Discovering the incident in Incident Response

MORE INFO MORE PROBLEMS

Starbucks currently uses Splunk as a Security Information and Event Management tool. Splunk aggregates machine data from all across the organization and indexes it to make it searchable for analysis and reporting.

The Splunk Enterprise Security application features various dashboards that communicate information about “notable” events. These events are defined by custom search queries that correlate events across different indexes or information sources.

"IT'S NOT INFORMATION OVERLOAD. IT'S FILTER FAILURE" - CLAY SHIRKY

Machines produce the data; machines can learn from the data

- Dimensionality Reduction with PCA
- Cluster Analysis with DBSCAN
- Random Forest Classification

EXPLORING THE DATA & DEFINING A USE CASE

Clustering identifies groups of events with similar characteristics. Can be used to identify and label events of interest.

Authentication Failure <= 5,500

User name correct but password wrong <= 0.5

Failure proportion <= 0.168

Prioritize outliers because they represent behavior that is unlike other notable events. Underlying assumption is that most notable events are actually normal behavior.

EVENT CLUSTERING

Cluster 1

Cluster 3

Cluster 5

Outlier

WHAT'S NEXT?

- The power of this analysis is that it identifies outliers with regard to other source behavior as well as a single source’s historical behavior. However, it is highly sensitive to the parameters used in clustering. Therefore exploring the characteristics of clusters is vital.
- A natural next step in this process involves identifying times when automated action could be taken. For example, if an event is classified as an outlier and meets various conditions we might suspend a user’s account.
- Machine learning can be applied to a wide range of security use cases. A particular use case of interest at Starbucks is using natural language processing to identify DNS exfiltration attacks.