Web Security Analysis Toolbox
About us

VP Research
@amirshaked
amirshk@perimeterx.com

CTO
@safruti
ido@perimeterx.com
Scalable, behavior-based threat protection platform for web, cloud and mobile
Prerequisites

- Get a GCP account
  - Go to [https://console.cloud.google.com/](https://console.cloud.google.com/) and follow instructions to create account
  - email access application to oreilly17@perimeterx.com
- Install the gcloud command-line tool by [installing Google Cloud SDK](https://cloud.google.com/sdk/docs/install)
- Install the datalab Cloud SDK component
- Create your own datalab instance by following these instructions: [https://cloud.google.com/datalab/docs/quickstarts](https://cloud.google.com/datalab/docs/quickstarts)
- Kibana - get access to Kibana instance in session
- [https://github.com/PerimeterX/web-security-analysis-toolkit-workshop](https://github.com/PerimeterX/web-security-analysis-toolkit-workshop)
Intro
Our Setup

- **Web Server**: Apache, Webapp, Filebeat
- **ELK**: Kibana, Logstash, Elastic
- **Big Query**: Data Lab, Data Studio
- **External Resources**: You

- **User**: JS
Collected Data

- Access logs
  - Specific HTTP headers
- Application logs
  - Controller route
  - Post data
- Browser fingerprinting
Traits of an analyst

Lots of data, but incomplete and untrusted

Heuristics and biases

Growing fields of required knowledgebase
Heuristics and biases

Availability heuristic
Gambler's fallacy
Conjunction fallacy
Insensitivity to sample size
Pluralistic ignorance

PLURALISTIC IGNORANCE

I'M NOT REALLY SURE WHAT'S GOING ON
SO I'LL JUST MIMIC EVERYONE ELSE
Kibana Intro
“Kibana is an open source analytics and visualization platform designed to work with Elasticsearch. You use Kibana to search, view, and interact with data stored in Elasticsearch indices. You can easily perform advanced data analysis and visualize your data in a variety of charts, tables, and maps.”

Kibana has no ACL - so play nice!
Learn by trying - [https://oreilly17.perimeterx.com/elk/](https://oreilly17.perimeterx.com/elk/)

- (lab:lab)
- Lucene syntax
- Which activity types we have

**K-Lab 1: Filter out bogon addresses**

*Beware of the quick distribution for large datasets*
K-Lab 2:
- Heat map os/browser
- Client map
- Bubble of return bytes vs paths

Add q query filter - path:login
Look and feel
Dashboards

K-Lab 3:
- Add all visualizations
- Add a filter - only China
get _template
Put _template
External Resources

Ipinfo.io
udger.com
maxmind.com
shodan.io
censys.io
pastebin.com

mxtoolbox.com
http://iplists.firehol.org/

Test tools:
Chrome Modheaders
Chrome DevTools
POSTMAN
Big Query
Best Practices

- Column based
- Single pass
- Minimum joins
- UDF
  - Emit rows
  - Generate new values
Standard SQL vs Legacy SQL
- UDF
- Different functions
- Table range vs wildcards

Views
Good bot detection

BQ-Lab 1: Find the spoofed bots by IP and classification
BQ-Lab 2:
1. Find bad ips
2. Use UDF to parse query string
3. Bonus - which value is least common, and if the “bad” ips reached it
Account Takeover (ATO)

BQ-Lab 3:
1. ATO - show histogram of actions per minute
2. ATO - find threshold per minute
3. Find IP addresses above threshold
4. Compare to is_bad field
5. Advanced - include the headers as well
Data Lab
External Resources

DL-Lab 1:
1. Add data from external resource (ipinfo)
2. Check for good bot user-agent spoofing
Anomalies on login

DL-Lab 2: Search for anomalies in #of login attempts
1. Graph histogram
2. Non standard user-agent or old user-agent ???
Data Studio
Elasalert - show demo files
Final Project
Final Project

Find the histogram of content-length return values from webserver.
Find the anomalies.
Check them on the kibana to see a flow.
Find the fingerprint.
Distribution of response bytes
Find the anomaly
Find the webshell you have on your server
ElastAlert - do even more, show examples (from our GIT)
Behavior-Based Web Protection

Amir Shaked, VP Research
amirshk@perimeterx.com
@amirshaked

Ido Safruti, CTO
ido@perimeterx.com
@safruti