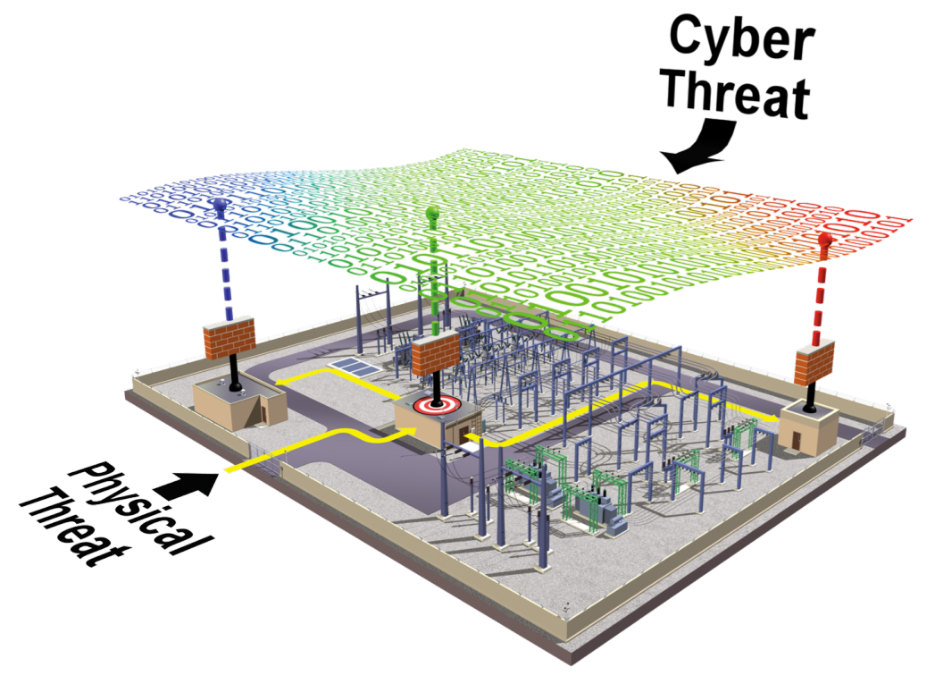
Detecting Anomalies in Industrial Control Systems

USING TIME-SAMPLED SENSOR AND ACTUATOR VALUES

Subin Sapkota, Nuhil Mehdy, Dr. Hoda Mehrpouyan | DEPARTMENT OF COMPUTER SCIENCE, COLLEGE OF ENGINEERING, BOISE STATE UNIVERSITY

BOISE STATE UNIVERSITY

INTRODUCTION



Industrial Control Systems (ICS) rely on sensors and actuators to collect the required information and manage the physical states of the processes. These systems are part of the nation's critical infrastructure, such as nuclear power plants, water treatment plant, etc.

ICS are increasingly vulnerable to cyber-physical attacks, where cyber vulnerabilities are exploited to gain control of ICS and cause catastrophic environmental as well as monetary losses. Therefore, to protect these critical infrastructures that are complex in nature, we required smarter tools and algorithms to monitor and detect new threats. While large number of research studies have been dedicated to Information Technology (IT) domain, few approaches have been solidified for Operational Technology (OT). Hence, this work concentrates on unique security requirements of the OT components i.e. sensors and actuators' data.

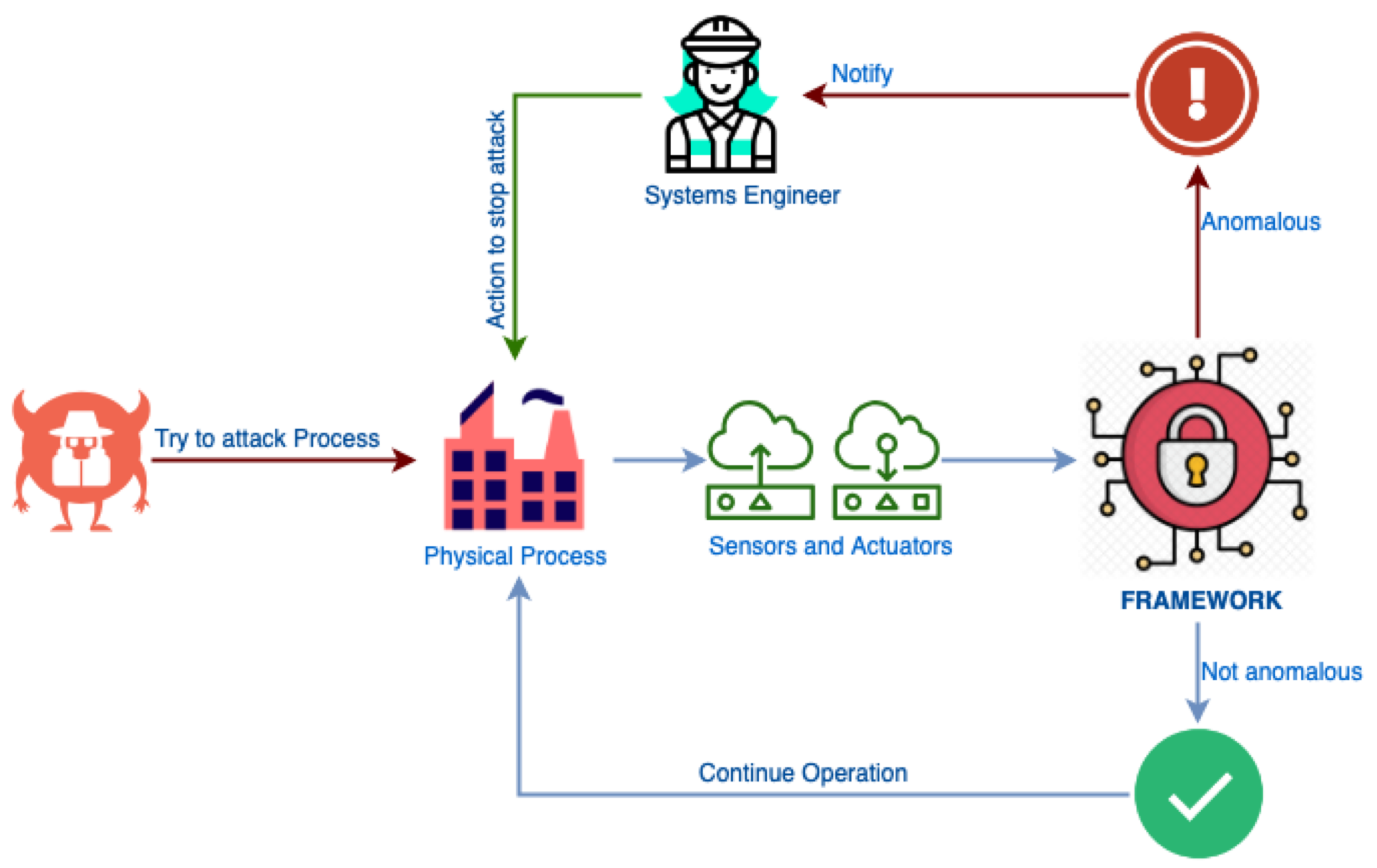
Attack Scenarios in ICS

In this poster, we focus on detecting attacks that affect the sensors and actuators' values as part of the control process. More specifically, we concentrate on attacks that results in the data manipulations of the sensor values, which will translate into a false or delayed actuation. This will force the physical processes to deviate to unsafe conditions. For example, slowly changing the sensor value of a water Level Indicator (LIT) to a value below its lower bound can cause the controller logic to send a command to turn on the pumps. If the wrong control command is not detected fast enough, it will result in tank overflow. Hence, designing and developing tools and mechanism to detect deviation from normal behavior is the focus of this work.

RESEARCH QUESTIONS

- In the domain of ICS, how can we build a forecasting model to predict future behavior of sensor and actuator values based on the historical data and failure behaviors of these components?
- How can this forecasting model be utilized to detect anomalies that results from the attacks, such as data injection and spoofing attacks, on sensors and actuators with high precision in real-time?
- What is the best design pattern for implementing an anomaly detection system in ICS w.r.t the large number of the training data, continuous versus discrete nature of the data, requirements for a quick detection, temporal properties, etc. ?

TOWARDS A RESILIENT AND SECURE CRITICAL INFRASTRUCTURE



RELATED WORKS

Type	Contribution	Weakness
Machine Learning	Artificial Neural Network, as well as various classification algorithms were introduced.	Problem of interpretability of detected anomalies. Scalability is not tested. Number of new attacks caught is not specified, and large false positives were discovered.
Formal Models Specification	System model in modelling language to verify system safety and security properties.	Large amount of manual intervention required to model in formal language. Questions of scalability are not answered.

Anomaly Detection using Machine Learning and Deep Learning:

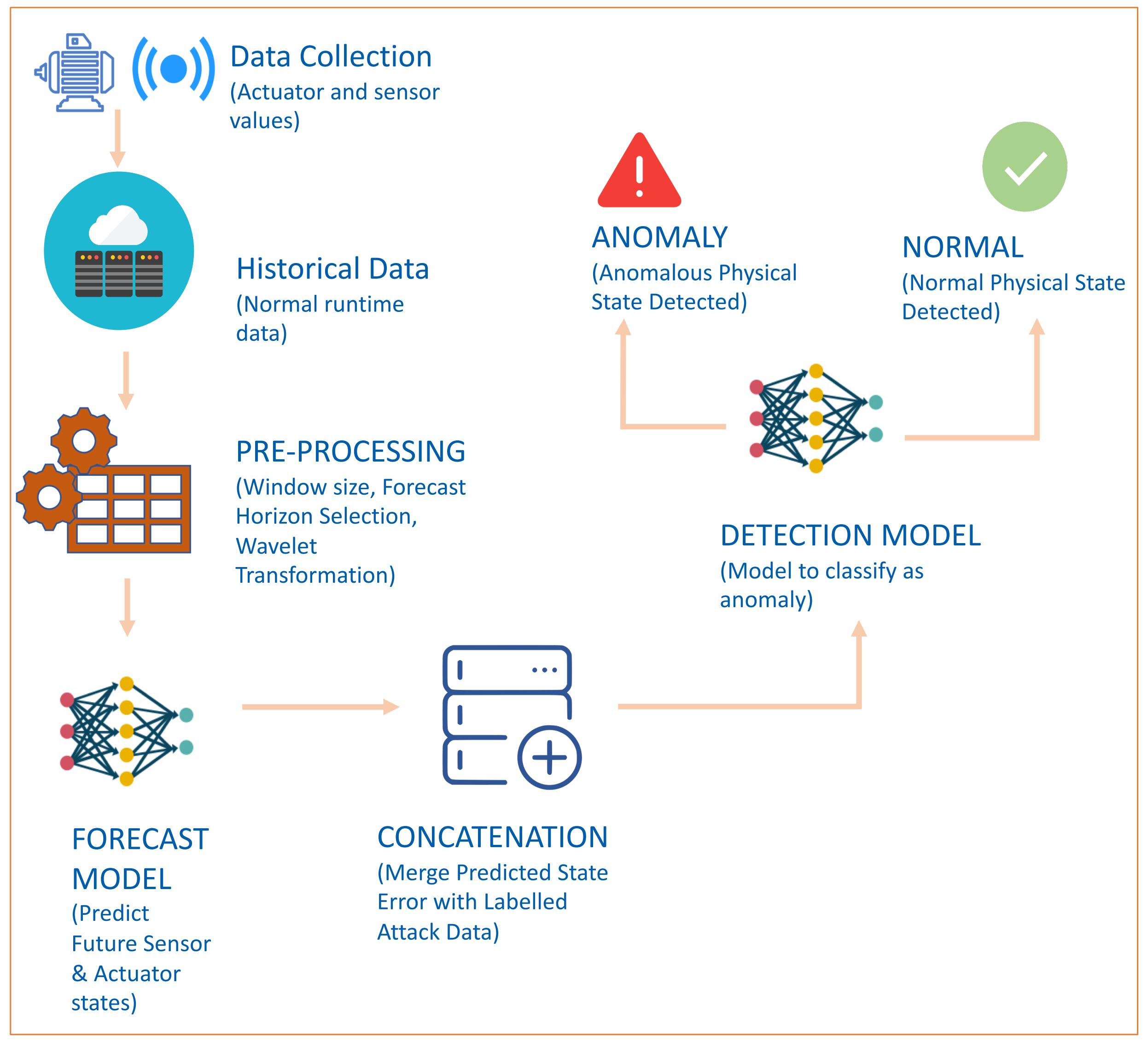
- Junejo, Khurum Nazir, and Jonathan Goh. "Behaviour-based attack detection and classification in cyber physical systems using machine learning." Proceedings of the 2nd ACM International Workshop on Cyber-Physical System Security. ACM, 2016.
- Shalyga, Dmitry & Filonov, Pavel & Lavrentyev, Andrey. (2018). Anomaly Detection for Water Treatment System based on Neural Network with Automatic Architecture Optimization.
- Goh, Jonathan, et al. "Anomaly detection in cyber physical systems using recurrent neural networks." 2017 IEEE 18th International Symposium on High Assurance Systems Engineering (HASE). IEEE, 2017.
- Inoue, Jun, et al. "Anomaly detection for a water treatment system using unsupervised machine learning." 2017 IEEE International Conference on Data Mining Workshops (ICDMW). IEEE, 2017.

Anomaly Detection using Formal Model Specifications

- Kang, Eunsuk, et al. "Model-based security analysis of a water treatment system." Proceedings of the 2nd International Workshop on Software Engineering for Smart Cyber-Physical Systems. ACM, 2016.
- Fauri, Davide, et al. "From system specification to anomaly detection (and back)." Proceedings of the 2017 Workshop on Cyber-Physical Systems Security and PrivaCy. ACM, 2017.
- Fauri, Davide, et al. "From system specification to anomaly detection (and back)." Proceedings of the 2017 Workshop on Cyber-Physical Systems Security and PrivaCy. ACM, 2017.
- Adepu, Sridhar, and Aditya Mathur. "Using process invariants to detect cyber attacks on a water treatment system." IFIP International Conference on ICT Systems Security and Privacy Protection. Springer, Cham, 2016.

FRAMEWORK OVERVIEW

Sensors and actuators data collected during normal system behavior of ICS are used to build a forecasting model to predict future values for sensors and actuators based on previous time window values. Wavelet transformation is applied to extract information from sensor values and add features. Error between predicted values and real values when the predicted time is reached is calculated. An anomaly detection model is trained by adding signals of attack on this error of predicted behavior to classify behaviors as anomalous/normal.


Data Collection
(Actuator and sensor values)
Historical Data
(Normal runtime data)
PRE-PROCESSING
(Window size, Forecast Horizon Selection, Wavelet Transformation)
FORECAST MODEL
(Predict Future Sensor & Actuator states)
CONCATENATION
(Merge Predicted State Error with Labelled Attack Data)
DETECTION MODEL
(Model to classify as anomaly)
ANOMALY
(Anomalous Physical State Detected)
NORMAL
(Normal Physical State Detected)

METHODOLOGY

- Sensor and actuator values are sampled every second during normal system runtime.
- Data is pre-processed to fit selected window-size and forecast horizon.
- Wavelet Transformation is used to decompose sensor signals to add features to each time window.
- Artificial Neural Network Model is built to forecast sensor and actuator values.
- Error of current state in terms of predicted state and actual states obtained.
- Error is concatenated to window-size data and labelled using the attack/normal signal from labelled attack dataset.
- Detection model that classifies a window as normal/anomalous is trained using Artificial Neural Network.

Case Study

Data

- Dataset obtained from a miniature water treatment plant: Secure Water Treatment System (SWaT) testbed.
- 52 sensor and actuator values sampled every second for 7 days normal behavior and 4 days attack data.
- Consists of 6 sub-processes.
- Combination of categorical and continuous values.

Attacks Analysis

- 36 attacks carried out with some form of impact.
- Targeting single points, single points in multiple processes, multiple points, or multiple points in multiple processes.
- Attacks take few minutes to 40+ minutes to make impact.

PRELIMINARY RESULTS

Figures in the right depict the comparative performance of our supervised models and Inoue et al. (Figure a). The proposed CNN+LSTM model failed to detect 9 out of 36 attacks, outperforming previous work that missed 23 attacks. We were able to achieve higher level of detection, while exponentially decreasing the training time of our models (Figure b). Currently, our model is dependent upon the signals provided by the labeled attack data, and future work will be done to explore a semi-supervised level of detection models.

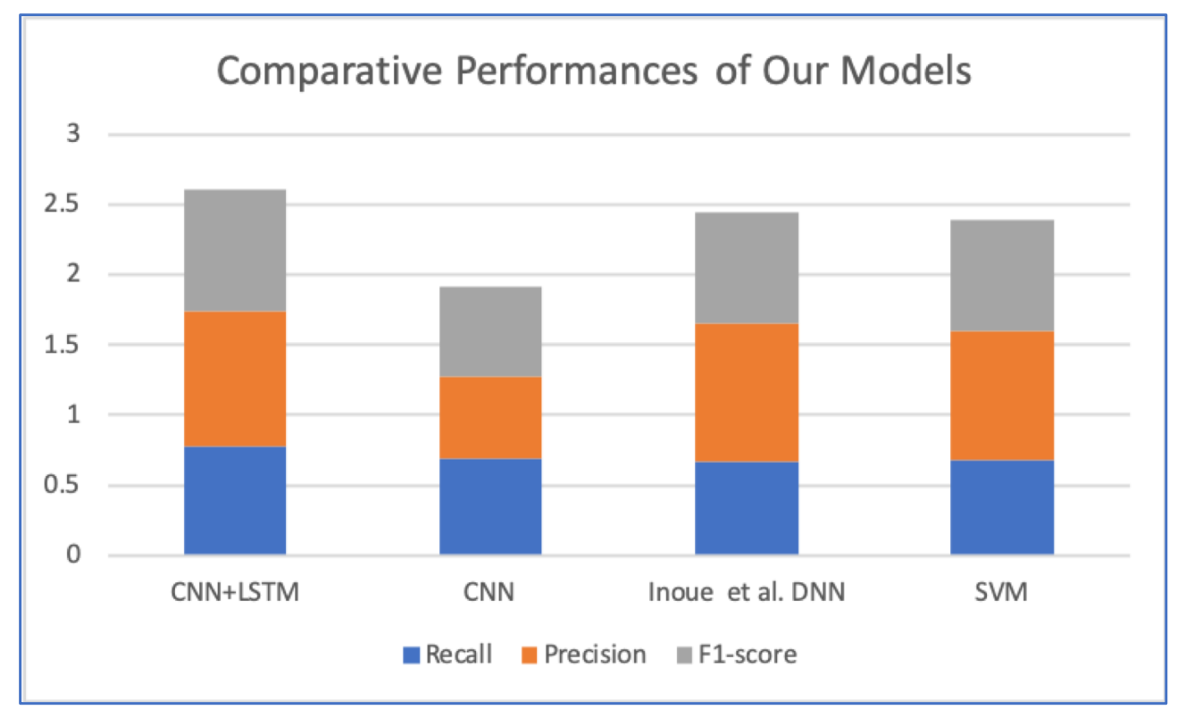


Figure (a)

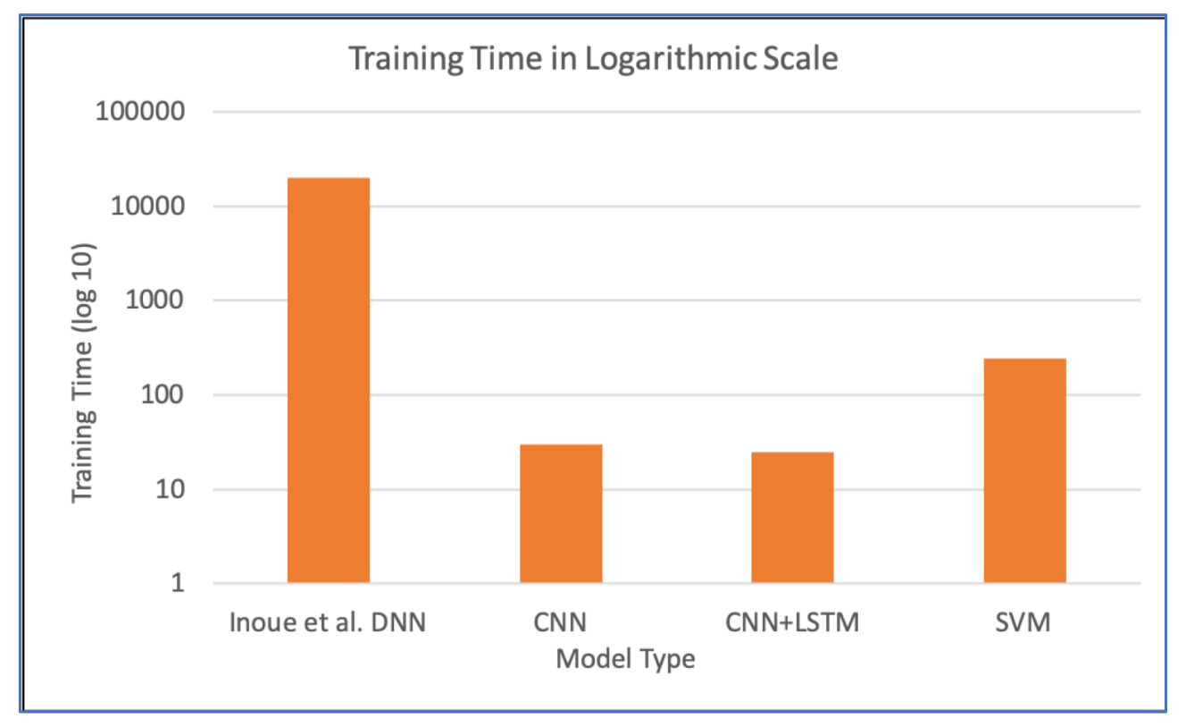


Figure (b)

CONCLUSION

Preliminary results demonstrate supervised anomaly detection methods are promising and could be utilized to provide a new layer of security in ICS. Future works include minimizing false alarms within the system, and building larger labelled dataset of the attack data. In addition, we will quantify detection time, training time, and model size for each variation of anomaly detection architecture and validate the best performing architecture utilizing the proposed methodology in this work.