Jupyter Notebooks & Pre-Recorded Data Sets

Threat Hunting
@Cyb3rWard0g & @Cyb3rPandaH

- Projects
  - @HunterPlaybook
  - @THE_HELK
  - ATTACK-Python-Client
  - @OSSEM_Project
  - @Mordor_Project
  - OpenHunt
  - Blacksmith & More
- Founders:
  - @HuntersForge

https://github.com/hunters-forg
Agenda

● Threat Hunting Challenges
● The Mordor Project
  ○ Datasets on-demand!
● Jupyter Notebooks Overview
  ○ Flexible language agnostic data analysis!
● Jupyter + Mordor Datasets
● Hunt the Planet! #HuntThePlanet
Threat Hunting Challenges
Threat Hunting: Day One!
## A Few Threat Hunting Challenges

<table>
<thead>
<tr>
<th>Challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rules Development</td>
</tr>
<tr>
<td>Data Quality</td>
</tr>
<tr>
<td>Detection Validation</td>
</tr>
<tr>
<td>Passion &amp; Talent</td>
</tr>
<tr>
<td>Leadership Support</td>
</tr>
<tr>
<td>Framework/Structure</td>
</tr>
<tr>
<td>Big Data</td>
</tr>
<tr>
<td>Data Documentation</td>
</tr>
<tr>
<td>Adversary Knowledge</td>
</tr>
<tr>
<td>Disparate Data</td>
</tr>
</tbody>
</table>
A Few Threat Hunting Challenges

Data Analytics

- Rules Development
- Framework/Structure
- Data Quality
- Big Data
- Detection Validation
- Data Documentation
- Passion & Talent
- Leadership Support
- Adversary Knowledge
- Disparate Data
Data Analytics Development

- Define a Research Goal
- Model Data
- Simulate Adversary
- Define Detection Model
- Validate Detection Model
- Document and Communicate Findings

https://www.youtube.com/watch?v=DuUF-zXUzPs
Data Analytics Development

1. Define a Research Goal
2. Model Data
3. Simulate Adversary
4. Define Detection Model
5. Validate Detection Model
6. Document and Communicate Findings

https://www.youtube.com/watch?v=DuUF-zXUzPs
Data Analytics Development

1. Define a Research Goal
2. Model Data
3. Simulate Adversary
4. Define Detection Model
5. Validate Detection Model
6. Document and Communicate Findings

https://www.youtube.com/watch?v=DuUF-zXUzPs
More than just testing security controls!

- Simulate Adversary
  - Test Security Controls
    - Endpoint Agent Detection
    - Analytics Platform Rules
    - Can I see it in my environment?
  - Model Adversary Behavior
    - Learn adversary behavior
    - Map data sources to adversary actions
    - Study derived techniques
Basic Adversary Simulation

Define Technique

Research Goal

Prepare Environment

Dev or Prod Setup

Telemetry Available (Full Auditing?)

Execute Simulation

Document Output

Prepare Data Exploration

Provide Metrics
Execute -> Collect -> Analyze -> Repeat

Simulating Adversarial Technique

Test Security Controls

Model Adversary Behavior

Data produced
Execute -> Collect -> Analyze -> Repeat

Simulating Adversarial Technique

Data produced

Test Security Controls

Model Adversary Behavior

TAKES TIME

Define Technique

Prepare Environment

Execute Simulation

Document Output

- Research Goal
- Dev or Prod Setup
- Telemetry Available (Full Auditing?)
- Prepare Data Exploration
- Provide Metrics

TAKES TIME
Same Technique + Some Variations

Data produced

Test Security Controls

Model Adversary Behavior
Same Technique + Some Variations

Define Technique
Prepare Environment
Execute Simulation
Document Output

Same Data?
Similar Events?

Test Security Controls
Model Adversary Behavior

Data produced
Lateral Movement via WMI (Win32_Process Create)

- `wmic /node:172.18.39.106 /user:Administrator /password:P1ls3n!` process call create cmd.exe
- `Invoke-WmiMethod -ComputerName 172.18.39.106 -Credential Administrator -Class Win32_Process -Name Create -ArgumentList notepad.exe`
- `SharpWMI.exe action=create computername=HR01.shire.com command="powershell.exe -enc ZQBj..."
- `./wmiexec shire.com/pgustavo@172.18.39.106`
Lateral Movement via WMI - Behavior

User

Host

Process

svchost.exe

Logon Session: Logon Type 3

creates

Process

Wmiprivse.exe

creates

Process

?.exe

Process

?.exe

https://twitter.com/HunterPlaybook/status/1171103081246969856
Lateral Movement via WMI - Behavior

Log Name | Event ID
---------|---------
Security  | 4624
Sysmon   | 3

Authentication Logon Type (3)

Process: svchost.exe
Logon Session: Logon Type 3

Process: Wmiprvse.exe
Process: ?.exe
Process: ?.exe

https://twitter.com/HunterPlaybook/status/1171103081246969856
Lateral Movement via WMI - Behavior

Logon Session: Logon Type 3

- User authenticates
  - Logon Type (3)
- Logon Type 3
  - Process creates
    - svchost.exe
  - Process creates
    - Wmiprvse.exe
  - Process creates
    - ?.exe

<table>
<thead>
<tr>
<th>Log Name</th>
<th>Event ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security</td>
<td>4624</td>
</tr>
<tr>
<td>Sysmon</td>
<td>3</td>
</tr>
</tbody>
</table>

Log Name | Event ID
---------|----------
Security  | 4688     
Sysmon    | 1        

https://twitter.com/HunterPlaybook/status/1171103081246969856
Same Technique + Some Variations

Data produced

Test Security Controls

Model Adversary Behavior
Same Technique + Some Variations

- Wmic
- Invoke-WMIMethod
- SharpWMI
- Wmiexec

Data produced

Similar Events

Test Security Controls

Model Adversary Behavior
Same Technique + Some Variations

<table>
<thead>
<tr>
<th>Data produced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Name</td>
</tr>
<tr>
<td>Security</td>
</tr>
<tr>
<td>Sysmon</td>
</tr>
<tr>
<td>Security</td>
</tr>
<tr>
<td>Sysmon</td>
</tr>
</tbody>
</table>

Wmic

Invoke-WMIMethod

SharpWMI

Wmiexec

Test Security Controls

Model Adversary Behavior
We are all doing it already!
We are all generating similar events!
What if we share our datasets?
From Zero to Data Analytics Development!
Nothing wrong with that... but

Simulating Adversarial Technique

Test Security Controls

Model Adversary Behavior

TAKES TIME
I would like to download on-demand datasets!

AND EXPEDITE DEVELOPMENT OF ANALYTICS!
Enter Mordor
Mordor Project @Mordor_Project

- Pre-recorded security events generated by simulated adversarial techniques in the form of JavaScript Object Notation (JSON)
- Pre-recorded data categorized by platforms, adversary groups, tactics and techniques defined by the Mitre ATT&CK Framework.
- Data represents not only specific known malicious events but additional context/events that occur around it.

https://github.com/Cyb3rWard0g/mordor
Mordor Standard Environments

- Environment designed to replicate a small research network
- Standardized and documented setup
- Platforms
  - Windows
  - Linux
- Endpoints Telemetry
  - Windows Security Auditing
  - Event Tracing for Windows (ETW) (NEW!!)
- Network Telemetry
  - Network Logs
- Environments Available: Shire and Erebor

Mordor Environments: The Shire
The Shire Design
The Shire Telemetry: Win Logs & Sysmon

https://github.com/Cyb3rWard0g/OSSEM/tree/master/data_dictionaries/windows/sysmon
The Shire: Event Log -> WEC -> HELK

https://mordor.readthedocs.io/en/latest/mordor_shire.html#
Mordor Environments: Erebor (Lonely Mountain)
Erebor Telemetry: ETW Events via SilkETW

- Reads Config
- Enables ETW Providers
- Creates Trace Sessions
- Real-Time Delivery

SilkServiceConfig.xml

Microsoft-Windows-DotNETRuntime
Microsoft-Windows-DNS-Client
Microsoft-Windows-LDAP-Client
Microsoft-Windows-RPC

Event Trace Sessions

Real-Time Delivery

SilkService-Log

Erebor: ETW Events -> Event Log -> WEC -> HELK
How do you collect data?

- We use **Kafkacat**!
- kafkacat is a generic non-JVM producer and consumer for Apache Kafka >=0.8, think of it as a netcat for Kafka.
- **In consumer mode**
  - Kafkacat reads messages from a topic and prints them to standard output (stdout). You can also redirect it to a file (i.e. JSON)
- **In producer mode**
  - Kafkacat reads messages from standard input (stdin). You can also send data to kafkacat by adding data from a file.

https://github.com/edenhill/kafkacat
Consuming Data (Taking a snapshot of data)

$ kafka-cat -b <Kafka-IP>:9092 -t <kafka-Topic> -C -o end > file.json

- `-b` : Kafka broker
- `-t` : Topic to consume from
- `-C` : Consumer Mode
- `-o` : Offset to start consuming from

Consuming Data -> Creating Mordor File (Video)

Grunt: f7f6fb7ad2

Status: Active
CommType: HTTP
DotNetFrameworkVersion: Net35
UserDomainName:
UserName: powershell

Children: Nothing selected
ValidateCert: False
Integrity: Medium
Process: None
UseCertPinning: False

Producing Data (Injecting Adversary Dataset)
Producing Data (Injecting Adversary Dataset)

$ kafkacat -b <Kafka-IP>:9092 -t <kafka-Topic> -P -l file.json

- **-b**: Kafka broker
- **-t**: Topic to produce to
- **-P**: Producer Mode
- **-l**: Send messages from a file
I just want to download all the datasets..

$ git clone https://github.com/hunters-forge/mordor.git

$ cd mordor/small_datasets/

$ find . -type f -name "*.tar.gz" -print0 | sudo xargs -0 -I{} tar xf {} -C .
I would like to download on-demand datasets!
I would like to download on-demand datasets!

YOU CAN DO IT NOW!

https://github.com/hunters-forge/mordor
I already built analytics for all of Mordor Datasets!
I already built analytics for all of Mordor Datasets!

1. Define a Research Goal
2. Model Data
3. Simulate Adversary
4. Validate Detection Model
5. Define Detection Model
6. Document and Communicate Findings

https://www.youtube.com/watch?v=DuUF-zXUzPs
I already built analytics for all of Mordor Datasets!

Mordor File

Validate Analytics

2 + 2 = 4
I have data with me, now what?
Ready to Analyze Data!

1. Define a Research Goal
2. Model Data
3. Simulate Adversary
4. Define Detection Model
5. Validate Detection Model
6. Document and Communicate Findings

https://www.youtube.com/watch?v=DuUF-zXUzPs
Developing & Documenting Analytics

- Define a Research Goal
- Model Data
- Simulate Adversary
- Validate Detection Model
- Define Detection Model
- Document and Communicate Findings

https://www.youtube.com/watch?v=DuUF-zXUzPs
Beyond building boolean *(A AND (B OR C))* queries!

- Creative ways to query data
- Language Agnostic
- Database Agnostic
- Flexible to integrate with other analytic platforms!

- Standardized and organized way to hunt!
- Easy to replicate findings
- Easy to access and share with others
- Easy to automate
Enter Jupyter Notebooks
What are Jupyter Notebooks?

- Think of a notebook as a document that you can access via a web interface that allows you to save:
  - **Input** (live code)
  - **Output** (evaluated code output)
  - **Visualizations and narrative text** (Tell the story!)

- Uses include:
  - Data cleaning and transformation
  - Statistical modeling
  - Data visualization
  - Machine learning, and much more

https://jupyter.org/
IPython -> Jupyter Notebook

The Jupyter Notebook project is the evolution of the **IPython Notebook library** which was developed primarily to enhance the default **python interactive console** by enabling scientific operations and advanced data analytics capabilities.

![Jupyter Notebook screenshot](https://jupyter.org/)

---

https://jupyter.org/
How Do Jupyter Notebooks Work?

- Jupyter Notebooks work with what is called a **two-process model** based on a **kernel-client** infrastructure.
- This model applies **Read-Evaluate-Print Loop (REPL):**
  - Takes a single user’s inputs
  - Evaluates them
  - Returns the result to the user

https://jupyter.org/
**Jupyter Notebooks Architecture**

- **Client**
- **Jupyter Server**
- **Kernel**

- **WebSockets**
- **ZeroMQ**

- **Notebook File (json)**
- **Jupyter Document**

https://jupyter.org/
Try it at home! Or Right Now... (Video)

https://github.com/hunters-forge/notebooks-forge
cyb3rward0g clones $
cyb3rward0g clones $
cyb3rward0g clones $ git clone
Jupyter Notebooks Architecture

- Client
- Jupyter Server
- Kernel
- WebSockets
- ZeroMQ
- Notebook File (json)
- Jupyter Document

https://jupyter.org/
Practice Like You Hunt!

\[ y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it} \]
Practice Like You Hunt!

\[ y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it} \]

pandas

Spark
Spark SQL
Spark Streaming
MLlib (machine learning)
GraphX (graph)

Apache Spark
Analyzing Data as DataFrames

- DataFrame is a 2-dimensional labeled data structure with columns of potentially different types.
- You can think of it like a spreadsheet or SQL table.

https://pages.databricks.com/gentle-intro-spark.html
Can we take a look at this?
Mordor -> Jupyter Notebooks
OpenHunt Library

- Via PIP:
  `pip install openhunt`

- Or Straight from Source
  `git clone https://github.com/Cyb3rPanda/openhunt`
  `cd OpenHunt && pip install .`
Notebook Walkthrough - LM via WMI (Video)
Analysis of Lateral Movement via WMI Dataset

* Import Python Libraries and Define Classes

```python
from openhunt.logparser import winlogbeat
from openhunt.analytics import summaryStats
from pyspark.sql import SparkSession
```

```python
win = winlogbeat()
st = summaryStats()
```

* Initialize Spark session

```python
spark = SparkSession.builder.appName("Mordor").config("spark.sql.caseSensitive", "True").getOrCreate()
print(spark)
</pyspark.sql.session.SparkSession object at 0x7f8e86e84c18>
```

* Import Mordor data set

```python
mordor_file = win.extract_nested_fields("mordor/small_datasets/empire_wmic_add_user_2019-05-18231333.json", spark)

[+] Processing a Spark DataFrame..
[+] Reading Mordor file..
```
Notebook Walkthrough - APT3 Campaign (Video)

APT 3 Emulation Plan

Phase 1
- C2 Setup
- Software Packing
- Obfuscate Files
- Initial Access

Phase 2
- Compromise Host
  - Defense Evasion
  - Discovery
  - Privilege Escalation
- Credential Access
- Persistence
- Lateral Movement
- Execution

Phase 3
- Collect Data
- Compress and Stage
- Exfiltrate

https://attack.mitre.org/docs/APT3_Adversary_Emulation_Plan.pdf
APT3 Data Set Analysis

* Import Python Libraries and Define Classes

```python
from openhunter.logparser import winlogbeat
from openhunter.analytics import summaryStats, visualizations, dataManipulation
from pyspark.sql import SparkSession
from pyspark.sql.functions import col, concat
import pyspark.sql.functions as F
import warnings
warnings.filterwarnings('ignore')
%matplotlib inline
```

```python
win = winlogbeat()
stat = summaryStats()
vis = visualizations()
dm = dataManipulation()
```

* Initialize Spark session

```python
spark = SparkSession.builder.appName("Mordor").config("spark.sql.caseSensitive", "True").getOrCreate()
print(spark)
< pyspark.sql.session.SparkSession object at 0x7f1d19fd6ab38>
```
Would everyone understand my notebook?
The ThreatHunter-Playbook @HunterPlaybook

- A Threat hunter's playbook to aid the development of techniques and hypothesis for hunting campaigns by leveraging security event logs from diverse operating systems.
- It documents detection strategies in the form of interactive notebooks to provide an easy and flexible way to visualize the expected output and be able to run the analytics against pre-recorded mordor datasets.

https://github.com/hunters-forge/ThreatHunter-Playbook
Notebook Format Walkthrough (Video)
A Threat hunter’s playbook to aid the development of techniques and hypothesis for hunting campaigns.

- threat-hunting
- sysmon
- hunting-campaigns
- hypothesis
- hunting
- dfr
- hunter
- mitre-attack-db
- mitre

296 commits
2 branches
0 releases
10 contributors
GPL-3.0

Cyb3rWard0g Update README.md

library
Updated Remote Service Control Manager Handle
last month

playbooks
Updated location for DPAPI MasterKey
5 days ago

pre-hunt
Quick Update
2 months ago

resources
Docker Update & Playbook Format
last month

signatures/sigma
ThreatHunter Playbook 2.0
last month

.gitignore
Revamping Project
5 months ago

Dockerfile
Update Dockerfile
4 hours ago

LICENSE
Book Test & License Update
2 months ago

README.md
Update README.md
1 hour ago
Is You Ready? haha
Hunt The Planet!
Threat Hunters Forge Community!

Threat Hunters Forge
Data Science, Threat Hunting & Open Source Projects
Threat Hunters Forge Slack Community!

**ThreatHunting**
Threat Hunters Forge

Join the Threat Hunters Forge Slack Community!

A community led effort to share detection strategies and to support open source projects to aid the development of security analytics and tooling for threat hunting!

[https://launchpass.com/threathunting](https://launchpass.com/threathunting)
Remember this initiative with Mordor?
What if everyone gets a notebook too?
Wait, Whaaat?
The Binder Project

- The Binder Project is an open community that makes it possible to create shareable, interactive, reproducible environments.
- The main technical product that the community creates is called **BinderHub**, and one deployment of a BinderHub exists at [mybinder.org](http://mybinder.org).
- Who is it for?:
  - Researchers, Educators, people analyzing data and people trying to communicate the data analysis to others!!

BinderHub connects several services together to provide on-the-fly creation and registry of Docker images. It utilizes the following tools:

- **A cloud provider** such as Google Cloud, Microsoft Azure, Amazon EC2, and others
- **Kubernetes** to manage resources on the cloud
- **Helm** to configure and control Kubernetes
- **Docker** to use containers that standardize computing environments
- **A BinderHub UI** that users can access to specify Git repos they want built
- **BinderHub** to generate Docker images using the URL of a Git repository
- **A Docker registry** (such as gcr.io) that hosts container images
- **JupyterHub** to deploy temporary containers for users

Open Infrastructure for Open Hunts!

https://github.com/hunters-forge/ThreatHunter-Playbook
Open Infrastructure for Open Hunts! (LIVE!)

https://mybinder.org/v2/gh/hunters-forg/ThreatHunter-Playbook/master
Threat Hunter Playbooks via Binder (Video)

https://github.com/hunters-forge/ThreatHunter-Playbook
A Threat hunter’s playbook to aid the development of techniques and hypothesis for hunting campaigns.

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>library</td>
<td>Updated Remote Service Control Manager Handle</td>
</tr>
<tr>
<td>playbooks</td>
<td>Updated location for DPAPI MasterKey</td>
</tr>
<tr>
<td>pre-hunt</td>
<td>Quick Update</td>
</tr>
<tr>
<td>resources</td>
<td>Docker Update &amp; Playbook Format</td>
</tr>
<tr>
<td>signatures/sigma</td>
<td>ThreatHunter Playbook 2.0</td>
</tr>
<tr>
<td>.gitignore</td>
<td>Revamping Project</td>
</tr>
<tr>
<td>Dockerfile</td>
<td>Update Dockerfile</td>
</tr>
<tr>
<td>LICENSE</td>
<td>Book Test &amp; License Update</td>
</tr>
<tr>
<td>README.md</td>
<td>Update README.md</td>
</tr>
</tbody>
</table>

Latest commit 4f7a799 16 minutes ago
Goal: Share and Empower the Community!
Let’s do it together!
Threat Hunters Forge References

- GitHub: https://github.com/hunters-forge
- Python Library: https://github.com/Cyb3rPanda/openhunt
- Slack Invitation: https://launchpass.com/threathunting
- Founders: @Cyb3rWard0g & @Cyb3rPandaH
- Official Twitter: @HuntersForge
- @HunterPlaybook
- @THE_HELK
- @OSSEM_Project, @Mordor_Project & More
Thank You! Muchas Gracias!