Threat Hunting 101

Marissa Page

Whoami

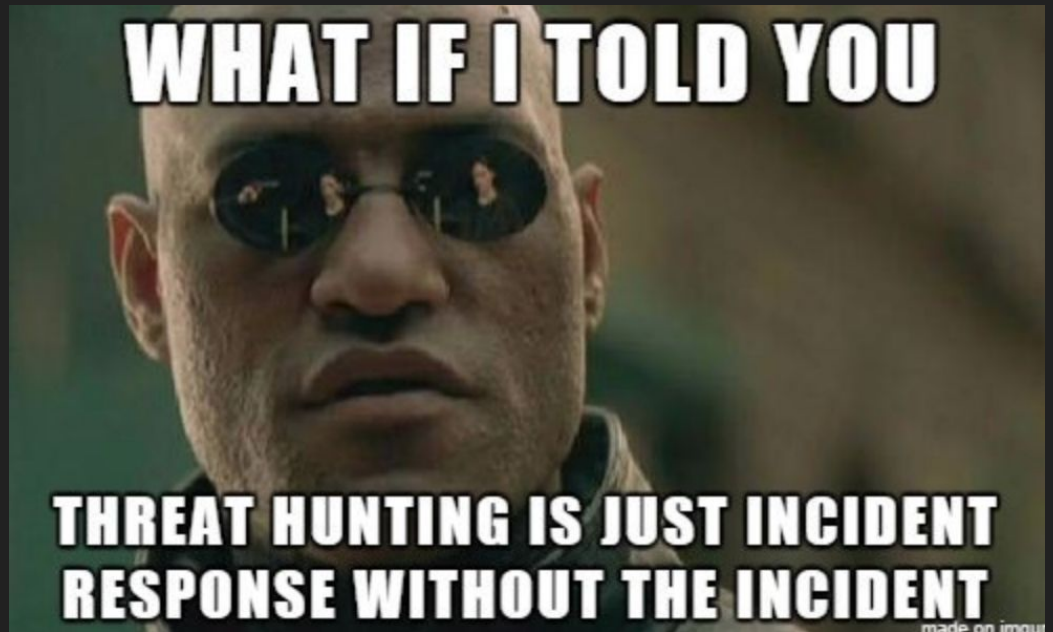
- Marissa Page
- 3.5 years of experience in Cybersecurity
- Experience in SOC analyst, Threat Hunting, and Malware Research
- President of the Columbus, GA ISSA Chapter
- Fun Fact - I performed in the 2016 Macy’s Thanksgiving Day Parade

Agenda

- General Overview
- How to Start a Successful Threat Hunting Program
- Threat Hunting Process

What is Threat Hunting?

- A proactive approach to identifying previously unknown, or ongoing non-remediated threats, within an organization’s network



Why is Threat Hunting Important?

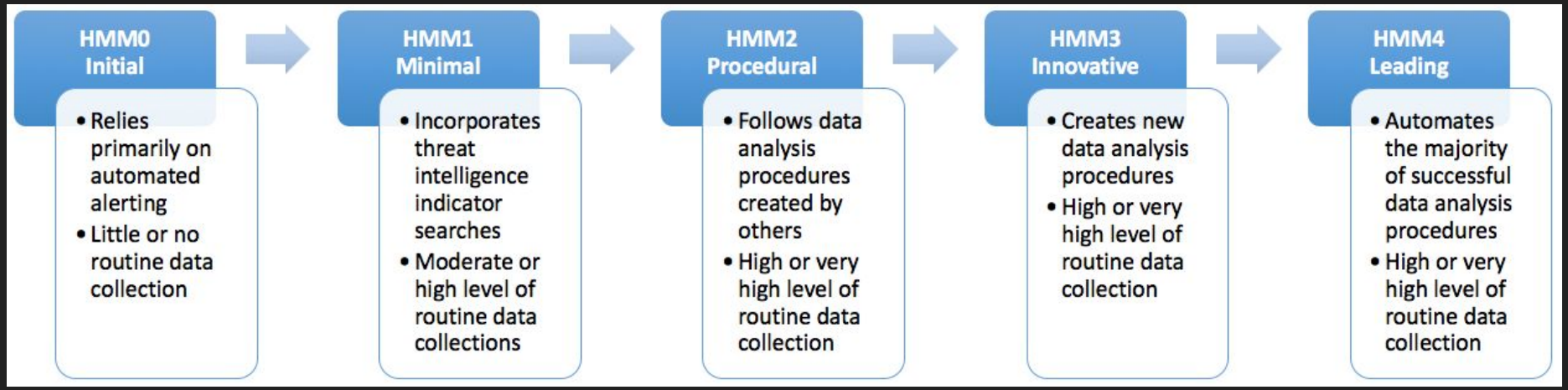
- 20% of threats are more likely to include sophisticated threats that can get passed automated cybersecurity & cause significant damage.
- On average threats can avoid detection for up to 280 days
- Effective threat hunting helps to reduce the time from intrusion to discovery

How Threat Hunting Works

- Cyber threat hunters bring a human element to security, complementing automated systems
- Threat hunters comb through security data, looking for:
  - Indicators of Compromise (IOCs)
  - Suspicious Processes

Types of Threat Hunting

- Structured Hunting
  - Based on an IoA and TTPs
  - Uses the MITRE ATT&CK framework, both PRE-ATT&CK and enterprise
- Unstructured Hunting
  - Initiated based on a trigger
- Situational or entity driven
  - Initiated by internal risk assessment or a trend and vulnerability analysis
  - these trends typically come from crowd-sourced attack data that reveal the latest TTPs



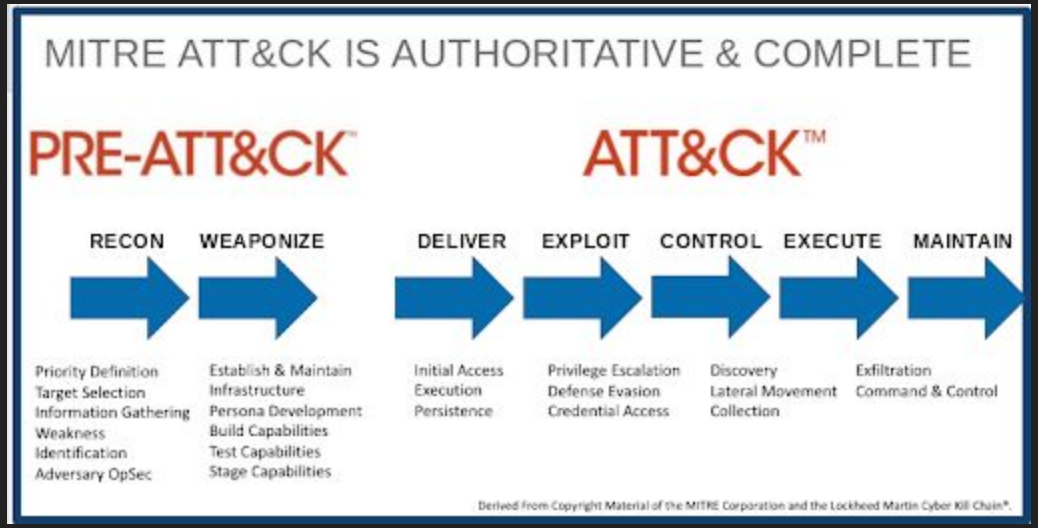
Hunt Maturity Levels Graph

Metrics for Measuring Your Hunting Success

- Number of Incidents by Severity
- Number of Compromised Hosts by Severity
- Dwell Time of any Incidents Discovered
- Number of detection gaps filled
- Logging Gaps Identified & Corrected
- Vulnerabilities Identified
- Insecure Practices Identified & Corrected
- Number of Hunts Transitioned to New Analytics
- False Positive Rate of Transitioned Hunts
- Any New Visibility Granted

Determining What to Hunt for & How Often

- Step 1: Choose Your Favorite Attack Model, for this presentation, we will be using the MITRE ATT&CK framework.
  - The kill chain will help you identify TTPs and attacker behaviors



Determining What to Hunt for & How Often

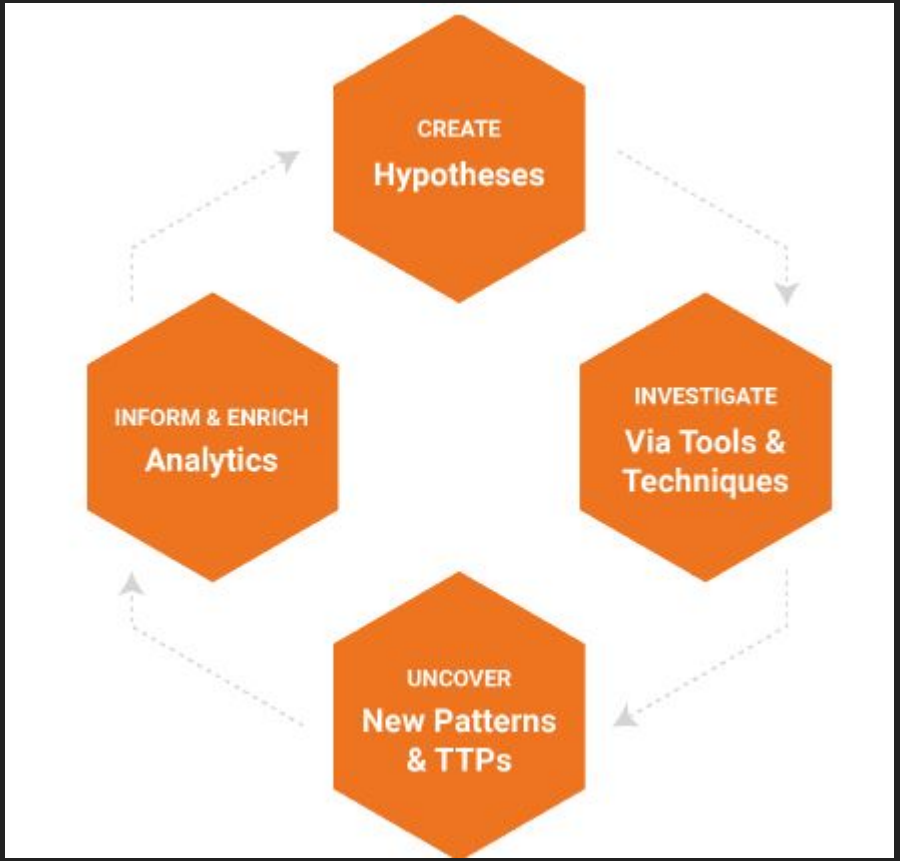
- Step 2: Identify Most Concerning Activities
  - Each phase in a model can include multiple categories of higher level tactics that an adversary might employ
  - An example list of potential attack activities & techniques:
    - Malware Beaconing
    - DLL Injection
    - Pass the Hash (PtH)
    - Shared Webroot
    - DNS Tunneling

**** Be sure to consider activities that are specific to your network environment & which assets you suspect an attacker would attempt to target ****

Determining What to Hunt for & How Often

- Step 3: Build Your Threat Hunting Calendar
  - Set a cadence for the frequency of your hunts
  - Organize each of the various phases by low, medium, and high impact activity
  - An example of weekly hunting sprints over two months:
    - Month 1:
      - Two weeks hunting High Impact Activity
      - Two week hunting Medium Impact Activity
    - Month 2:
      - Two weeks on High Impact Activity
      - One week on Medium Impact Activity
      - One week predicting attacks

Threat Hunting Loop



Four Primary Hunting Techniques - Searching

- The simplest method of hunting; querying data for specific results or artifacts
- Requires finely defined search criteria to prevent result overload.
- There are two primary factors to keep in mind when carrying out a search:
  - searching too broadly may produce far too many results
  - searching too specifically may produce few results

Four Primary Hunting Techniques - Clustering

- A statistical technique, often carried out with machine learning, that consists of separating groups based on certain characteristics
- Used for outlier detection: can accurately find aggregate behaviors
- Effective with a large data groups that do not explicitly share behavioral characteristics

Four Primary Hunting Techniques - Grouping

- Consists of taking a set of unique artifacts and identifying when they appear together based on specific criteria
- May potentially represent a tool of a TTP that an attacker might be using.
- Works best when you are hunting for multiple, related instances of unique artifacts

Four Primary Hunting Techniques - Stack Counting

- One of the most common techniques carried out by hunters
- Involves counting the number of occurrences for values of a particular type, analyzing the outliers
- Most effective with a thoughtfully filtered input
- Friendly intelligence can be used to defined filters for your input

Datasets for Hunting - Endpoint Data

Process execution metadata	Contains information on processes run on specific hosts. Critical metadata associated with process execution includes command-line commands/arguments and process filenames and ID.
Registry access data	Contains data related to registry objects, including key and value metadata
File data	Information on stored files and artifacts kept on a local host. This can include when files were create or modified, as well as size, type, and storage location information
Network data	Identification of the parent process for network connection
File prevalence	Information on how common a file is in your environment

Datasets for Hunting - Network Data

Network session data	Contains information on network connections between hosts. Critical metadata associated with network connections including the source IP address, destination IP address, destination port, start time of the connection, and end time/duration of the connection. This includes Netflow, IPFIX, and similar data sources.
Bro logs	A widely recommended network monitoring tool that collects connection-based flow data and application protocol metadata (HTTP, DNS, SMTP), specialized for security application
Proxy logs	HTTP data that contains information on outgoing web requests, including Internet resources that internal clients are accessing
DNS logs	Contains fata related to DNS domain resolution activity, including domain-to-IP address mappings and identification of internal clients making resolution requests
Firewall logs	Connection data that contains information on network traffic at the border of a network, focused on blocked connections
Switch and Router logs	Internal netflow, also known as east/west traffic, in your environment that shows what is going on inside the network behind your perimeter security

Datasets for Hunting - Security Data

Threat Intelligence	A broad category of information that includes the indicators and TTPS used by attackers, as well as the operations and campaigns they carry out
Alerts	The automated warnings or notifications created by correlation engine tools like a SIEM or IDS, indicating that a given rule set was violated or certain pattern identified, which might indicate a potential incident
Friendly Intelligence	Another broad category of information about an organization's own IT infrastructure, security ecosystems, critical assets, employee information, and business processes. Friendly intel helps hunters orient and understand the environment in which they are hunting and contextualize their investigation

High Impact Activity to Hunt For: Internal Reconnaissance

- Host Enumeration:
  - Determines local host details, establishes an understanding of local user context/ local host configuration.
- Network Enumeration:
  - Establishes which hosts are remotely accessible from the local host.

High Impact Activity to Hunt For: Persistence

- Scheduled Task Execution:
  - Process of queueing up programs or scripts that can be operationalized at a later point.
- DLL Injection:
  - Allows adversaries to hide malicious activity by incorporating it as part of a benign or routine process.
  - Allows attackers to access a system's process memory and permissions.
- Registry Modification:
  - RUN and RUNONCE registry key values allow for malware binaries to execute upon system boot and session login.
  - Common technique for persistence seen in the last decade.

High Impact Activity to Hunt For: Command & Control

- Common Protocol, Common Port:
  - This method seeks to hide traffic in plain sight by blending in with routine network traffic.
- Uncommon Protocol, Uncommon Port:
  - This technique bypasses heavily monitored ports and sends data through uncommon ports
  - Evades human operators and routine detection systems

High Impact Activity to Hunt For: Lateral Movement

- Pass the Hash (PtH)
  - Captures valid password hashes via a Credential Access technique
  - Bypasses user's cleartext password
- Remote Desktop Protocol
  - Allows access a computer's desktop interface
- Shared Webroot
  - Upload malicious content to internal website, execute using a web browser
  - Can result in the attacker gaining local or administrative privileges.
- Path Interception
  - Places an executable file in a specific path so that it is mistakenly run by a legitimate application.

High Impact Activity to Hunt For: Exfiltration

- DNS Tunneling
  - Allows the transfer of encoded data inside of DNS queries
  - Can bypass common security controls
- SFTP/SCP Exfiltration
  - Allows for usage of SSL to hide details about the traffic

Example Hunt - Hunting for Internal Reconnaissance

What are you looking for? (Hypothesis)	Hypothesis: An attacker conducting internal reconnaissance would attempt to carry our host enumeration and automate these commands with a script. Look for these commands to be spawned by a script: ● Whoami ● Net user ● Useraccount (WMIC) ● Get-NetIPConfiguration (Powershell) ● Hostname ● Ipconfig ● Nicconfig (WMIC)
Investigation (Tools and Data)	Determine what datasets you are using: ● Process execution metadata ● Process filenames ● Process file hashes
Uncover Patterns and IOCs (Techniques)	Using grouping, search for the above artifacts in process execution metadata. Specify that the commands should need to be executed within a given time frame. Doing this, you discover a previously unidentified script that contains commands to enumerate host information and saves the results in a unique file.
Inform and Enrich Analytics (Takeaways)	Taking the script and output files, you can now add those file names to your indication database and automated detection tool's watchlist. In this way, if the attacker continues to try and use this script on another host it will be detected automatically. The indicators can also be used to identify other previously compromised systems.

Example Hunt - Hunting for Command & Control

What are you looking for? (Hypothesis)	Hypothesis: Attackers may be operating on a C2 channel that uses custom encryption (uncommon protocol) on a common network port. Look for: ● Anomalies in monitored network port channels, i.e. connections that do not have protocol artifacts related to the common port you are looking at. ● For example, look for connections that have no identifiable HTTP metadata over port 80/TCP.
Investigation (Data)	Determine what datasets you are using: For identifying use of common protocols, you will want to focus primarily on application protocol metadata, including: ● Proxy logs, IIS logs ● DNS resolution logs ● Bro HTTP, SSL, DNS, SMTP logs
Uncover Patterns and IOCs (Techniques)	1. Use a search to identify legitimate protocol connections on a common port you will be inspecting, by looking at protocol metadata ■ If looking at port 80, search for any HTTP protocol records that exist for a given time period. 2. Use a second search to identify all network session metadata (e.g., Netflow, Firewall, etc.) on the common port for the same time period used in step 1. 3. Using the output of steps 1 and 2, remove the legitimate protocol connections from the session data, this should leave uncommon protocol connections on the common port 4. Take the results of step 3 and stack the data for what is useful to investigating your hypothesis ■ For example: destination IP, bytes transferred, connection duration/length, etc.
Inform and Enrich Analytics (Takeaways)	The destination IP address involved in the C2 activity you have discovered can be taken as IOCs and added to an indicator database in order to expand automated detection systems. You can also create packet-level signatures to trigger alerts for cases where the custom protocol you have discovered may appear again.

Top Considerations for Technology - Questions

Investigation Capabilities	1. *Which of the standard hunting techniques does the tool generally enable us to carry out?* 2. *How does the tool support the creation of hypotheses on which to base a hunt?* 3. *What ability does the tool have to import outside intelligence or custom indicators in order to assist analysts with the investigation of hypotheses?* 4. *What capabilities does the tool have that allow an analyst to pivot through different data sets?* 5. *How does the tool support the collection and storage of new indicators of Compromise that might be found over the course of a hunt?*
Analytics Supported	6. *What kind of analytics does the tool support that will help us facilitate more streamlined proactive investigation?* 7. *Does the tool enable the creation and customization of detection analytics?* 8. *Does the tool utilize any machine learning or data science techniques?*
Deployment & Data	9. *What data sources does the tool support?* 10. *To what extent is the tool able to scale its data storage capacity?* 11. *Through what process does it ingest or stream data?* 12. *What integrations with other security tools does it support?*