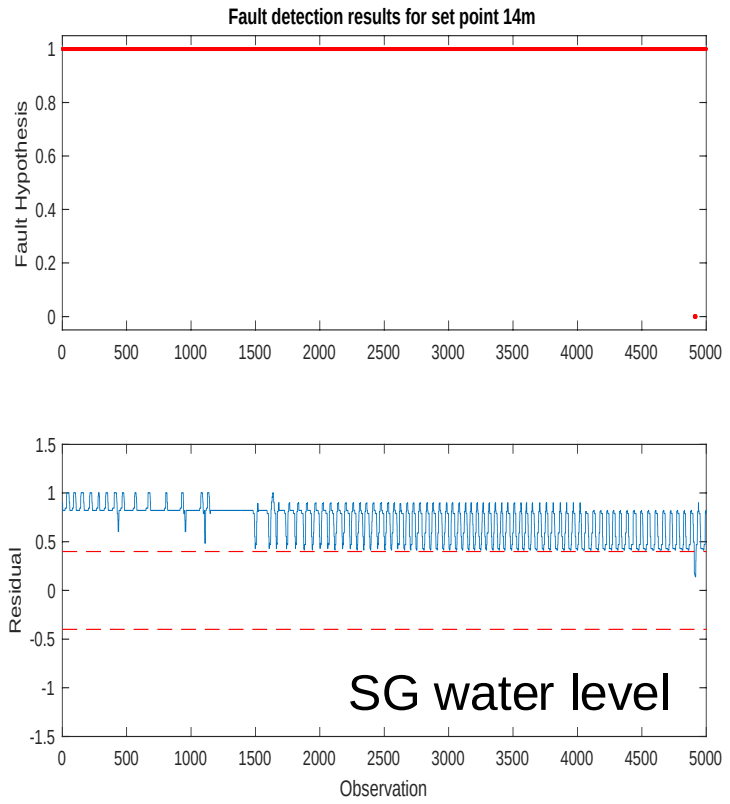
Localized detection results



Scenario I: input 15.9m



Scenario II: adding 0.9m

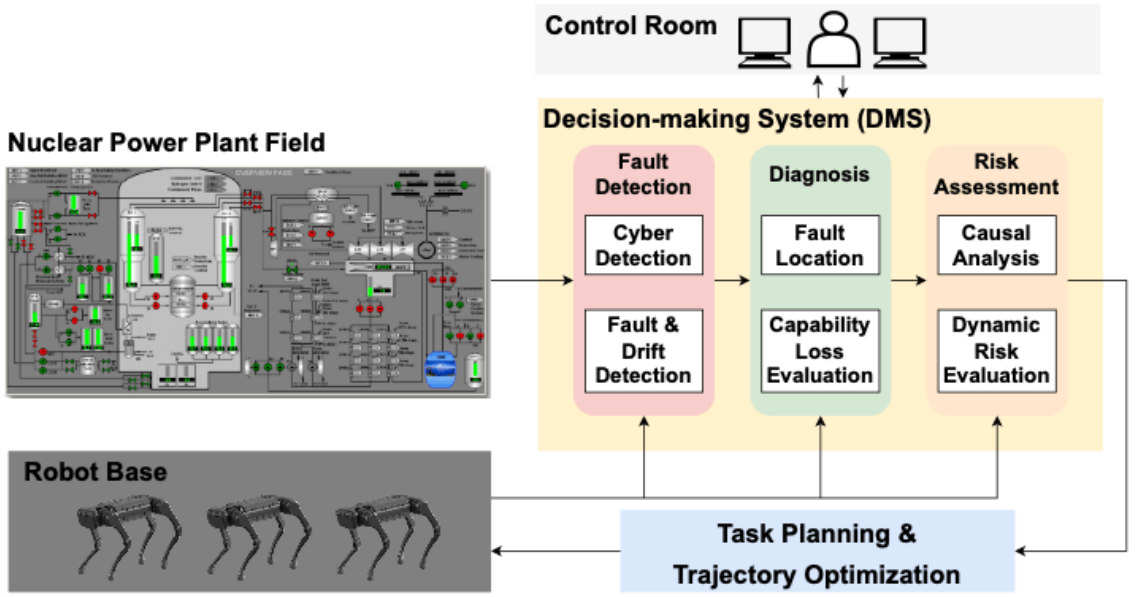


Scenario III: set point 14m

Robot Assisted Online Monitoring, inspection, & Maintenance

Motivation & Goals

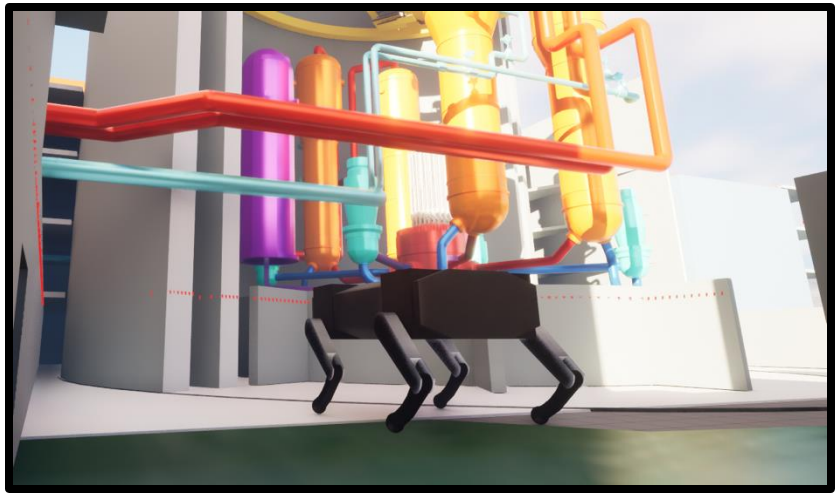
- **Limited sensors**
  Online monitoring system relying on the limited number of installed sensors
  ▶ **Additional on-demand data**
- **Risk of human**
  Dispatching human workers to facility is expensive, and involves risks of injury
- **Slow and inaccurate**
  human-involved process delays decision and lower the accuracy
  ▶ **Fully automated diagnosis**



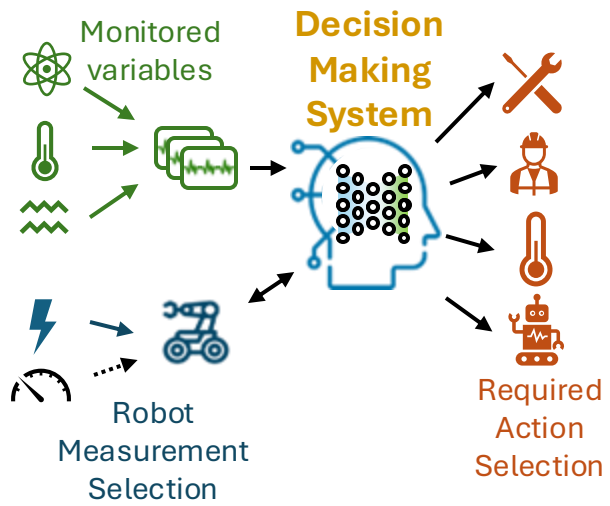
Key Features

- **Diagnosis with variable input sizes**
  Diagnosis will be performed different sets of robot measured input data
- **ML model for non-synchronous data**
  Data from robots will not be synchronously acquired and ML model needs to consider it
- **Feature selection reducing uncertainty**
  It will choose variables to measure which can clarify confused predictions most efficiently
- **Preference weighted**
  Robots will in the order of importance of additional measurements

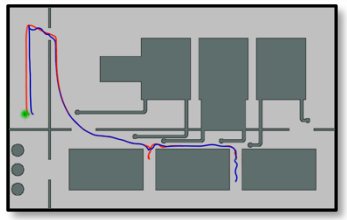
Digital Twin Simulator



Decision Making System



Robot simulation



Preference based task assignments
Cost based Path optimization

Conclusions

A cybersecurity solution platform by data aggregation

- Aggregate cyber and process data
- Perform cybersecurity functions with a defense-in-depth concept
- Wide attack-detection coverage
- Situational awareness for both IT and OT teams

A localized cybersecurity strategy for Stuxnet type of attack

- Small attack surface and good data availability and quality
- A close loop HIL testbed is built
- False data injections
- A low computational cost unsupervised model

Acknowledgement

The work presented was partially supported by Lloyds Register Foundation and the International Joint Research Center for the Safety of Nuclear Energy. Lloyd’s Register Foundation helps to protect life and property by supporting engineering-related education, public engagement, and the application of research. This work was partially supported Oracle Corporation. Some of the work was supported by the U.S. Department of Energy, Office of Nuclear Energy.

Most of the work in this presentation was completed when Dr. Zhang was at University of Tennessee, Knoxville.