Samenvatting

Met het snel veranderende landschap van de cybersecurity verschijnen er voortdurend nieuwe tools, aanvalstechnieken en exploits om misbruik van te maken. Onderzoek en ontwikkeling van nieuwe verdedigingsmiddelen is essentieel om de vertrouwelijkheid, integriteit en beschikbaarheid van onze netwerken en waardevolle data te waarborgen.

Het stoppen van een cyberaanval begint bij de detectie ervan. Hiervoor zijn Network Intrusion Detection Systems ontwikkeld, die automatisch aanvallen kunnen detecteren en een alarm kunnen genereren. Met de opkomst van artificiële intelligentie worden steeds vaker machine learning-technieken in dit domein toegepast. Echter, hoogwaardige modellen vereisen hoogwaardige data. Er zijn veel datasets geconstrueerd, maar elk van hen kampt met tekortkomingen die de bruikbaarheid beperken. Bovendien kunnen deze modellen, door de snelle evolutie van tools en technieken, snel verouderd raken.

Deze scriptie onderzoekt een nieuwe methode om de robuustheid van machine learning-modellen voor netwerkdetectie te verbeteren via het ConCap-framework. Aan de hand van CIC-IDS-2017 als casestudy wordt deze dataset gereconstrueerd en uitgebreid om te verifiëren of ConCap op haalbare wijze als vervanger voor CIC-IDS-2017 kan dienen en hoe het uitbreiden van CIC-IDS-2017 met ConCap de robuustheid van modellen kan versterken.

Summary

With the fast-evolving landscape of cybersecurity come new tools, new attack techniques and new exploits to abuse. Research and development of new defensive tools is imperative for maintaining the confidentiality, integrity and availability of our networks and our precious data.

The process of stopping any cyberattack starts at its detection. For this purpose, Network Intrusion Detection Systems have been developed to automatically detect attacks and raise an alarm. With the advent of artificial intelligence, more and more machine learning techniques are being used in this area. However, high quality models require high quality data. Many datasets have been constructed, but each came with a set of issues that reduce its usability. Furthermore, with the fast evolving tools and techniques, these models can be quickly rendered obsolete.

This thesis studies a new way of improving robustness of machine learning models intended for network intrustion detection through the ConCap framework. Using CIC-IDS-2017 as a case study, it gets reconstructed and extended to verify that ConCap can feasibly be used as a stand-in for CIC-IDS-2017 and how extending CIC-IDS-2017 through ConCap can enhance model robustness.