Observations on Automating Intrusion Analysis using Machine Learning

Loai Zomlot, Sathya Chandran, Doina Caragea, Xinming Ou
Kansas State University
Manhattan, KS, USA
{lzomlot, sathya, dcaragea, xou}@ksu.edu

Intrusion analysis, *i.e.*, the process of combing through IDS alerts and audit logs to identify real successful and attempted attacks, remains a difficult problem in practical network security defense. The major contributing cause to this problem is the high false-positive rate in the sensors used by IDS systems to detect malicious activities. The goal of our work is to examine whether a machine-learned classifier can help a human analyst filter out non-interesting scenarios reported by an IDS alert correlator, so that analysts' time can be saved. Furthermore, we would like to present some observations from our experience of applying machine learning approaches in intrusion analysis. This research is conducted in the open-source SnIPS intrusion analysis framework. Throughout observing the output of SnIPS running on our departmental network, we found that an analyst would need to perform repetitive tasks in pruning out the false positives in the correlation graphs produced by SnIPS. We hypothesized that such repetitive tasks can yield (limited) labeled data that can enable the use of a machine learning-based approach to prune SnIPS output based on the human analysts' feedback, much similar to spam filters that can learn from users' past judgment to prune emails. Our goal is to classify the correlation graphs produced from SnIPS into "interesting" and "non-interesting", where "interesting" means that a human analyst would want to conduct further analysis on the events. We spent significant amount of time manually labeling SnIPS output correlation graphs based on this criterion, and built prediction models using both supervised and semi-supervised learning approaches. Our experiments revealed a number of interesting observations that give insights into the pitfalls and challenges of applying machine learning to intrusion analysis. The experimentation results also indicate that semi-supervised learning is a promising approach towards practical machine learning-based tools that can aid human analysts, when a limited amount of labeled data is available.