MLSecu – Project

Year : 2023-2024
Lecturer : Pierre Parrend

Project objectives

The goal of the project is to design, deploy and evaluate a data chain for the analysis of cybersecurity data. The data treatment will be performed as batch.

Choose the objective of your analysis:

- Objective 1
  - Anomaly detection for tracking attacks
- Objective 2
  - Adversarial attacks against classification
- You can do Objective 1 + Objective 2, which gives a 25% bonus

Choose the dataset to analyse among Cybersecurity Datasets for core networks:

- Hardware In The Loop
  - Data and doc : WDT2022
- Secure Water Treatment
  - SWaT.A7_June 2020-20221017T161421Z-001.zip
  - Doc : go2016SWAT.pdf

Launch

Launch your groups:

- Set group number from 1 to 12 in: MLSecu_SCIA_2023-2024_project_groups.xlsx
- Groups will be frozen on 21/9/2023 evening

Intermediate presentation

Present your first results on 26/10/2023

Deliverables

The deliverables are:

- Analysis notebook, shared on google collab
- Analysis report (20 pages)
- Final oral group presentation (10 min) + demonstration (5 min)

Your report will present the detailed specification and implementation details on:

- The complete deployment of the data handling chain (including classification + anomaly detection)
- Characterization of the dataset under study
- Benchmark of 3 complementary analysis algorithms, including confusion matrix, precision, recall, AUPRC, Matthews Correlation Coefficient
- Conclusions about cybersecurity events in the dataset

The oral presentation is a security analysis review based on the report.

Specifications

The data handling chain will comply with following requirements:

The choice of the dataset, the design of the data handling chain as well as the choice of the analysis algorithm is part of the work.