Thesis Topic Example 1

Evasion Attacks Against ML-based Malware Detectors

Keywords: cybersecurity, machine learning, adversarial AI

Background:
Malware detectors, as well as other ML-based detection systems, have proven to be susceptible against adversarial examples, i.e specialized inputs created with the sole purpose of confusing a detection system, resulting in the misclassification of the crafted input. Most attacks presented in the literature assume partial or complete knowledge of the target detectors, which is not available to the attackers in a real-world scenario. When targeting a Next-Gen Antivirus (NGAV), attackers do not have information about the algorithm, its parameters, the input and the output of the detection system. In addition, attackers might be restricted to only a very limited number of queries to the classifier. This project involves understanding how NGAV work, studying current methods to generate adversarial malware examples, and adapting SOTA methods or devising new methods to work in a query-limited scenario.

Aim:
The expected outcome of the project is a method to generate adversarial malware examples in a query-limited scenario, a report describing the method's details, the findings and the challenges, and a repository with the source code of the method.