THREAT DETECTION THROUGH CORRELATION OF NETWORK FLOWS AND LOGS

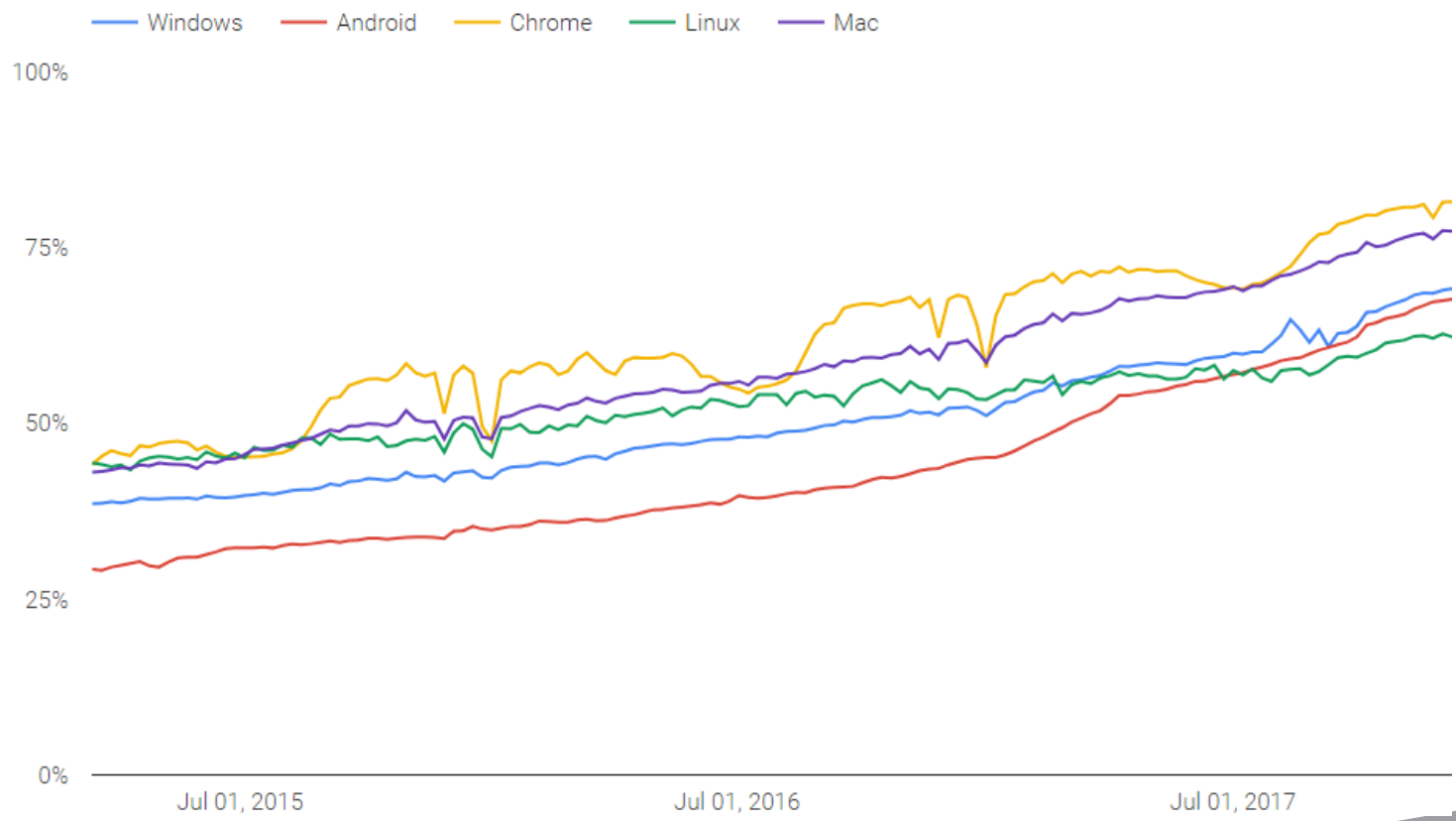
Tuesday 5th June, 2018

Stanislav Spacek
Pavel Celeda

CSIRT-MU

Rise of Encrypted Traffic

Percentage of pages loaded over HTTPS in Chrome by platform

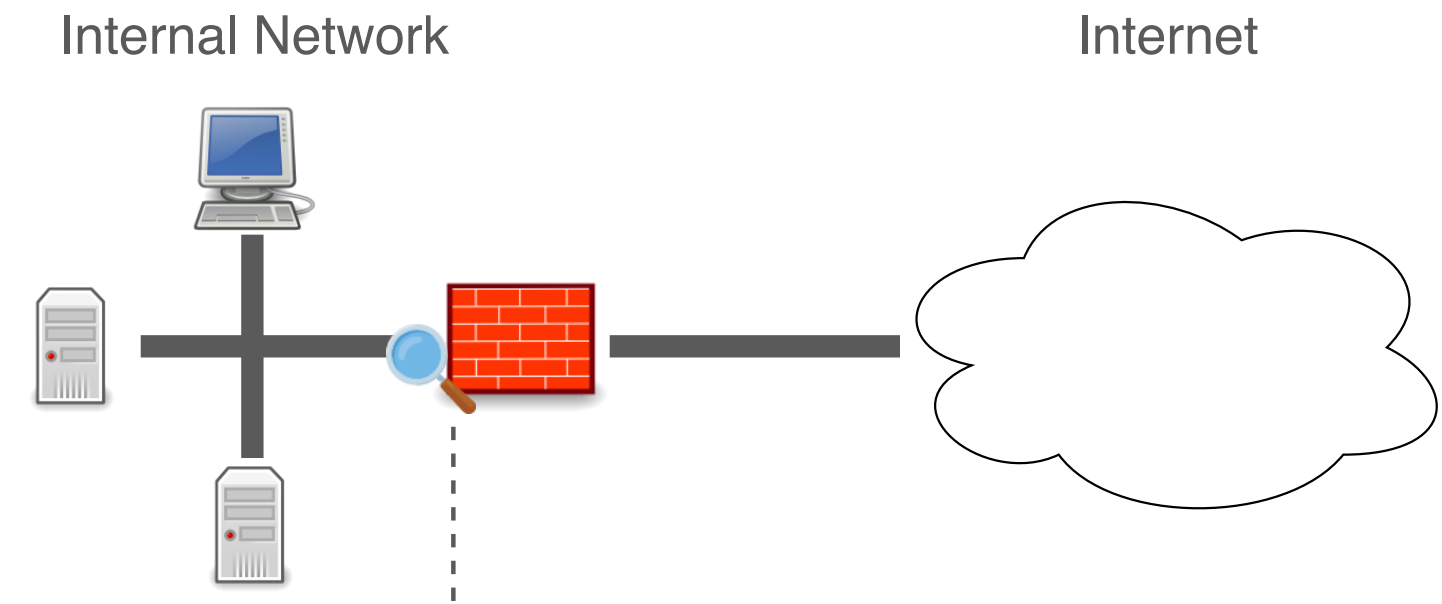


CSIRT-MU

Research Goal

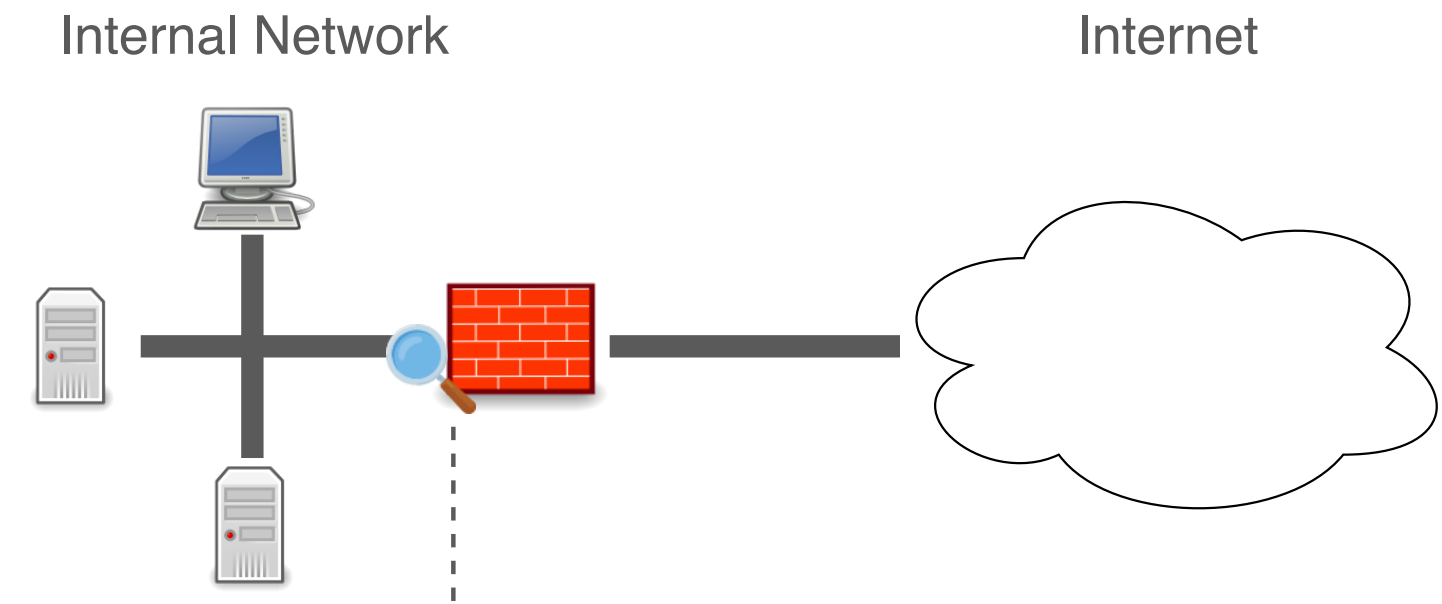
- Enriching flow analysis with information stored in logs
- Detect network threats using the combined base of data
- Tackle the visibility restriction in encrypted traffic

Flow Analysis



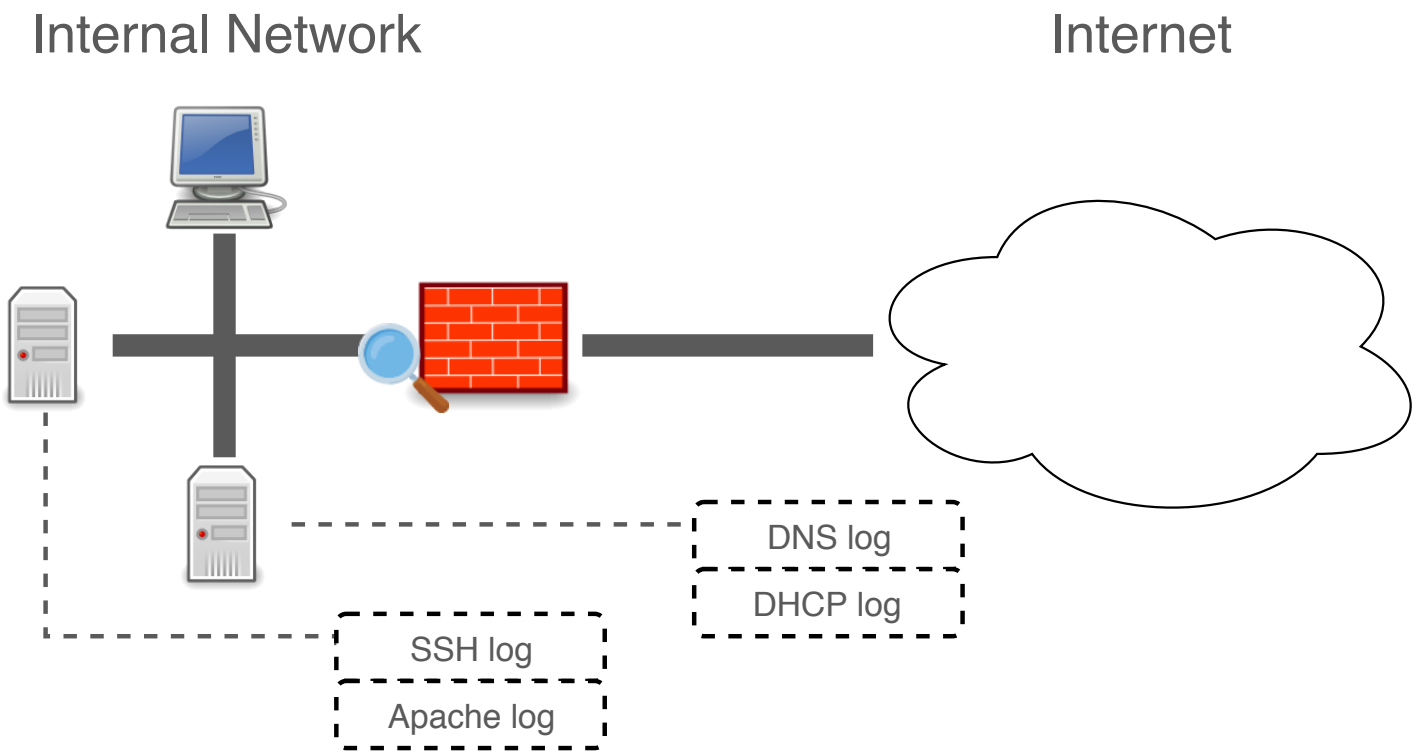
Start	Duration	Source	Destination	Packets	Bytes	Application Ext.
10:38:01.141	0.3	10.1.1.15	192.168.1.2	12	28	HTTP ...
10:38:03.158	0.8	192.168.1.2	10.1.1.15	5	38	HTTP ...
10:42:56.247	3621.5	192.168.12.3	10.3.8.158	4229681	550362187	BitTorrent ...
10:42:56.985	3622.3	10.3.8.158	192.168.12.3	3542189	458713698	BitTorrent ...

Flow Analysis Enrichment



Start	Duration	Source	Destination	Packets	Bytes	Application Ext.
10:38:01.141	0.3	10.1.1.15	192.168.1.2	12	28	SSL/TLS
10:38:03.158	0.8	192.168.1.2	10.1.1.15	5	38	SSL/TLS
10:42:56.247	3621.5	192.168.12.3	10.3.8.158	4229681	550362187	SSL/TLS
10:42:56.985	3622.3	10.3.8.158	192.168.12.3	3542189	458713698	SSL/TLS

Log Analysis Enrichment



CSIRT-MU

Research Questions

- How can we ameliorate the network flow analysis with logs to detect threats in a constantly evolving environment?
  - Logs and flows are captured in different forms
  - The logs are heterogeneous
  - A common base of data must be established
- How can the network flow and log correlation improve the threat detection rates in encrypted traffic?
  - Detection method design
  - Detection accuracy testing

Proposed Approach I

Log representation

- Transform logs into a unified form
- Treat each log as a stream of defined events

Flow and log correlation

- Correlation based on shared parameters – source, destination and timestamp
- Timestamp tolerance range will need to be specified

Proposed Approach II

Detection method design

- Will be based on association rule learning algorithms
- It will need to balance:
  - The detection speed and accuracy
  - The accuracy itself (F-score)

Detection accuracy testing

- Testing in a cloud-based testbed for simulating cyberattacks
- Testing dataset based on real network traffic with injected attacks
- The dataset will be made public

Summary

- Mass usage of encryption restricts network flow monitoring
- Correlation of flows and logs should provide additional insight
- Threat detection method based on the correlated data should provide better accuracy

CSIRT-MU

THANK YOU FOR YOUR ATTENTION

csirt.muni.cz

@csirtmu

Stanislav Spacek

spaceks@ics.muni.cz