Contribution ID: **7** Type: **not specified**

Enhancing Industrial Control System Cybersecurity Using Artificial Intelligence for Critical Infrastructure: A Case Study on Modbus TCP Attack Detection

As Industrial Control Systems (ICS) continue to digitize and integrate with broader networks, cybersecurity threats targeting communication protocols have become a critical concern. Among these protocols, Modbus TCP—widely used in Supervisory Control and Data Acquisition (SCADA) systems. Artificial Intelligence (AI) offers a promising solution to enhance intrusion detection by learning complex patterns in network traffic. Machine learning models can effectively distinguish between legitimate and malicious Modbus TCP behavior in real time, providing a proactive cybersecurity layer for critical infrastructures such as nuclear facilities and industrial plants.

This study implemented a Python-based pipeline using a subset of a publicly available dataset from the Canadian Institute for Cybersecurity [1] [2]. The dataset contained labeled Modbus TCP packet captures categorized as normal or attack traffic. Multiple files were imported, merged, and extensively preprocessed through feature engineering. Extracted features captured both temporal dynamics and protocol-specific patterns, including inter-packet time deltas, packet length, Modbus function codes, categorized function types (e.g., read/write), sender and receiver ports, TCP control flags (SYN, ACK, FIN, RST), broadcast indicators, and packet frequency per source IP. Additionally, entropy measures and rolling statistics were computed to detect irregular traffic distributions. After feature preparation, class imbalance was mitigated using SMOTE oversampling, and the dataset was split into training and test sets. A Random Forest classifier with 100 estimators was trained to map these features to a binary classification task identifying malicious Modbus TCP traffic.

Evaluated on 1.4 million Modbus TCP packets, as shown in table 1, the model achieved balanced precision and recall values ranging from 0.95 to 0.96 for both normal and attack classes, demonstrating its ability to accurately identify malicious traffic while minimizing false positives. The confusion matrix in figure 1 confirmed this robustness, with 677,348 normal and 666,384 attack packets correctly classified and relatively few misclassifications. These results underscore the model's suitability for real-time intrusion detection in critical infrastructure networks.

Beyond standard classification metrics, model performance was further validated through Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curves. The ROC curve in figure 2 showed an Area Under the Curve (AUC) of 0.99, highlighting the model's exceptional capability to distinguish normal from attack traffic with very low false positive rates. The PR curve in figure 3 achieved an average precision score of 0.991, confirming the model's effectiveness in maintaining high precision even as recall increases. These curve-based evaluations reinforce the numerical findings and affirm the model's robustness for deployment in nuclear cyber defense applications.

[1]: Canadian Institute for Cybersecurity, CIC Modbus Dataset 2023.

[2]: Kwasi Boakye-Boateng, Ali A. Ghorbani, and Arash Habibi Lashkari, "Securing Substations with Trust, Risk Posture, and Multi-Agent Systems: A Comprehensive Approach", 20th International Conference on Privacy, Security and Trust (PST), Copenhagen, Denmark, August.2023.

Author: ABUTO'AMAH, MO'ATH (Eng)