Boosting IoT Intrusion Detection systems with LLM-based dataset augmentation

Deep Learning and other Machine Learning algorithms have been extensively used in recent research literature to facilitate traffic classification and anomaly detection-based intrusion detection systems for the Internet of Things (IoT). However, often such systems suffer from a high number of false positive alarms due to insufficient datasets used in training the detection models. In recent work, we have studied the applicability of Large Language Models (LLM) for augmenting IoT IDS datasets with promising results. The task in this topic would be to investigate the use of pre-trained foundation LLM models for the purpose of augmenting existing IoT datasets and evaluate the performance on state-of-the-art IoT intrusion detection systems with the goal of increasing the overall resilience of these systems.

To be able to successfully work on this topic, you should have a basic understanding about recent AI algorithms like Large Language Models, some familiarity with AI frameworks like PyTorch[1] or TensorFlow [2], and a willingness to learn and quickly adopt knowledge on AI, IoT security and intrusion detection.

Note: This represents only a preliminary high-level description of the topic. Details and actual scoping of the thesis will be agreed on with separately with the supervisor.

Thesis supervisor:
Prof. Dr.-Ing. Markus Miettinen
markus.miettinen@fra-uas.de
+49 69 1533 3969

Secondary supervisor: Maurizio Petrozziello
petrozziello@fra-uas.de

Thesis language: English or German

Topic updated: June 11, 2025

[1] https://pytorch.org/
[2] https://www.tensorflow.org/