Project Description: Behavioral Models for the Detection of Unusual Asset Access

Company: Interset
Supervisors: Shaun Pilkington, Data Science, Interset

The Company

Interset provides a highly intelligent and accurate insider and targeted outsider threat detection solution that uses data science, behavioral analytics, machine learning and big data to protect critical data.

The Science

Our approach is to identify risk within the organization by ingesting activity data from various sources (such as desktops, servers and data repositories) and then detecting abnormal, suspicious and high risk behaviors of user accounts, machines, and intellectual property assets (such as files). We do this by learning normal, baseline patterns from the data (using machine learning and mathematical models), and then quantify abnormal behavioral through probabilistic models that look for differences between observed behavior and baseline.

The Project

We incorporate new data sets by doing exploratory data analysis (EDA) on actual log datasets (provided by our customers) that describe human activity surrounding valuable data repositories to protect, such as source code audit logs, cloud repository activity information, and database query logs.

For this project, models will be developed to determine when an unusual asset or resource is accessed. This represents activity that is very useful for the predictive detection of theft through a compromised account: for example, John Smith’s account is hacked, and is now found accessing areas of the database that are unusual for John.

This can be determined based on the user's past behavior, or the behavior of other users (e.g. collaborative filtering). The datasets for this will include logs from:

- Active Directory
- Source Code Repository
- Sharepoint IIS logs

The Team

We catch bad guys with math. Our mission is to enable our clients to protect their data using new insights and advanced analytics. We are smart and creative. We are empowered to make a difference.