Network Intrusion Detection using pre-trained tabular representation models

Network Intrusion Detection (NID) is a common topic in cybersecurity. However, it is not trivial to find a solution when facing the complicated network environment nowadays. Often a complex system is needed to process enormous volume of data stored in databases. This thesis proposes to use Deep Learning models to tackle the NID problem in a pre-train/fine-tune manner. As the new paradigm of transfer learning, the process of pre-training follows by fine-tuning has achieved huge success in many areas such as vision and NLP. We aim to study whether those trending models still perform well on large-scale structured data such as network security logs. It is plausible to leverage the strong learning ability of DL models to learn table representations and separate anomaly from benign records based on the learned information.

Prerequisites

- Machine learning knowledge
- Programming skills (Python, Git)
- Computer networking knowledge

Advisors

Hasan Yagiz Özkan, Cristian Bermudez Serna
Dr. Haojin Yang (HPI)