**Control Logic Based Cyber Attacks in Industrial Control Systems**

**Laurence Browne**

**Supervised by: Dr. Sridhar Adepu**

**Abstract:**

Cyber Physical Systems (CPS) use a cyber supervisory system to monitor sensors and control actuators in order to interact with the physical world. Smart devices and the Internet of Things (IoT) mean CPS are common in everyday life, from home automation to Critical National Infrastructure (CNI). As such, understanding the threats posed to these system by emerging technologies is a key area of research for both the public, private and defence sectors.

This paper illustrates the application of machine learning techniques to passively model the dependencies and relationships within a typical CPS and use these to form the beliefs of an intelligent agent. The intelligent agent probes the system to verify these beliefs then pushes the system beyond it’s normal operational parameters- a cyber-attack.

Data from the Secure Water Treatment Testbed (SWaT) was used as this is a common industrial process and is a good proxy for other CPS as it uses a range of common components such as motors/ pumps, switches, metering and sensors.

A physics-informed intelligent agent was used which could identify components by their characteristics. A Classification Neural Network was trained on normal and attack data, this anomaly detector is a common Intrusion Detection System (IDS) and used to provide feedback to the intelligent agent.

The agent was tested against simulations of the SWaT using traditional and deep learning models.

It can be seen that even basic statistical tools give a reasonable approximation of an unknown systems and would be sufficient to draft simple attacks.

More advanced methods which model temporal relationships were combined to produce a tool which was highly effective in generating simple cyber-attacks and facilitating more complex, human in the loop attacks. This tool is intended to be system agnostic so can be used to create attacks against other CPS with minimal modification.

**Ethics statement:**

This project does not require ethics approval, as reviewed by my supervisorDr. Sridhar Adepu.

I have completed the ethics test on Blackboard. My score is 12/12.

**Project plan:**

The iTrust Centre for Cyber Security operates a Secure Water Treatment Testbed (SWaT) which is used to “support research into the design of secure, public infrastructure” [1]. SWaT is a Cyber Physical System (CPS) so consists of the physical side which implements a process and a cyber side which performs control, monitoring and security [2].

The testbed produces clean water by using both ultra filtration and reverse osmosis which are implemented through a six stage, distributed control system.

Each stage is managed by a Programmable Logic Controller (PLC) which applies pre-programmed logic to switches and actuators based on the input signals ( sensor readings) or control messages from other PLCs. This logic is in the form of ‘Ladder’ diagrams which are a graphical programming language derived from electrical switch schematics [3].

The effect is that the PLC is a standalone micro-controller which is able to control one stage of an industrial process autonomously.

The PLCs communicate with each other using the manufacturers proprietary protocol based on TCP/IP (level 1 in the OSI model) [1] in order to pass or request data pertinent to their own stage of the process e.g. request a faster flow of water or indicate a batch is ready to move to the next stage.

The internal logic (rules) of the PLCs are designed to account for all possible system states and includes parameters such as maximum temperatures, pressures and flow rates [1]. Should a stage of the system move outside of these parameters the PLC will take appropriate action and raise an alert.

A cyber vulnerability methodology was used to identify vulnerabilities suitable for machine learning techniques to be applied to [4]. Here we don’t assume full ‘insider’ level of access- where the attacker can manipulate the rules and logic programmed into the PLC’s.

It is assumed instead that only access to the signals between PLCs are available. The TCP/IP packets which carry the control and monitoring messages are first analysed in order to passively model normal system function- identifying the rules within the PLCs as these are the parameters which dictate normal operation.

The modelling is accomplished using the normal operation data from the SWaT dataset as a proxy for the live sampling of the system. Applying unsupervised techniques such as clustering, correlation coefficients and mutual information will suggest similar or interrelated components/ stages.

The SWaT is a deterministic system as all system states are entirely predicted by the state of the other components, though it is possible to use a probabilistic approach to the interaction of components, temporal analysis is more appropriate to model the causality.

The data may exhibit seasonality due to demand patterns in output stage which could identify distinct modes of operation such as high demand/ refill stage. Auto correlation of features, the correlation of a random variable with a lagged version of itself, is a common approach- methods such as “Autoregressive Integrated Moving Average (ARIMA) and Exponential Smoothing (ES) … are widely used due to their simplicity and interpretability” [5].

Extending this auto-correlation to the multivariate dataset and recording if a time lag produces stronger correlations will be used to illustrate dependencies between components- such as a tanks water level lagging the flow in its feed pipe.

Fourier analysis, “… a decomposition of the series into a sum of sinusoidal components” [5] will be used to achieve the same aim. By considering the whole data set as complex waveforms, the fundamental frequencies and harmonics can be found for each component. Components which exhibit similar fundamentals with a consistent phase difference are likely to be functionally dependent [6].

The model of the system which results from this passive analysis will be used as the initial beliefs of an intelligent agent. The agent will test and update its beliefs by injecting values into the SWaT and comparing the result to its expected value. Models of the SWaT will be a proxy for the real system, these will be implemented using traditional and deep learning methods.

In a real-world application the agent will interact with the CPS by ‘spoofing’ the packets sent between PLCs. Using the Man-On-The-Side (MOTS) attack methodology it seeks to inject a packet into the communication channel just before the real value is sent [7]. The control system accepts this first packet and rejects the subsequent, real packet as a duplicate. For the purpose of validating the methods the agent will craft values to act as input to the ML model and compare the models output to its own predictions.

The SWaT is a competitive, multi-agent environment because the control system is trying to maintain a system state the attack agent is trying to disrupt [8]. It is assumed that the agent is able to spoof multiple packets in a single system cycle in order to suppress the effect of PLCs in other stages counteracting the attack (and so control all input values in the case of the simulation models).

To constrain the agent from choosing values which would be unusual in the real-world system, a classifier based on a Recurrent Neural Network (RNN) will be trained on the normal and attack datasets. This anomaly detector provides feedback to the agent, improving its ability to choose the rate at which it affects system values and so avoid detection. Anomaly detectors, often Support Vector Machine or Deep Neural Networks, are common in CPS as they act as Intrusion Detection Systems [2]. These anomaly detectors may reside on the SCADA system or possibly on a stand-alone (edge) processor at PLC level [2].

The intelligent agent is a simple physics-informed model- it includes heuristics for common industrial components, for example a swich is likely to have two discrete states whereas a sensor is likely to have a fixed range of values and, due to the need for digitized data, a fixed step size. The heuristic will also include temporal relationships- a switch event is likely to precede a change in another components state.

The agent will be able to create a network diagram of all components, their interactions and normal operating parameters which it has found by testing system limits. Simply exceeding these limits or increasing the duty cycle of components ( rapid switching etc.) constitute common cyber attack methods. The explainable, in depth knowledge of a previously unknown system would facilitate complex attack when there is a human in the loop.

# References

|  |  |
| --- | --- |
| [1] | iTrust Centre in Cyber Security, “Introduction to SWaT Testbed,” 5 April 2016. [Online].  Available: https://www.youtube.com/watch?v=2r1ctjULCnI. |
| [2] | J. W. C. M. P. S. C. J. S. Y. C. Yifan Jiaa,  “Adversarial Attacks and Mitigation for Anomaly Detectors of Cyber-Physical Systems,”  *International Journal of Critical Infrastructure Protection,* 2021. |
| [3] | M. Walker and C. a. M. J. Bissell, “The PLC: a logical development,” 2010.  [Online]. Available: https://www.researchgate.net/publication/  48989787\_The\_PLC\_a\_logical\_development. |
| [4] | H. B. H. O. D. T. K. K. L. V. K. Shaymaa Mamdouh Khalil,  “Threat Modeling of Cyber-Physical Systems - A Case Study of a Microgrid System,”  January 2023. [Online].  Available: https://www.sciencedirect.com/science/article/pii/S016740482200342X. |
| [5] | W. F. a. A. R. Syeda Sitara,  “A Review of Time-Series Forecasting Algorithms for Industrial Manufacturing Systems,”  2024.  [Online]. Available: https://www.researchgate.net/publication/381151823  \_A\_Review\_of\_Time-Series\_Forecasting\_Algorithms\_for\_Industrial  \_Manufacturing\_Systems.  [Accessed Aug 2024]. |
| [6] | P. Bloomfield, “Fourier Analysis of Time Series,” 2000. [Online].  Available: https://books.google.co.uk/books?hl=en&lr=&id=zQsupRg5rrAC&oi=fnd&pg=  PR11&dq=Bloomfield,+P.,+1976.+Fourier+Analysis+of+Time+Series:+An+Introduction  .+John+Wiley,+New+York.&ots=SzwmipFAWm&sig=JxHU2tf9DEfQEo3krj5LicCRA6g  &redir\_esc=y#v=onepage&q&f=false.  [Accessed 8 2024]. |
| [7] | P. M. a. K. McLaughlin, “Towards Understanding Man-on-the-Side Attacks (MotS) in SCADA,”  2004. [Online]. Available: https://arxiv.org/pdf/2004.14334. [Accessed Aug 2024]. |
| [8] | S. R. a. P. Norvig, Artificial intelligence: A modern approach, Boston: Pearson, 2020. |

**Risk Assessment:**

**1. Technical Risks**

* **Risk:** Secure Water Treatment Testbed (SWaT) dataset may have incomplete, noisy, or corrupted data making automated description of the system difficult.
* **Likelihood:** Low
* **Impact:** High
* **Mitigation:** Data cleaning and exploration to ensure statistical methods can simulate known structure.
* **Risk:** Inadequate resources for training models due to large, time-series dataset.
* **Likelihood:** Medium
* **Impact:** Medium
* **Mitigation:** Investigate methods such as preprocessing, kernel functions and Montecarlo sampling to minimise computational requirements.

**2. Modelling and Algorithmic Risks**

* **Risk:** Intelligent Agent is unable to improve on system knowledge generated through statistical methods.
* **Likelihood:** Medium
* **Impact:** High
* **Mitigation:** Apply regression or deep learning models in its place.

**4. Ethical and Legal Risks**

* **Risk:** Ethical concerns regarding the facilitation of future Cyber Attacks
* **Likelihood:** Low
* **Impact:** High
* **Mitigation:** Focus on cyber defence application of techniques and consider limiting distribution of final report.

**5. Timeline Risks**

* **Risk:** Complex project with multiple dependent stages. Delays in a single stage will effect the whole project.
* **Likelihood:** Medium
* **Impact:** Medium
* **Mitigation:** Build slack time into the timeline adjust milestones as necessary. Document difficulties and how they have led to a different overall outcome.

**Project Timeline:**

A screenshot of a spreadsheet

Description automatically generated