Leveraging Data from Open-Source Intrusion Detection Systems for Enhancing Security of Systems

Viren Chhabria, Master of Science in Computer Science

University of Dublin, Trinity College, 2019

Supervisor: Dr. Stephen Farrell

Bad actors are omnipresent across the Internet trying to gain unauthorized access to systems for malicious activities. Intrusion Detection Systems(IDS) are like burglar alarms in computer science designed the detect and prevent intrusion attempts. This makes them a widely researched and important topic in the field of computer science. In this research, 26 months of historical data gathered from live internet facing servers running popular open-source IDS - Fail2Ban and DenyHosts is analyzed. A scoring mechanism ranging from 0-4 was devised to associate a threat level to each intrusion detected. Exploratory Analysis for finding geo-spatial, temporal and other underlying patterns in the data. Predictive Analysis was carried out using supervised machine learning techniques like time-series forecasting and classification to verify whether it is possible to efficiently predict intrusion attempts or the threat associated with them. The result was a software pipeline being created to ingest data from IDS and produce a security dashboard reflecting the security state of the systems and patterns in the attacks. The predictive analytics results were interesting. The time-series forecasts generated using fbProphet were inaccurate, highlighting the uncertainty and variability in attacks. The XGBoost classification model was able to predict the threat level associated with each attack with 75.46% accuracy on the validation set.