Accelerating Incident Response: Applying Confidence Aggregation to Defensive Artifacts

Author: Andrew Russett, andrew.j.russett@gmail.com

Advisor: Tanya Baccam

Accepted: *February 16th, 2024*

Abstract

Incident response plays an integral role in cybersecurity today. Despite the successes of the cybersecurity industry, the time between an intrusion and its detection remains relatively high. The following research is designed to combat the rising Mean Time to Identify (MTTI) security breaches and will explore the development of a tool to provide clarity. This tool coalesces an analytic assessment of the MITRE ATT&CK framework and provides confidence ratings for each tactic, technique, and procedure (TTP). This research will examine two threat groups through attack emulation, map the results to the MITRE ATT&CK framework, and provide a qualitative assessment of each TTP concerning the attack chain. This assessment aims to clarify the pivotal time between when a breach occurs and when it is identified. By applying a confidence statistic across each TTP, a visualization of the attack can provide context to each artifact discovered. When combined with a tested threat emulation model, this approach can increase the confidence a breach has occurred and prepares an incident response team earlier than traditional security operations would alone.

1. Introduction

Incident Response is a critical function of an organization’s cyber security operations. An incident response team is poised to act as firefighters during a cybersecurity incident, and they are tasked with containing and eradicating threats identified during a breach. This effect can directly influence an organization's final cost of a breach. Costs can be difficult to estimate as they come in many forms. Damage from ransomware, stolen data, loss of customer information, and damage to their reputation are just a few ways that breaches can affect an organization.

Before an incident response team can be effective, however, an organization must first identify that a breach has occurred. As of 2023, the average time to identify and contain a breach was 277 days, 204 days to identify the breach, and 73 days to contain and remove it from the network (IBM Security, 2023). This statistic has led to an average cost of $4.45m, a 15.3% increase since 2020 (IBM Security, 2023). This response time varies heavily depending on the organization's cyber security capabilities. Rising costs of security breaches and an increased time to identify and contain those breaches over the past few years have left more to be desired from cyber security teams worldwide.

1.1 History of Response Times

Today, the most significant hurdle for the incident response process is an organization’s ability to determine if a breach has occurred. The annual report from Mandiant indicates the average time that attackers will dwell within a network before deploying ransomware is merely 16 days on average; this is down from 21 days in 2022 (Mandiant, 2023). This metric brings to context the fact that over the last year, 67% of all ransomware-related attacks were not discovered until the ransomware notified the organization that their data had been encrypted (Mandiant, 2023).

There are many factors that determine the success of an organization's ability to detect and respond to a cybersecurity incident. Between March 2022 and March 2023, IBM Security and the Ponemon Institute conducted a study of 553 organizations affected by security breaches (IBM Security, 2023). Of the 27 factors they identified in the annual data breach report, having a “planned and tested incident response process” alone reduced losses by 34% or $1.49m when compared to the average (IBM Security, 2023). One of the issues facing the industry today is that the fires being fought by incident response teams are becoming increasingly more challenging to detect.

1.2 A Unified Approach

This research paper aims to design a tool to help cybersecurity personnel make quick decisions related to security events and ultimately reduce the time required to identify these threats. A Security Operations Center (SOC) has become a standard posture for many organizations and defines the process that teams use to provide cyber security. Regardless of the implementation, internal or external, a SOC provides a cumulative approach to defending an organization from external threats.

This study will examine the use of cyber threat emulation against a simulated network. The emulated attacks will be used to collect the artifacts necessary to build a more robust response tool. Two cyber-criminal organizations were chosen for emulation and pulled from the MITRE Engenuity Adversary Emulation Library. To further contextualize the results of the cyber threat emulation attacks, a heatmap of confidence ratings will be used to cross-reference collected TTPs (Tactics, Techniques, and Procedures) and their relation to each related TTP.

Generally, TTPs are identified and cataloged into a framework by MITRE and are often used to attribute individual events. However, to create a repeatable process, the confidence heatmap and the cyber threat emulation execution will be documented in a way through TTP analysis. This will allow a SOC to produce a tested and planned incident response process, improve its ability to contextualize individual detections, and reduce the time it takes to identify a breach.

2. Research Method

The MITRE ATT&CK framework is composed of 14 Tactics, 234 techniques, and 569 sub-techniques (*MITRE ATT&CK*® 2023). This structure is broken down into more manageable pieces when applying the filter of threat groups to the framework. No individual group uses every Tactic, Technique, or Sub-Technique and has its own signature that can be used for attribution, and as of 2023, there are a total of 143 threat groups being tracked by MITRE (*MITRE ATT&CK*® 2023). This study is built around the cyber threat emulation of two groups: Advanced Persistent Threat 29 and Wizard Spider.

2.1. Cyber Threat Emulation Scenarios

Cyber threat emulation is a concept of mimicking the attacks and behaviors of APTs to test how the defensive mechanisms within a defended network will react. To be effective in incident response, a SOC must understand what these attacks look like and how they interact with devices on the network it defends. One of the open-source tools that defenders can use is the Adversary Emulation Library designed and supported by MITRE Engenuity. These attack scenarios must be heavily customized as they emulate the toolset of each threat group within the library but not the actual tools themselves.

For the purposes of cyber threat emulation, the following simulated network was built to enable all aspects of a corporate network, including a third-party website that employees frequent. The network consists of two Windows machines, Windows 7 and 10, a router/firewall, an email server, and two domain controllers. The attacker's platform consists of multiple Linux machines, Kali and Redhat, and leverages the emulated tool sets for each cyber threat group.

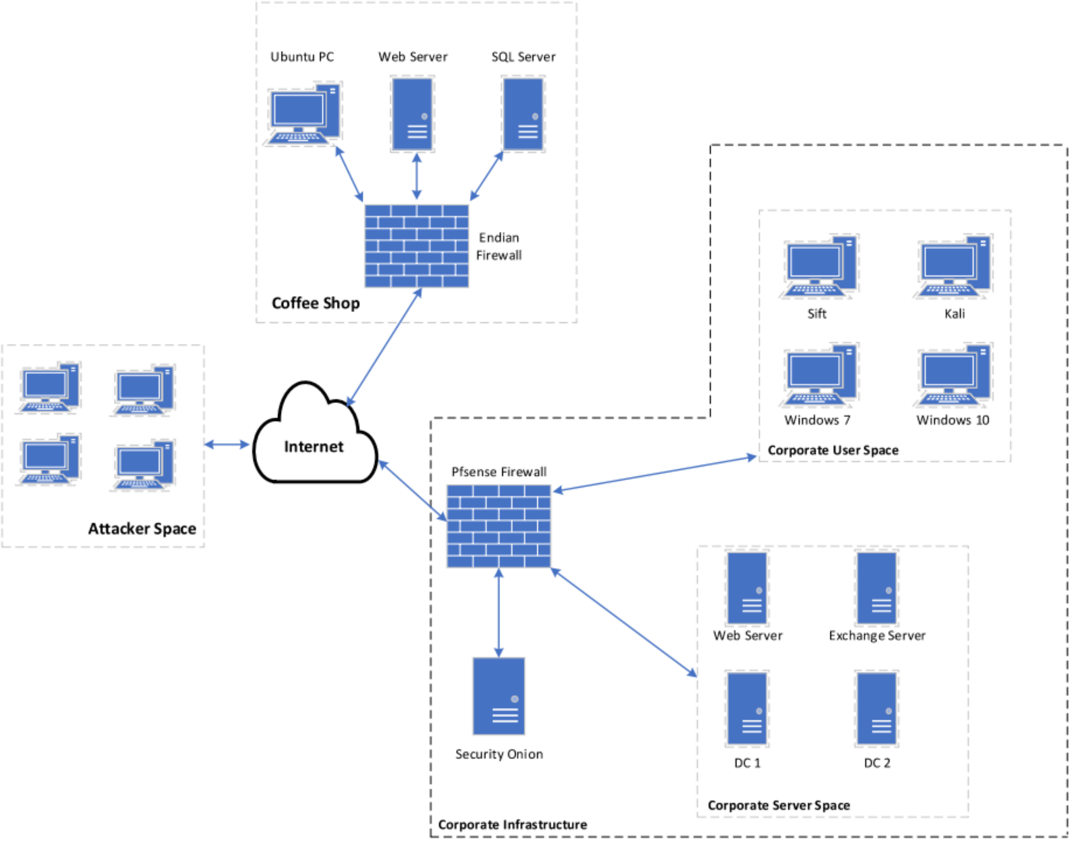


Figure 1. Simulated Network Diagram

2.1.1. APT 29

APT 29 is a threat group associated with the Russian Foreign Intelligence Service. Initially identified in 2008, this group likely targets foreign military and research institutions (*MITRE ATT&CK®* 2023). This APT was attributed to the SolarWinds Compromise and has a variety of malware they normally use to gain access to intelligence sources. The emulation scenario consists of two phases, starting with a watering hole attack to gain a remote shell into the victim's corporate network. The second phase consists of collection, lateral movement, and long-term persistence in the network.

The APT 29 attack begins with a malicious document embedded in a webpage that corporate users frequently visit. The victim will download the menu from the website and execute the malicious payload when opening the document. The malware will create a pupy reverse shell back to the command and control (C2) server the attacker controls. The attacker will then leverage this shell to perform collection and exfiltration of the infected host. A second-stage stealth toolkit is installed on the victim's host, allowing the attacker to abuse user access controls for elevated privileges.

The second phase of this attack begins with removing artifacts from the victim's host computer, establishing persistence with the new toolkit as a service, and stealing credentials from web browsers. Finally, the attacker will perform network discovery lateral movement with the SeaDuke malware and persist in the network indefinitely, collecting information.

2.1.2. Wizard Spider

Wizard Spider is a Russia-based threat group that targets major corporations and hospitals (*MITRE ATT&CK*®, n.d.). Initially identified in 2016, the Wizard Spider group is known for using malware like Emotet and TrickBot to perform attacks that ultimately end with the deployment of ransomware. The emulation scenario will include three phases: initial access, lateral movement, and deployment of the Ryuk ransomware package to encrypt the network.

The Wizard Spider attack scenario begins with the deployment of a phishing attack against a corporate admin. By leveraging a macro-enabled document, the malware pulls the Emotet client from the attacker's server after executing. At this time, the malware opens an SSH tunnel for a remote desktop backdoor, enables persistence through an autorun registry key, and scrapes credentials from the Outlook client on the victim's computer. Using the Emotet backdoor, the attacker can pull down TrickBot to perform network discovery and lateral movement before they steal AES hashes from the domain controller. Finally, the attacker deploys the Ryuk ransomware to the Domain Controller before the rest of the network for encryption.

3. Findings and Discussion

The scenario attacks were executed over the course of two months, starting November 2023 through January 2024. Artifact collection occurred periodically during and after the attacks using the assumption of a compromise and open-source tools. During the investigation, any artifacts identified were collapsed into signatures that could be used to identify the activity across each of the defense tools available to the defense team.

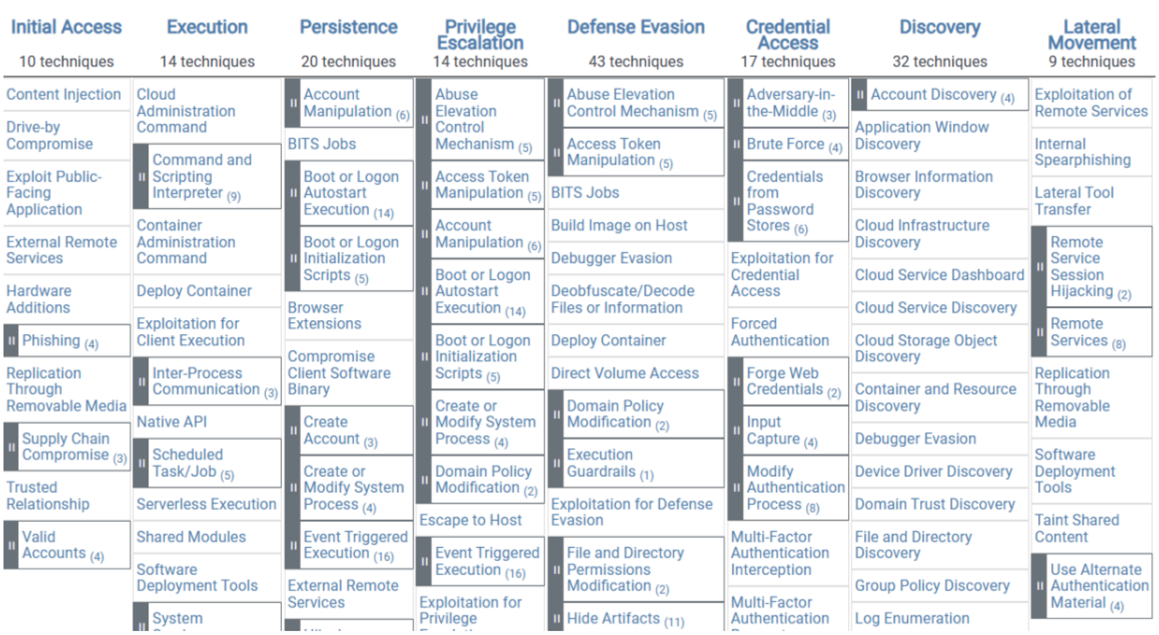
3.1. Artifact Collection and Analysis

The defense team primarily used four tools for the autonomous collection of artifacts and logs: Suricata, Sysmon, Windows Event Logs, and Zeek logs. Suricata and Zeek were captured off of a span port at the corporate firewall. Sysmon and Windows Event Logs were shipped via Beats to the Security Onion collector. Additionally, Pcap captures at the firewall were taken every 24 hours and assessed by the security team.

A threat-hunting team conducted periodic assessments of the corporate network each week and logged any changes on the hosts that could be considered suspicious or malicious with the defense team. These results were added to the daily defense reports and added to possible signatures.

3.1.1. The MITRE ATT&CK Framework

Artifacts identified by the defense team are structurally broken up into signature components based on known techniques from the MITRE ATT&CK Framework. Following Red Canary, this study is limiting the framework to the tactics that have the most impact on the successful outcome of an incident response. By focusing on the initial stages of a breach, between initial access and lateral movement, an organization can attempt to avoid the costly impacts such as exfiltration or ransomware encryption (Red Canary Threat Detection Report, 2023).



Initial Access 10 techniques	Execution 14 techniques	Persistence 20 techniques	Privilege Escalation 14 techniques	Defense Evasion 43 techniques	Credential Access 17 techniques	Discovery 32 techniques	Lateral Movement 9 techniques
Content Injection	Cloud Administration Command	Account Manipulation (6)	Abuse Elevation Control Mechanism (5)	Abuse Elevation Control Mechanism (5)	Adversary-in-the-Middle (3)	Account Discovery (4)	Exploitation of Remote Services
Drive-by Compromise	Command and Scripting Interpreter (9)	BITS Jobs	Access Token Manipulation (5)	Access Token Manipulation (5)	Brute Force (4)	Application Window Discovery	Internal Spearphishing
Exploit Public-Facing Application	Container Administration Command	Boot or Logon Autostart Execution (14)	Account Manipulation (6)	BITS Jobs	Credentials from Password Stores (6)	Browser Information Discovery	Lateral Tool Transfer
External Remote Services	Deploy Container	Boot or Logon Initialization Scripts (5)	Boot or Logon Autostart Execution (14)	Build Image on Host	Exploitation for Credential Access	Cloud Infrastructure Discovery	Remote Service Session Hijacking (2)
Hardware Additions	Exploitation for Client Execution	Browser Extensions	Boot or Logon Initialization Scripts (5)	Debugger Evasion	Forced Authentication	Cloud Service Dashboard	Remote Services (8)
Phishing (4)	Inter-Process Communication (3)	Compromise Client Software Binary	Create or Modify System Process (4)	Deobfuscate/Decode Files or Information	Forge Web Credentials (2)	Cloud Service Discovery	Replication Through Removable Media
Replication Through Removable Media	Native API	Create Account (3)	Domain Policy Modification (2)	Deploy Container	Input Capture (4)	Cloud Storage Object Discovery	Software Deployment Tools
Supply Chain Compromise (3)	Scheduled Task/Job (5)	Create or Modify System Process (4)	Escape to Host	Direct Volume Access	Modify Authentication Process (8)	Container and Resource Discovery	Taint Shared Content
Trusted Relationship	Serverless Execution	Event Triggered Execution (16)	Event Triggered Execution (16)	Domain Policy Modification (2)	Multi-Factor Authentication Interception	Debugger Evasion	Use Alternate Authentication Material (4)
Valid Accounts (4)	Shared Modules	External Remote Services	Exploitation for Privilege	Execution Guardrails (1)	Multi-Factor Authentication	Device Driver Discovery	
	Software Deployment Tools			Exploitation for Defense Evasion		Domain Trust Discovery	
	System			File and Directory Permissions Modification (2)		File and Directory Discovery	
				Hide Artifacts (11)		Group Policy Discovery	
						Log Enumeration	

Figure 2. MITRE ATT&CK Framework (*MITRE ATT&CK®* 2023)

Removal of the first two tactics, Reconnaissance, and Resource Development, stems from two separate factors. Reconnaissance tends to be saturated with thousands of standard scanning artifacts. Companies such as Google, Microsoft, or even Security Companies like Norton and McAfee commonly scan resources on the internet. Some do this for security reasons, others to update resources like the Google search engine or routing protocols. Sorting out malicious scanning from non-malicious scanning can be arduous and time-consuming with little to no benefit.

Resource Development on the other hand is a tactic that often leaves no artifacts for a security team to identify. The techniques generally consist of setting up tools, buying or securing infrastructure to launch their attacks from, and are easily burned or dissolved quickly when discovered. By eliminating these two tactics the confidence assessment can narrow its focus to tactics that impact an organization in real time.

Collection, Command and Control, Exfiltration, and Impact are tactics that indicate the completion of an attack chain. If an attacker has already entered this stage of a breach an Incident Response team has little room to maneuver and is essentially cleaning up the aftermath.

3.1.2. Tool Assessment

Each tool was assessed in three categories: Availability, Reliability, and Location. Each metric created a confidence rating in the tool that could be used to assess artifacts found by each tool. Tools with low internal visibility, such as Zeek and Suricata, were rated with a lower confidence rating for techniques where their visibility would be reduced or entirely absent. This metric also rated the threat hunting team lower due to availability.

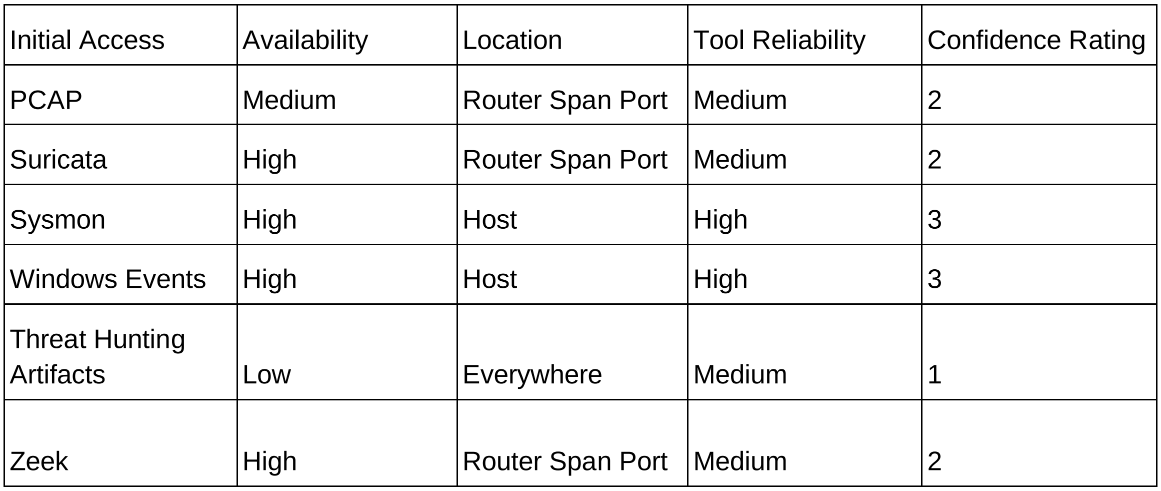
Availability is determined by two factors in this study. The first is how readily available the tools output is for the security defense team. The secondary factor is determined by the tool's availability at the time of execution. A good example of this is PCAP captures from tcpdump. While this tool is available after recording it cannot be analyzed until the following day after the capture has been completed. Comparing this to the Zeek logs that filter into Suricata every 20 minutes it has a lower availability score for real time assessments.

Reliability is also broken out into multiple factors. The first, and most pressing is how reliable to tool has been for the organization in the past. Issues such as implementation, compatibility, and even network issues factor into the tool's reliability statistic. A tool that is normally very reliable in the industry may not be reliable in an organization due to these factors, and each organization should perform these assessments on each of their tools periodically to ensure the tools are performing as expected.

For example, in this study, Suricata and pcap had a reliability issue due to their implementation with ESXI, a virtualization platform. A load balancing process would move virtual machines to hardware for increased efficiency, but when a sensor was not on the same hardware as the router, they would not capture traffic the same due to an underlying ESXI function. This load balance would effectively turn off promiscuous mode and prevent the sensor from receiving traffic that was not specifically sent to its IP address. This was solved by ensuring the virtual machines that required promiscuous mode for collection remained on the same hardware as the router span they collected from.

The final assessment category for each tool was the location where collection occurred. This affected the confidence rating of each tool based on the perspective the tool had within the environment. A higher confidence rating was given to host-based tools. Network tools on the router span port had a more limited perspective in testing and analysis compared to host-based tools like Sysmon. This was due to the implementation of the sensors. Additional sensors within the network that could be placed between vlans, for example, could have a higher confidence rating, but they were not used in these scenarios.

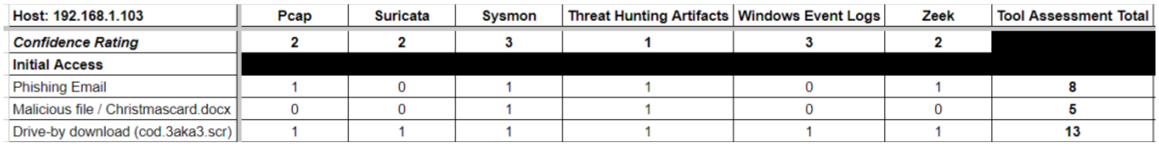
With a limited perspective, the network tools could not see information in encrypted protocols. They were unaware of activities within the network and could not be used to determine techniques used in tactics such as lateral movement, discovery, or credential access attacks against the domain controllers. This limited perspective reduces their overall confidence rating as a tool and can be used to identify gaps in security capabilities by the defense team.



Initial Access	Availability	Location	Tool Reliability	Confidence Rating
PCAP	Medium	Router Span Port	Medium	2
Suricata	High	Router Span Port	Medium	2
Sysmon	High	Host	High	3
Windows Events	High	Host	High	3
Threat Hunting Artifacts	Low	Everywhere	Medium	1
Zeek	High	Router Span Port	Medium	2

Figure 3. Defensive tool assessment

The combination of the Availability, Reliability, and Location provided a confidence multiplier that was used on signatures, logs, or artifacts found by that tool. An example of this would be the Phishing email signature that was captured by Sysmon. Sysmon has high availability, and reliability and is positioned internally in the network, giving it the highest score possible. Tools were assessed on a confidence scale of 1 to 3. The phishing artifacts identified in Sysmon logs would then receive a 3x multiplier before other confidence metrics were applied.



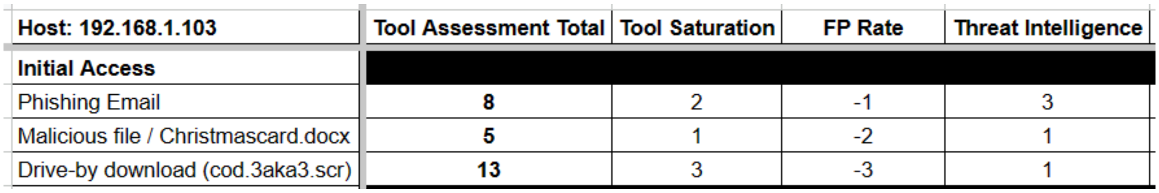
Host: 192.168.1.103	Pcap	Suricata	Sysmon	Threat Hunting Artifacts	Windows Event Logs	Zeek	Tool Assessment Total
Confidence Rating	**2**	**2**	**3**	**1**	**3**	**2**	
Initial Access							
Phishing Email	1	0	1	1	0	1	**8**
Malicious file / Christmascard.docx	0	0	1	1	0	0	**5**
Drive-by download (cod.3aka3.scr)	1	1	1	1	1	1	**13**

Figure 4. Confidence Matrix (Tool Assessment component)

Figure 4 has each confidence rating enriched into the cells below each tool. The 1 and 0 scores in each cell denote whether the tool had artifacts associated with the technique in its row (identified in the furthest left column). A score of 1 would indicate that artifacts were present in the tool in the top row of that column and the multiplier from Figure 3 would be used to determine the “Tool Assessment Total.” This column was the sum of the confidence ratings of all tools that had a score of 1 in that row. This metric would later be modified by the Tool Saturation metric but denotes the raw confidence scores at this stage of artifact analysis.

3.2. Confidence Matrix

Tool Assessment is one of the four metrics used as weights. These weights provide a confidence rating to be applied to the total number of artifacts found each day within the tactic categories. The confidence matrix is derived from four categories: Tool Assessment Total, Tool Saturation, False Positive Rate, and Threat Intelligence Scores.



Host: 192.168.1.103	Tool Assessment Total	Tool Saturation	FP Rate	Threat Intelligence
Initial Access				
Phishing Email	**8**	2	-1	3
Malicious file / Christmascard.docx	**5**	1	-2	1
Drive-by download (cod.3aka3.scr)	**13**	3	-3	1

Figure 5. Confidence Matrix Weights

While these weights may seem arbitrary, they are custom tuned to the environment, and when applied to the artifact totals, they provide context that an organization can use to implement their incident response teams sooner than when dealing with each artifact on its own.

3.2.1. Weights and their Value

Tool Saturation, False Positive Rate, and Threat Intelligence Scores are used to tune an organization's confidence matrix and provide an additional layer of context to the artifacts and detections within their defense toolkit. Tool saturation is a generally straightforward metric. The more tools that identify the same signature, the higher the confidence that the signature receives. Conversely, the False Positive Rate must be tuned to the organization.

Naturally, security teams must be aware of false-positive and false-negative detections. The cross-section of these two rates where both false positives and false negatives are equivalent is the Crossover Error Rate or CER. Many tools, like pcap, Zeek logs, and even Sysmon to a lesser extent, produce these false positives regularly. An example of this is identified in the drive-by download technique identified in Zeek logs. Of the hundreds of logs identified with a similar signature only a few were actual artifacts of the APT 29 attack.

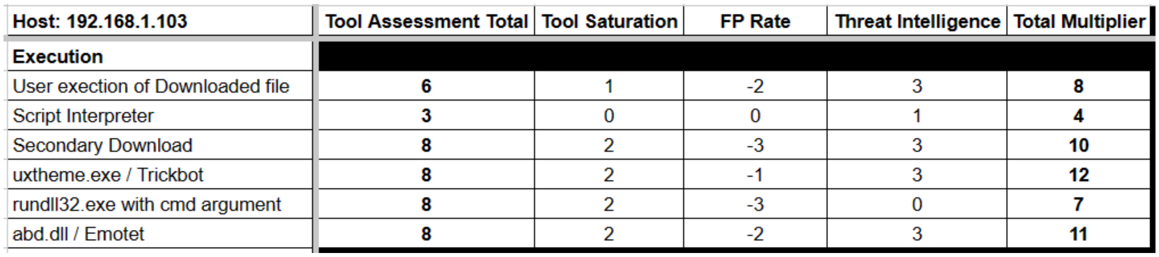
Due to the rate of false positives collected during both scenarios, each tool and technique must account for the possibility of logs, detections, or artifacts that can get caught up within a signature but are not part of the actual attack. This is also true of a false negative, an attack that should have been caught but, for some reason, did not fit into the signature. While these two concepts are critically important to a defense team, tool, or capability, they are less important in the confidence matrix.

The confidence matrix is primarily concerned with displaying the information that the security team has differently. However, as the matrix does directly deal with the possibility of the CER impacting results, we account for this within the matrix. Each signature must be compared to the tool and artifact type to determine the probability of false positives. The best example of this comes from the first scenario. The Wizard Spider emulation attack leverages the rundll32 executable to launch their Emotet client and begin command and control communications. rundll32, however, is a common executable, and if the signature is not precise, it can catch non-malicious occurrences and inflate the confidence heatmap with benign data.

False negatives on the other hand are a result of poor reliability within a tool or toolset and should be used by a security team to identify weaknesses in their data. For both scenarios there was very little defense evasion techniques caught by the security team or their sensors. This false negative rate heavily impacts a team's ability to identify threats quickly and should be monitored and evaluated regularly to improve the confidence matrix.

Finally, to measure each artifact, the application of a threat-hunting approach assumes a compromise has already occurred and leverages open-source intelligence to identify potential groups that have historically used the technique that is associated with the artifact that is found. This metric applies known techniques to the signatures within the matrix. The closer an artifact relates to a known technique, the higher the confidence score.

The confidence matrix for the emulation scenarios uses weights between -3 and 3. Combined with the tool assessment metric, the total confidence score becomes the final multiplier for each signature. Figure 6 below indicates how these metrics are applied to the total multiplier field.

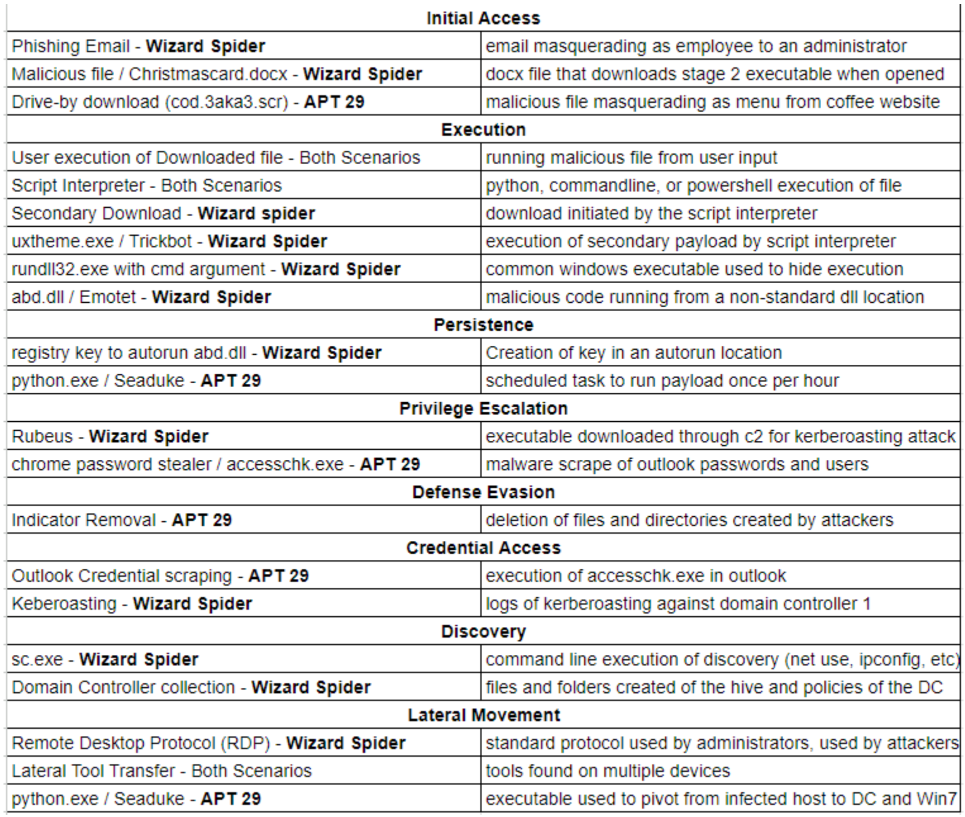


Host: 192.168.1.103	Tool Assessment Total	Tool Saturation	FP Rate	Threat Intelligence	Total Multiplier
Execution					
User exection of Downloaded file	**6**	1	-2	3	**8**
Script Interpreter	**3**	0	0	1	**4**
Secondary Download	**8**	2	-3	3	**10**
uxtheme.exe / Trickbot	**8**	2	-1	3	**12**
rundll32.exe with cmd argument	**8**	2	-3	0	**7**
abd.dll / Emotet	**8**	2	-2	3	**11**

Figure 6. Total Confidence Multiplier

3.2.2. Artifact Collection Results

A breakdown of the final metric calculations requires examining how each artifact contributes to the final score of the tactic it is associated with. For example, Scenario 1 execution began 20 November 2023 and completed on 22 December 2023. Artifact signatures were identified within the dataset by the security team, broken down into a useable threat signature, and labeled with the associated tactic.

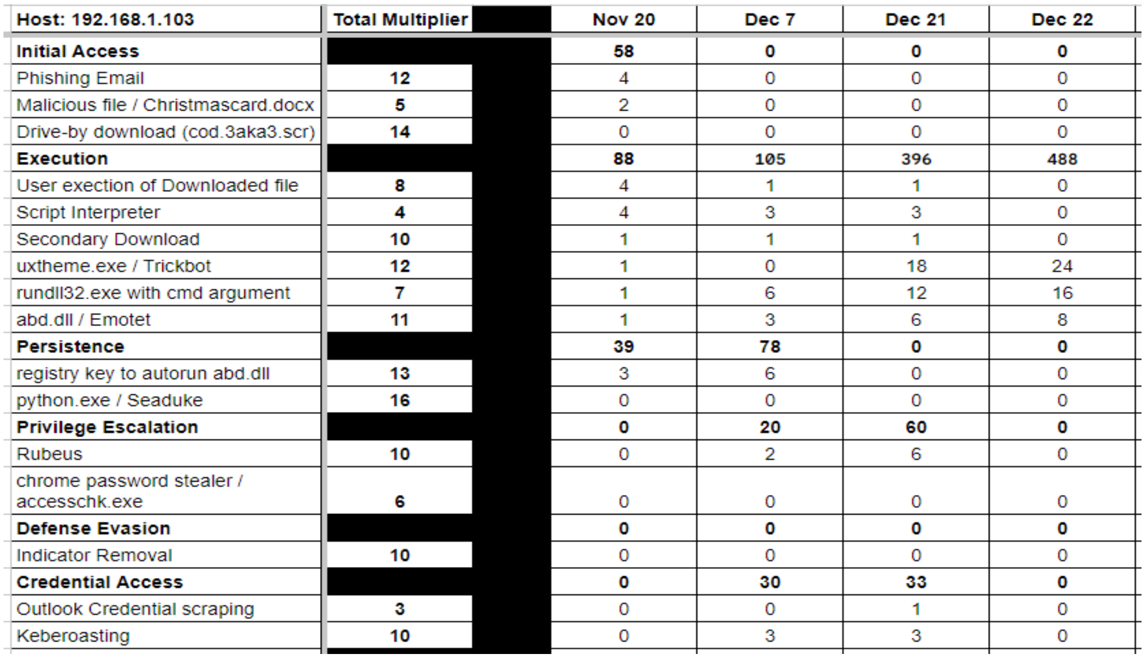


Initial Access	
Phishing Email - **Wizard Spider**	email masquerading as employee to an administrator
Malicious file / Christmascard.docx - **Wizard Spider**	docx file that downloads stage 2 executable when opened
Drive-by download (cod.3aka3.scr) - **APT 29**	malicious file masquerading as menu from coffee website
Execution	
User execution of Downloaded file - Both Scenarios	running malicious file from user input
Script Interpreter - Both Scenarios	python, commandline, or powershell execution of file
Secondary Download - **Wizard spider**	download initiated by the script interpreter
uxtheme.exe / Trickbot - **Wizard Spider**	execution of secondary payload by script interpreter
rundll32.exe with cmd argument - **Wizard Spider**	common windows executable used to hide execution
abd.dll / Emotet - **Wizard Spider**	malicious code running from a non-standard dll location
Persistence	
registry key to autorun abd.dll - **Wizard Spider**	Creation of key in an autorun location
python.exe / Seaduke - **APT 29**	scheduled task to run payload once per hour
Privilege Escalation	
Rubeus - **Wizard Spider**	executable downloaded through c2 for kerberoasting attack
chrome password stealer / accesschk.exe - **APT 29**	malware scrape of outlook passwords and users
Defense Evasion	
Indicator Removal - **APT 29**	deletion of files and directories created by attackers
Credential Access	
Outlook Credential scraping - **APT 29**	execution of accesschk.exe in outlook
Keberoasting - **Wizard Spider**	logs of kerberoasting against domain controller 1
Discovery	
sc.exe - **Wizard Spider**	command line execution of discovery (net use, ipconfig, etc)
Domain Controller collection - **Wizard Spider**	files and folders created of the hive and policies of the DC
Lateral Movement	
Remote Desktop Protocol (RDP) - **Wizard Spider**	standard protocol used by administrators, used by attackers
Lateral Tool Transfer - Both Scenarios	tools found on multiple devices
python.exe / Seaduke - **APT 29**	executable used to pivot from infected host to DC and Win7

Figure 7: Scenario Artifacts

Signatures were derived from Figure 7 by context clues within the artifacts that could be leveraged across the defense team's toolset. An example of this was the Sysmon logs for keberoasting. Host 192.168.1.103 was the initial access vector for the Wizard Spider scenario, its connection to 192.168.1.103 while already logged in was suspicious. The fact that it connected over port 88 with a host process of Rubeus.exe constituted an escalation of the artifact to potentially malicious. This signature was passed to the threat hunting team who were able to match the Windows event logs on the domain controller to the credential harvesting attack.

With signatures made for each artifact collection of each artifact, the day they occurred in the network, and the confidence metrics could be used to calculate the confidence matrix.



Host: 192.168.1.103	Total Multiplier	Nov 20	Dec 7	Dec 21	Dec 22
Initial Access		**58**	**0**	**0**	**0**
Phishing Email	**12**	4	0	0	0
Malicious file / Christmascard.docx	**5**	2	0	0	0
Drive-by download (cod.3aka3.scr)	**14**	0	0	0	0
Execution		**88**	**105**	**396**	**488**
User exection of Downloaded file	**8**	4	1	1	0
Script Interpreter	**4**	4	3	3	0
Secondary Download	**10**	1	1	1	0
uxtheme.exe / Trickbot	**12**	1	0	18	24
rundll32.exe with cmd argument	**7**	1	6	12	16
abd.dll / Emotet	**11**	1	3	6	8
Persistence		**39**	**78**	**0**	**0**
registry key to autorun abd.dll	**13**	3	6	0	0
python.exe / Seaduke	**16**	0	0	0	0
Privilege Escalation		**0**	**20**	**60**	**0**
Rubeus	**10**	0	2	6	0
chrome password stealer / accesschk.exe	**6**	0	0	0	0
Defense Evasion		**0**	**0**	**0**	**0**
Indicator Removal	**10**	0	0	0	0
Credential Access		**0**	**30**	**33**	**0**
Outlook Credential scraping	**3**	0	0	1	0
Keberoasting	**10**	0	3	3	0

Figure 8. Confidence Matrix Application

The figure above demonstrates how the final tactic totals are calculated. Initial Access on Nov 20th had a total score of 58, this was derived from the total number of artifacts multiplied by the confidence matrix total multiplier field for that signature (4*12 + 2*5+0*14). This calculation is done for every signature and recorded the day the signatures were logged. A total of 19 artifact signatures were identified between the two emulated scenarios (refer to Figure 7).

The table in Figure 8 was used to formulate the total scores for each tactic associated with the date the artifact occurred. This table was then formatted into a two-dimensional list before applying the results to a heatmap.

3.3. Confidence Heatmap

Taking the data collected by the defense teams from their automated collections, parsed logs, and active collections, the confidence matrix weights are then applied to the data as a weighted vector. To plot this data for visualization purposes the tables are organized in a two-dimensional array. Figure 9 is plotted using seaborn, matplotlib, and python, and creates a visualization of all data collected by date and the associated Tactic within the MITRE ATT&CK Framework.

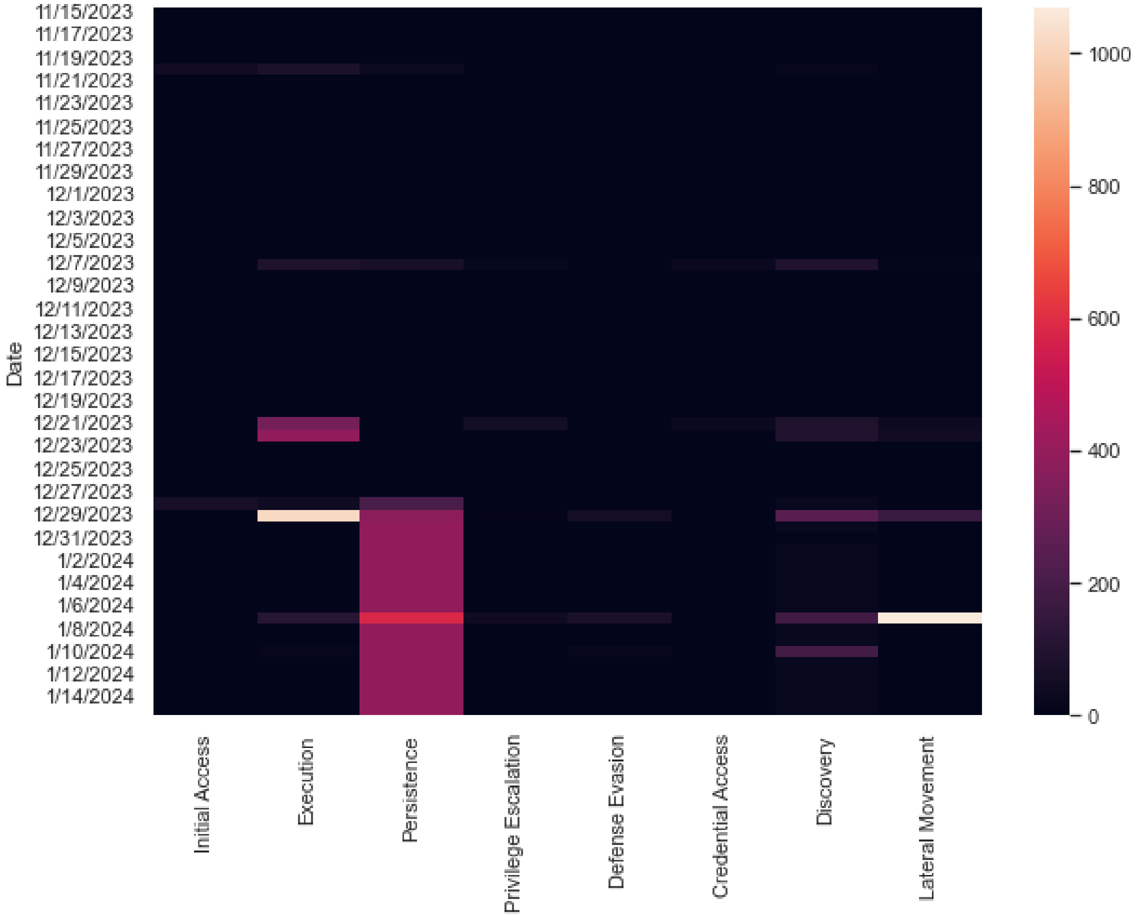


Figure 9. Quarterly Confidence Heatmap Nov 2023 - Jan 2024

At first glance it is apparent that some tactics are prone to produce a larger response. This is indicative of techniques that generate a larger signature and produce more defensive artifacts. It also highlights the nature of the two attacks. The minimal footprint in the Initial Access column indicates that both attacks probably originated internal to the network and that no brute force attacks were utilized.

Additionally, the visualization in Figure 9 can be used to demonstrate what the tools the organization is using are strong at detecting, and what detections or tools might need additional funding or support. This was most striking in the defense evasion column as both attacks had some form of defense evasion, but very little to none of those techniques were identified by the defensive toolset. There were six specific clusters of activity that were identified after the Wizard Spider and APT 29 attacks.

- November 20, 2023 – an administrator logs into their email with a privileged account and downloads a Christmas card email they though was from a co-worker. This executable, when opened downloaded malware from the Wizard Spider control server and gave them access to the host. They quickly established persistence.
- December 7, 2023 – after a computer reset the persistence registry key executes the dll downloaded on November 20th. The callout connects to a listener and allows Wizard Spider to perform discovery and exfiltration of the infected host.
- December 21-22, 2023 – Another computer reset establishes a session the attackers use to move laterally through the network. Credentials that were stolen via kerberoasting on Dec 7th have been decrypted and allow the attackers to gain access to the domain controller before they begin trying to use the Ryuk malware to encrypt hosts and servers throughout the organization. Fortunately, this part of the attack was prevented.
- December 28-29, 2023 – a user accesses the website of a nearby coffee shop, they download a menu before ordering something to go. The menu was compromised with a malware executable that immediately connected to an APT 29 command and control server. The attackers immediately begin to install persistence in the form of a scheduled task, before they

begin to perform discovery of the host and network they now have access to.

- January 7, 2024 – After identifying the company the network belonged to APT 29 catches one of the hourly callouts from the scheduled task. They begin to perform lateral movement through the SeaDuke malware that masquerades as python.exe. Establishing persistence throughout the network.
- January 10, 2024 – APT 29 begins to regularly exfiltrate data from the network, the only footprint is the hourly scheduled task and the amount of data leaving the network.

These two scenarios represent two threat groups with very different goals, methods, and intent. Though the vast difference in attack patterns has also caused the results of the second emulation to overshadow the first, this can be remedied by isolating techniques into two individual heatmaps.

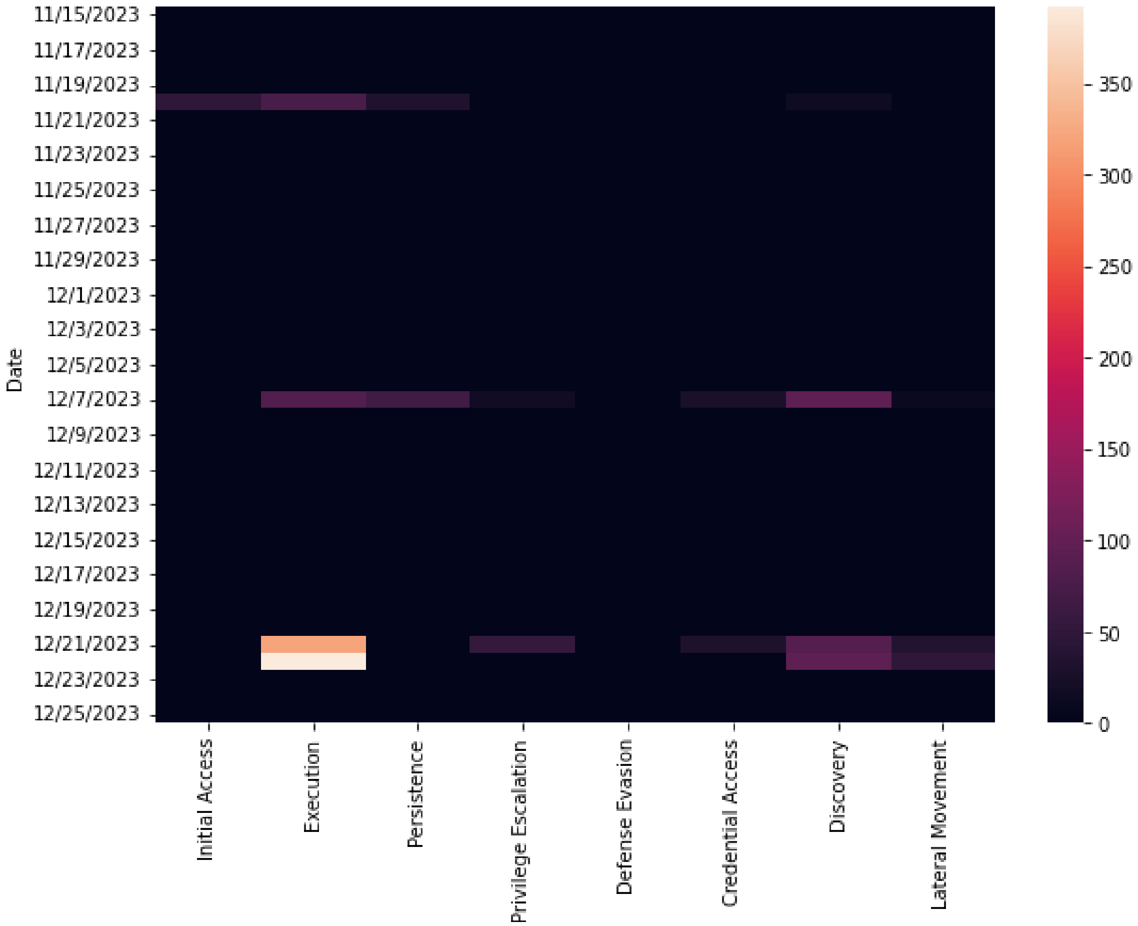


Figure 10. Scenario 1 - Wizard Spider Confidence Heatmap

After isolating the techniques most commonly associated with the Wizard Spider attack the heatmap in Figure 10 provides a much clearer picture of the three phases the attack was emulated in. The first phase of gaining and maintaining access in late November, the persistence and discovery phase in early December, and finally the final attack phase where the attacker moved additional tools into the environment and spread out through the network to begin leveraging the Ryuk ransomware tool.

The Wizard Spider attack is a quiet spear phishing attack with clear goals. This caused there to be fewer artifacts present in the network before the execution of the ransomware tool encrypting devices throughout the network. In contrast, Figure 11 below is a heatmap of the APT 29 attack. Comparing these attacks demonstrates the differences in intent, and how the attackers gained access to the corporate network.

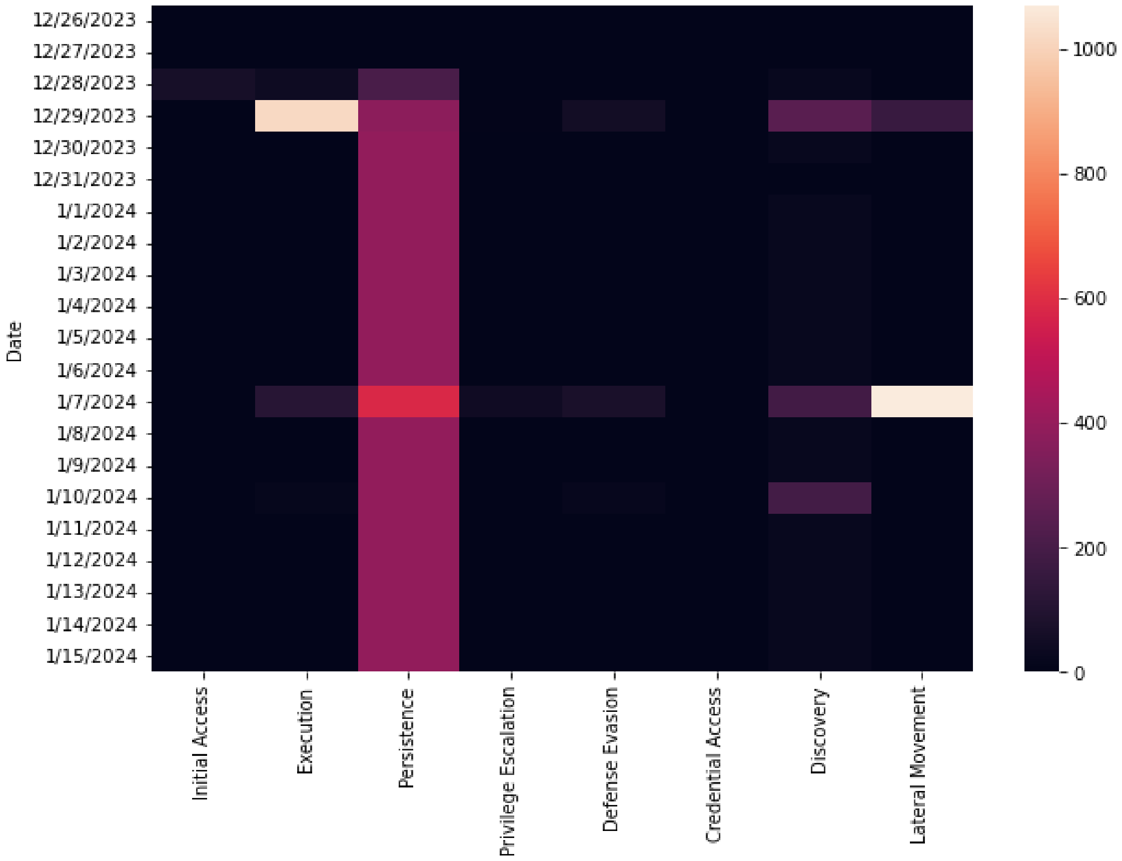


Figure 11. Scenario 2 - APT 29 Confidence Heatmap

APT 29 was able to gain access to the corporate network via drive-by compromise. Due to an earlier compromise of a public coffee shop's website the malware was downloaded and executed by a corporate administrator giving the attacker access to a network they had not anticipated. The initial attack is loud, with multiple layers of attacks occurring very quickly to determine where the malware had been executed and what the attacker could gain access to. This can be seen in Figure 11 in the upper left quadrant and is relevant as this is the initial compromise.

The use of scheduled tasks to maintain persistence was significant in this scenario. Though the scheduled task only executed every hour, the signature and tools associated with artifact collection highlighted those artifacts in the heatmap. This is something that could easily be lost in a normal alert based visualization but is easily visible within the confidence heatmap. Without the confidence matrix, this attack could be very easy to miss.

4. Recommendations and Implications

Applying a confidence matrix as a weight against collected artifacts can provide important clarity in the metrics an organization collects regularly. Additionally, applying that filter to a visualization like a heatmap can assist defenders in making decisions when it matters most. One challenge in emulating the attacks in this study has been the depth of the defensive network.

In a large organization the heatmaps can be further divided by subnet, VLAN, or even each host if the network is regularly under attack in some spaces such as the DMZ. This would also allow an organization to keep an eye on critical assets to ensure that if a breach does occur, they could apply the signatures across critical assets to determine if they were at risk.

4.1. Recommendations for Practice

During this study each artifact was manually collected from the security onion sensor, however in a live environment it would be beneficial to directly pull data from the elastic stack, or even building the heatmap and confidence matrix into their SIEM to provide real time analytics. While this study is primarily concerned with visualizing tactics that would give an Incident Response Team the most time to be effective, organizations can also apply these confidence analytics against the entire MITRE ATT&CK Framework.

Any organization looking to implement this tool should identify what each analytic within the confidence matrix means to their internal structure. The organization should have a baseline of their network, an understanding of what their defensive posture currently looks like, and a defined goal for implementing this tool.

4.2. Implications for Future Research

No one tool or technique will overcome the difference between average dwell time and the average Mean-Time-To-Identify organizations face daily around the world. Even the most effective artificial intelligence tools only reduce the MTTI by 40% (IBM Security, 2023). This study attempts to introduce a new tool that organizations can use to further improve their response time to security breaches, but realistically it must be implemented with other tools and strategies.

Future Research topics associated with confidence metrics could investigate methods for automating collection of artifacts and application of the vector weights from the confidence matrix to those artifacts. The option of turning this tool into a plugin for common tools like elastic could also be explored.

Exploration of new confidence weights could also be considered. Concepts like total number of hosts with the same artifacts discovered, or even relational time between artifacts to determine if the artifact was generated by a script or a human could be beneficial to many organizations. Each of these confidence weights could also be applied to a machine learning framework as they are meant to be adjusted by the organization that implements them.

Finally, this concept needs to be applied to a live network. Applying confidence ratings to artifacts in a simulated environment can build towards theory of execution, but only data from a live environment where false positives and steady baselines exist will the tools presented in this study find their footing.

5. Conclusion

This study introduces a new concept of applying confidence ratings to individual artifacts to improve the time between a breach occurring and when it is identified. By organizing the data an organization already possesses, and rearranging the data, the confidence heatmap aims at providing a different perspective. This tool is designed to be very adaptable and, based upon the needs of the organization, can be skewed to highlight subtle differences between multiple endpoints, VLANs, or even subnets.

Two threat groups were identified and selected for cyber threat emulation to test the concepts behind this tool. Findings indicate that some techniques and tactics are far more visible with this tool than they would be from a simple artifact collection perspective. By focusing the weights of each confidence analytic, an organization can find and track techniques they are most concerned with.

There are several challenges that remain to be fleshed out, including automation and integration of the tool with common tools such as elastic. Some recommendations to address these challenges are critical for implementation and future research into this concept.

As adversaries across the domain continue to refine their capabilities and reduce their dwell time, so too must defensive tools and capabilities adapt to these emergent threats. The rising cost of security breaches cannot be ignored, and organizations must find novel ways to improve their capabilities of identifying these threats before the critical line between initial access and lateral movement. Once an attacker gains access to the network at that depth, they begin to take actions that leave an impact.

This research is the first step in creating a more robust capability in identifying these threats earlier. Additional tools and capabilities need to be integrated into a security toolkit either universal in its approach or custom tailored for an organization as unique as the critical data and infrastructure they are trying to protect. Using confidence assessments stands as a strong tool aimed at reducing the MTTI that organizations face each year. Providing context to data could significantly reduce response times and improve the incident response process.

References

Ahmad, A., Hadgkiss, J., & Ruighaver, A. B. (2012). Incident response teams – Challenges in supporting the organisational security function. *Computers & Security*, *31*(5), 643–652. https://doi.org/10.1016/j.cose.2012.04.001

Ajmal, A. B., Khan, S., Alam, M., Mehbodniya, A., Webber, J., & Waheed, A. *(2023). Toward Effective evaluation of Cyber Defense: Threat Based Adversary Emulation approach. IEEE Access, 11, 70443–70458. https://doi.org/10.1109/access.2023.3272629*

Best Practices for MITRE ATT&CK® Mapping | CISA. (2023, January 17). Cybersecurity and Infrastructure Security Agency CISA. https://www.cisa.gov/news-events/news/best-practices-mitre-attckr-mapping

Center-For-Threat-Informed-Defense. (n.d.). *GitHub - center-for-threat-informed-defense/adversary_emulation_library: An open library of adversary emulation plans designed to empower organizations to test their defenses based on real-world TTPs.* GitHub. https://github.com/center-for-threat-informed-defense/adversary_emulation_library/tree/master

IBM Security. (2023). Cost of Data Breach Report. *ibm.com*. Retrieved November 1, 2023, from https://www.ibm.com/account/reg/us-en/signup?formid=urx-52258

Mandiant. (2023). Mandiant M-Trends 2023. *https://www.mandiant.com/resources/reports*. Retrieved November 1, 20231, from https://www.mandiant.com/resources/reports/get-your-copy-m-trends-2023-today

MITRE ATT&CK®. (n.d.). https://attack.mitre.org/

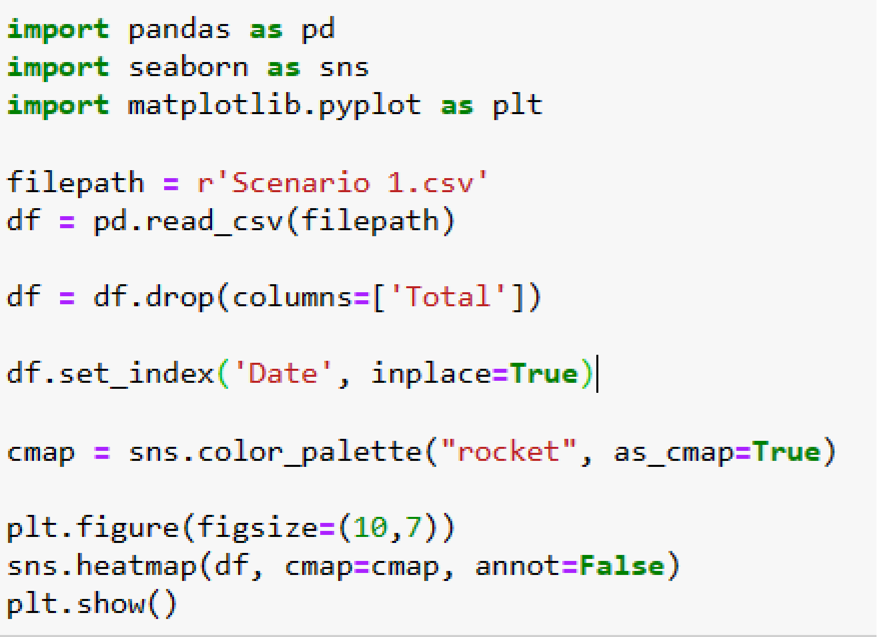
Poley, S. (2023). The Cyber Data Paradox: Storing Less, Discovering More. *Sans Cyber Security Research Papers*. https://www.sans.edu/cyber-research/cyber-data-paradox-storing-less-discovering-more/

Red Canary Threat Detection Report. (2023, April 22). Red Canary. https://redcanary.com/resources/guides/threat-detection-report/

Vielberth, M., Böhm, F., Fichtinger, I., & Pernul, G. (2020). Security Operations Center: A systematic study and open challenges. *IEEE Access*, *8*, 227756–227779. https://doi.org/10.1109/access.2020.3045514

Appendix

Python Heatmap Script



```python
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

filepath = r'Scenario 1.csv'
df = pd.read_csv(filepath)

df = df.drop(columns=['Total'])

df.set_index('Date', inplace=True)

cmap = sns.color_palette("rocket", as_cmap=True)

plt.figure(figsize=(10,7))
sns.heatmap(df, cmap=cmap, annot=False)
plt.show()
```

Wizard Spider Intelligence Summary

Associated Country: Russia
First Identification: 2016
Associated Group Names: UNC1878, TEMP.MixMaster, Grip Spider, FIN12, GOLD BLACKBURN, ITG23, Periwinkle Tempest
Description:
Wizard Spider is a crime group associated with the TrickBot malware. Described as financially motivated, the group often targets large organizations with the intent of conducting financial fraud. Primary attacks involve a ransomware mechanism such as Ryuk, Conti, and Bazar. The crime group has been labeled as a Russian state-sponsored cyber actor by the Cybersecurity & Infrastructure Security Agency (CISA).
Associated Software:

- AdFind
- Bazar
- BITSAdmin
- BloodHound
- Cobalt Strike
- Conti
- Dyre
- Emotet
- Empire
- GrimAgent
- LaZagne
- Mimikatz
- Net
- Nltest
- Ping
- PsExec
- Rubeus
- Ryuk
- TrickBot

Intelligence References:
https://attack.mitre.org/groups/G0102/
https://www.cisa.gov/news-events/cybersecurity-advisories/aa22-110a
https://www.crowdstrike.com/blog/wizard-spider-adversary-update/
https://www.theregister.com/2022/05/18/wizard-spider-ransomware-conti/
https://www.trendmicro.com/en_us/research/19/b/trickbot-adds-remote-application-credential-grabbing-capabilities-to-its-repertoire.html
https://intel471.com/blog/ransomware-as-a-service-2020-ryuk-maze-revil-egregor-doppelpaymer

APT 29 Intelligence Summary

Associated Country: Russia
First Identification: 2008
Associated Group Names: YTTRIUM, The Dukes, Cozy Bear, CozyDuke, IRON RITUAL, Dark Halo, SolarStorm, UNC3524, UNC2452
Description:
APT 29, formally named in a white house release in 2021, is associated with the Russian Foreign Intelligence Service (SVR) due to their pervasive espionage campaign and its connection to the SolarWinds attack. The group has a wide scope targeting NATO-aligned countries and governments. Their primary goal appears to be data theft and continuous monitoring.
Associated Software:

- AADInternals
- AdFind
- BloodHound
- BoomBox
- CloudDuke
- Cosmic Duke
- Cobalt Strike
- CozyCar
- Gemini Duke
- HAMMERTOSS
- Meek
- mimikatz
- MiniDuke
- OnionDUke
- PinchDuke
- POSHSPY
- PowerDuke
- PsExec
- SDelete
- SeaDuke
- SUNBURST
- Tor

Intelligence References:
https://www.mandiant.com/resources/blog/dissecting-one-ofap
https://attack.mitre.org/groups/G0016/
https://www.whitehouse.gov/?s=Fact+Sheet%3A+imposing+costs
https://www.mandiant.com/resources/blog/unc2452-merged-into-apt29\