Senior Project– Computer Science- 2015

Classification of System Call Sequences using Anomalous Detection

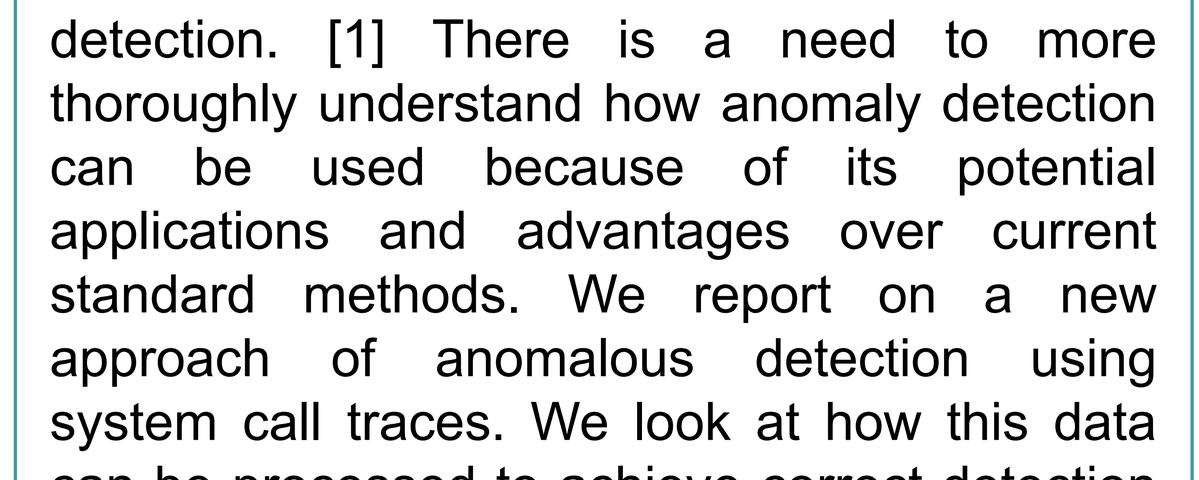
William Doyle
Advisor – Aaron Cass

Abstract

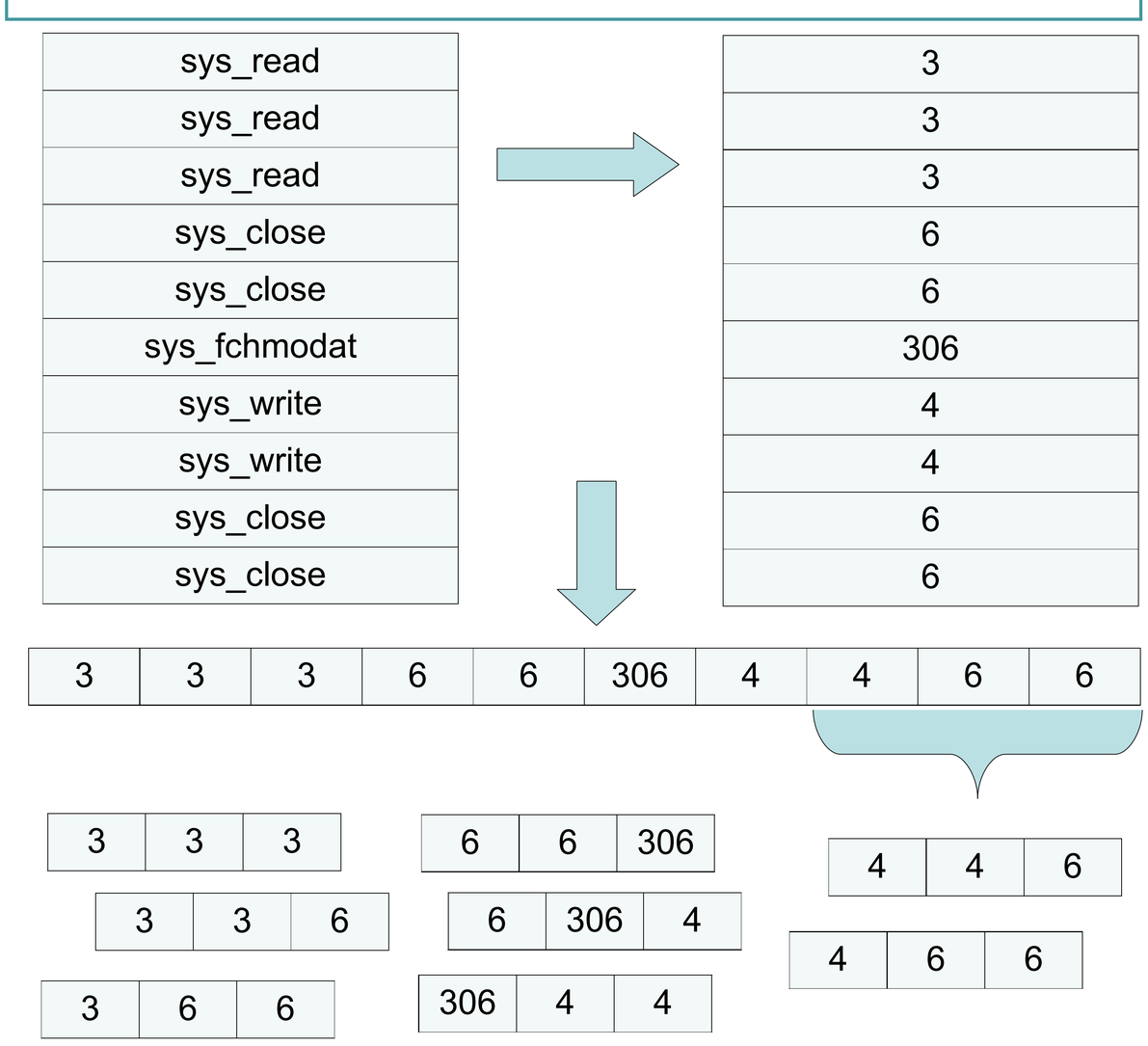
Recently, work has been done within the intersection of security and machine learning to better understand anomalous intrusion detection. [1] There is a need to more thoroughly understand how anomaly detection can be used because of its potential applications and advantages over current standard methods. We report on a new approach of anomalous detection using system call traces. We look at how this data can be processed to achieve correct detection of intrusions on a system. Our goal is to outline ways in which system call traces can be leveraged as well as what we can do and learn from these results.

Compressing and Results

There are many possible sequences as a result of there being a large collection of individual system calls for the operating system. “Combining” some of the system calls into one compressed system call during the preprocessing can aid in computation and classification. We observed statistical improvement when using subsets of the data. Including all of the data at training time does not yield different models, but similar ROC curves.



ROC curve classification of un/compressed system call traces



Example: Generating an n-gram from system call trace

Avoiding Security Risks

We can “learn” what a typical sequence of events an attacker would take. We accomplish this task by looking at traces of system calls on a system. Operating systems observe a sequence of events in the form of these system calls. Using these system calls we can analyze what makes up an attack versus normal behavior. [2]

Future Work

Optimizing the classifier being used to further separate attack and normal characteristics. Explore changing the length of n-gram, metric for compression, and the classification threshold during the data processing and classification phases of the experiments.

References:
[1] J. Hu. “Host-based anomaly intrusion detection”, in Handbook of Information on Communication Security, Springer Berlin Heidelberg, 2010, pp. 235-255
[2] S. Forrest, S. A. Hofmeyr, A. Somyaji, and T.A. Longstaff, “A sense of self for unix processes” in Security and Privacy, 1996, Proceedings, IEEE Symposium, pp. 120-128