LA-UR-17-21663
Approved for public release; distribution is unlimited.

Title: User Behavior Analytics

Author(s): Turcotte, Melissa
Moore, Juston Shane

Intended for: Report

Issued: 2017-02-28

User Behavior Analytics

Melissa Turcotte
Juston Moore

Advanced Research in Cyber Systems
Los Alamos National Laboratory

February 27, 2017

User Behaviour Analytics

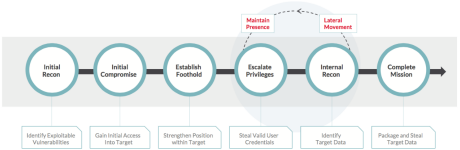
User Behaviour Analytics is the tracking, collecting and assessing of user data and activities.

- Goal: Detect misuse of user credentials by developing models for the normal behaviour of user credentials within a computer network and detect outliers with respect to their baseline.

Los Alamos NATIONAL LABORATORY

Adversaries and user credentials

- External adversary
  - Reusable user credentials are one of the most powerful items an attacker can obtain
  - Adversaries generally have to get access to user credentials to move through the network
- “ insider threat” or rogue user.
  - May result in credential abuse, i.e accessing unauthorized file shares, exfiltrating data

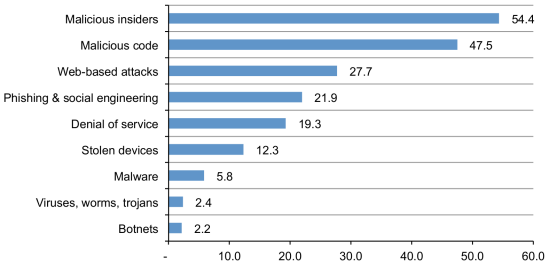


Mandiant M-Trends 2016 Report

“63% of confirmed data breaches involved weak, default or stolen passwords.”

Verizon 2016 Data Breach Investigations Report

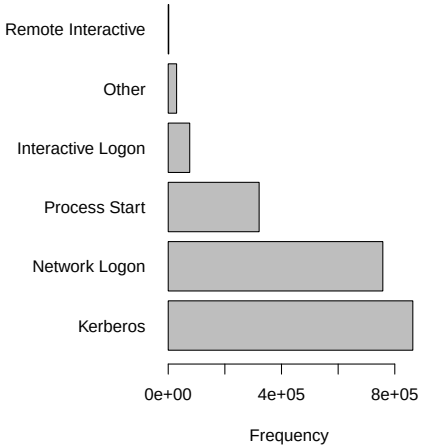
Figure 13. Some attacks take longer to resolve
Estimated average time is measured for each attack type in days
Consolidated view, n = 252 separate companies



Ponemon 2015 Cost of Cyber Crime Study

Los Alamos NATIONAL LABORATORY

Data Source



Computer event logs are a critical resource for investigating security incidents.
They can give detailed information about what is happening at a machine level.

- authentication, logons
- processes
- applications/services

Many of these log entries are tied to a user credential action.

Future Data Sources

- Badge reader data
- HR data
- proxy logs
- e-mail logs

Los Alamos NATIONAL LABORATORY

Approaches

Many rule-based approaches for looking at computer event logs to detect security incidents, which require knowing what indicators attackers generate (reactive).

Two complimentary statistical-based approaches have been considered:

User behaviour anomaly detection

- View the computer event logs as a multivariate stream of data with different characteristics associated with each user credentials.

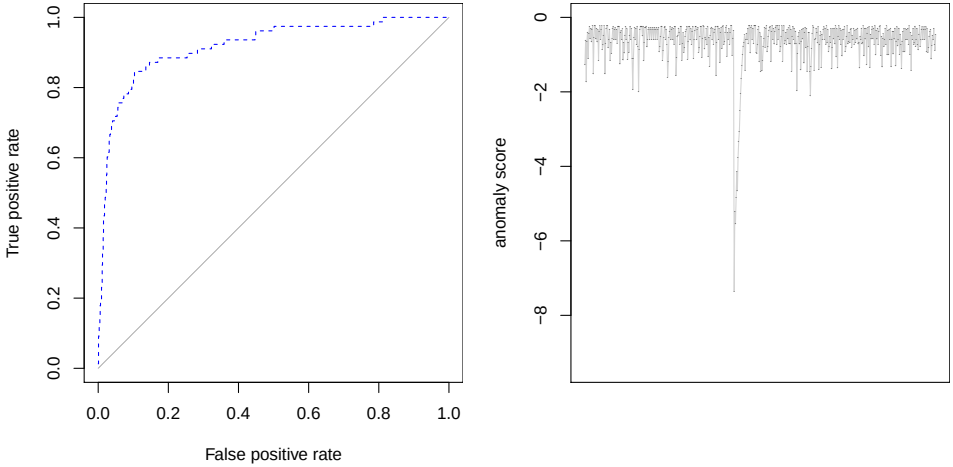
$$\{(X_t, Y_t, E_t) : t = 1, 2, \ldots\}$$

  X_t = client, Y_t = server, E_t = event type.
- Build probability models for normal user credential behaviour based on their historical and current network usage.
- For each new observed event, use the probability models to obtain a score for how likely the observed event is according to the users historical behaviour.

Los Alamos NATIONAL LABORATORY

Detection of the 2013 red team attack

ROC curve and anomaly scores over time for a compromised user.



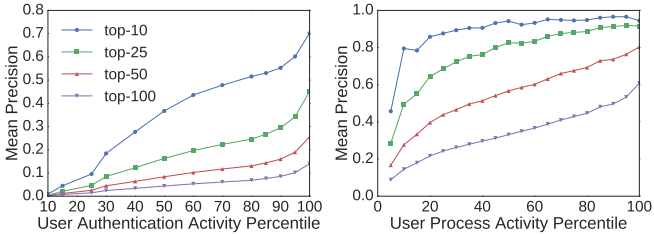
Peer-based anomaly detection: Recommender systems

Early research.

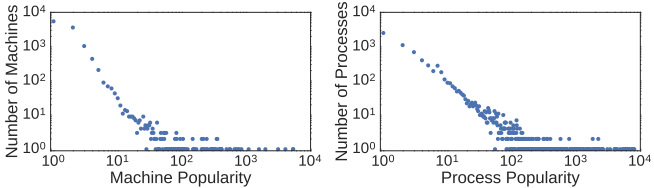
- Leverage behaviour of similar users (peers) to better predict individual actions, reducing false alarms.
- Utilise recommender system algorithms to predict user actions that are unlikely based on peer-group preferences.
- Allows for different peer groups depending what features of the data are being considered.

Los Alamos NATIONAL LABORATORY

Model fit

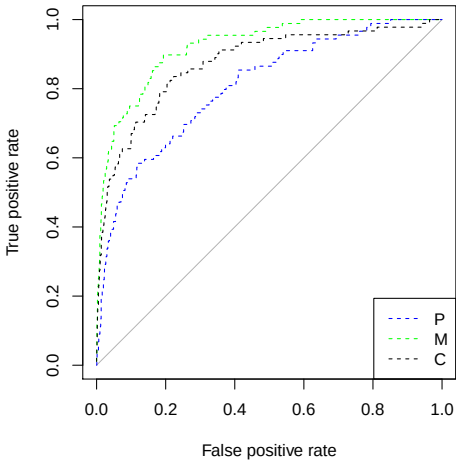


Mean precision for users with varying levels of activity



Log-log plot of the empirical distribution of the popularity of processes and machines

Detection of the 2013 red team attack



- Low false-alarm rate is paramount if any anomaly detection systems are to be used by an operations analyst.
- Four out of the top 10 most anomalous users were known compromised credentials.

Path Forward

- Combine the two approaches above to provide a robust overall model for UBA.
- Utilise more data sources to get a more holistic view.
- Software development for UBA with a commercial partner.

Los Alamos NATIONAL LABORATORY