Human Machine Teaming

Traffic Characterization Challenge

November 2019

1 Context

IT analysts at a company have received alerts of suspicious activity from the intrusion detection system (IDS) that monitors company network traffic. After reviewing the alerts, the analysts believe there are two cases that require further investigation. First, they suspect that an employee has been streaming Netflix in violation of company policy. Second, they suspect that malware has been downloaded on an employee computer on October 31st, 2019 approximately between 12:00-2:00PM. The analysts inspected all employee network traffic and were able to rule out almost all employees as potential suspects except for the five employees that have been using Virtual Private Networks (VPNs) to obscure their network activity.

This challenge is an opportunity for researchers to utilize classification methodologies to categorize VPN traffic by applica-tion of origin. We also encourage researchers to explore the human-machine teaming aspect of this challenge by providing a means for expert analysts to interpret the results.

2 Data

Training Dataset

The training dataset consists of VPN traffic flows and is labeled by application of origin. The traffic flows correspond to single applications. The included applications are:

1. VoIP
2. SSH
3. RSync
4. Netflix
5. YouTube
6. Chrome
7. Skype Chat

The training dataset will be a JSON file with the following fields, see Table 1

DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited. This material is based upon work supported under Air Force Contract No. FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Air Force.

Table 1: Training Dataset Features

Feature	Description
connection	String containing source IP, source port, destination IP, destination port, and packet protocol.
timestamps	Array of all the packet timestamps, in order of arrival, for the given connection.
sizes	Array of all the packet payload sizes, in order of arrival, for the given connection.
directions	Array of Booleans indicating packet flow direction, in order of arrival, for the given connection. 1 indicates a packet traveling from the source IP to the destination IP and 0 indicates the reverse.
application	Name of application associated with the given connection.

Test Dataset

The test dataset will consist of network traffic captured for each suspected employee from 12:00-2:00PM on October 31st, 2019. Unlike the training dataset, the test dataset will not be labeled and will consist of one stream of timestamps, sizes, and directions of at most two overlapped applications running consecutively for each employee, see Figure 1.

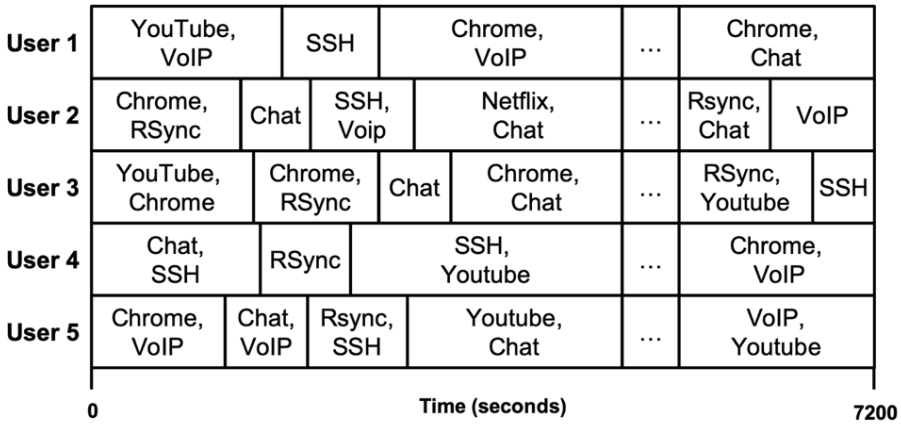

Figure 1: Example Schedule

The test dataset will be a JSON or pickle file with the same fields as the training dataset but will not include application labels.

3 Task

The goal of this challenge is to classify consecutive applications running under a VPN to identify the employee that has been streaming Netflix in violation of company policy and the employee that has compromised the security of the company network by downloading malware.

4 Submission

Researchers will submit a brief one-page document providing the following information:

- A concise explanation of the techniques used to classify the test dataset.
- A list of the applications used by each employee by second, for example, see Figure 2.
- The identity of the employee that was streaming Netflix and the employee that downloaded malware with justification for each.
- Human-interpretable representation to display classification results to analysts (e.g., visualization)

time_in_seconds	app1	app2
0	netflix	skype-chat
1	skype-chat	voip
2	voip	netflix
3	skype-chat	voip
4	rsync	-
5	chrome	youtube
6	ssh	-
...	...	...
7197	netflix	netflix
7198	chrome	-
7199	rsync	ssh
7200	youtube	netflix

Figure 2: Example Submission

5 Evaluation

Submissions will be judged on overall classification accuracy as well as the quality of the justification used to determine the identity of the two noncompliant employees.