MITRE | ATT&CK® Evaluations

USING RESULTS TO EVALUATE ENDPOINT DETECTION PRODUCTS

Table of Contents

Background

MITRE evaluates cybersecurity products using an open methodology based on MITRE ATT&CK®. ATT&CK is a globally-accessible knowledge base of adversary tactics and techniques based on real-world observations. ATT&CK serves as a common language between vendors and their customers so they can better understand how the vendors' products address malicious behaviors. To gather product specific insights, we use ATT&CK as a foundation and build an emulation plan to test products "in the style of" a specific adversary. The emulation plans are sourced with public cyber-threat intelligence reporting, mapped to a subset of ATT&CK techniques, and used to replicate the behaviors that generate objective insights into how well products perform.

Our goals are to:

- Empower end-users with objective insights into how to use specific commercial security products to detect known adversary behaviors
- Provide transparency around the true capabilities of security products and services to detect known adversary behaviors
- Drive the security vendor community to enhance their capability to detect known adversary behaviors

These evaluations are not a competitive analysis. We show the detections we observed without providing a "winner." There are no scores, rankings, or ratings. Instead, we show how each vendor approaches threat detection through the language and structure of ATT&CK.

So how does this help you? The evaluation results give you a basis to perform processing and analysis to find insights useful to your organization.

This guide helps you understand how to use the evaluation results to assess security products and select an endpoint threat detection tool.

Let's get started.

Understanding Your Needs

Before diving into the results, it's necessary to understand the needs of your organization and how ATT&CK Evaluations can help.

Primary criteria include:

- **Does this tool detect known threats to your organization (i.e., ATT&CK technique coverage)?**
- **How does the tool present the data to your analysts (i.e., Graphical User Interface [GUI])?**
- How much does the tool cost?
- How does the tool integrate with your other tools?

ATT&CK Evaluations can help with the first two!

Let's break down those first two questions to determine what you need ATT&CK Evaluations to help answer.

- **Does this tool detect the threats targeting your organization?**
  - What vendors provide the most visibility across adversary techniques?
  - What vendors best address the techniques that the threats use?
  - How many false positives will you have to deal with?
  - How often is the tool updated to cover new techniques?
- **How does the tool present the data to your analysts (i.e., GUI)?**
  - Are you going to use the GUI, or are you just interested in the raw events to feed a Security Information Event and Management (SIEM) or orchestration tool?
  - What is the skill-level of the analysts who'll be using this tool?
  - Do you need a turn-key solution for less experienced analysts?
  - Do you need features to allow experienced analysts to hunt through raw data and create their own detections?

Before starting any analysis of technique coverage, it's important to think about what techniques are most relevant to your organization, based on the adversary groups and threats your organization tends to face.

This answer is outside the scope of ATT&CK Evaluations. You could use a variety of sources to get this information, whether the ATT&CK framework, public/commercial threat reporting and analysis, or your own intel to determine this subset of techniques.

Once you have this list, you can start analyzing the results of those techniques within ATT&CK Evaluations.

Addressing Your Needs

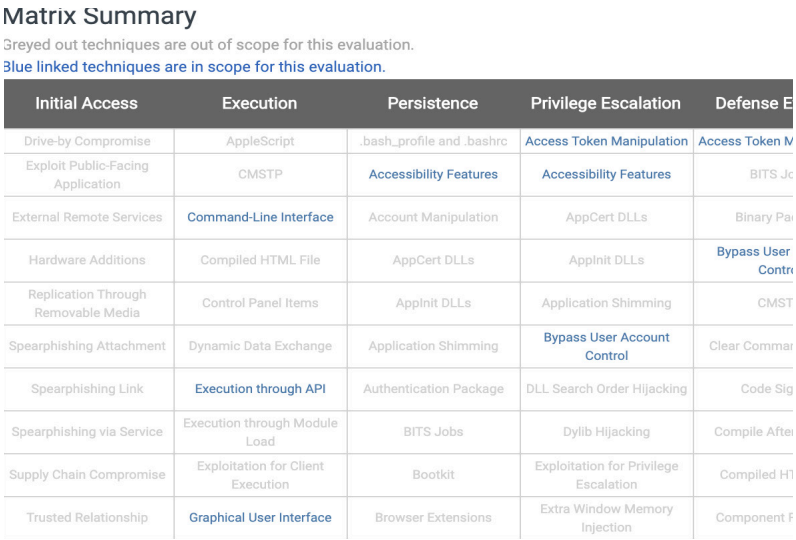
Once you have answered these high-level questions, you can use ATT&CK Evaluation results to help select a security product.

GETTING THE RESULTS

The methodology and results from all of the ATT&CK Evaluations are freely available at **https://attackevals.mitre.org**.

There are two main graphical views of the results:

1. The ***Matrix Summary view*** allows you to see all the techniques in scope for a particular round of evaluations, in a format that mirrors the ATT&CK matrix. Click on each individual technique to see all the related results, which include procedures (i.e., descriptions of the adversary behaviors), detection categories,[1] notes, and screenshots from each product.



Matrix Summary

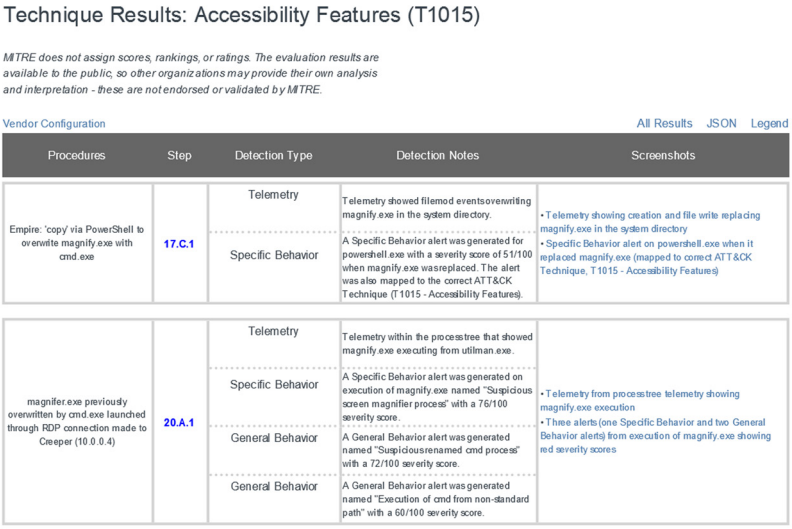
Greyed out techniques are out of scope for this evaluation.

Blue linked techniques are in scope for this evaluation.

Initial Access	Execution	Persistence	Privilege Escalation	Defense E
Drive-by Compromise	AppleScript	.bash_profile and .bashrc	Access Token Manipulation	Access Token M
Exploit Public-Facing Application	CMSTP	Accessibility Features	Accessibility Features	BITS Jo
External Remote Services	Command-Line Interface	Account Manipulation	AppCert DLLs	Binary Pa
Hardware Additions	Compiled HTML File	AppCert DLLs	AppInit DLLs	Bypass User Contr
Replication Through Removable Media	Control Panel Items	AppInit DLLs	Application Shimming	CMST
Spearphishing Attachment	Dynamic Data Exchange	Application Shimming	Bypass User Account Control	Clear Commar
Spearphishing Link	Execution through API	Authentication Package	DLL Search Order Hijacking	Code Sig
Spearphishing via Service	Execution through Module Load	BITS Jobs	Dylib Hijacking	Compile After
Supply Chain Compromise	Exploitation for Client Execution	Bootkit	Exploitation for Privilege Escalation	Compiled HT
Trusted Relationship	Graphical User Interface	Browser Extensions	Extra Window Memory Injection	Component F

FIGURE 1: PARTIAL MATRIX SUMMARY VIEW FOR AN EVALUATION

1 *https://attackevals.mitre.org/methodology/round2/detection-categories.html*



Technique Results: Accessibility Features (T1015)

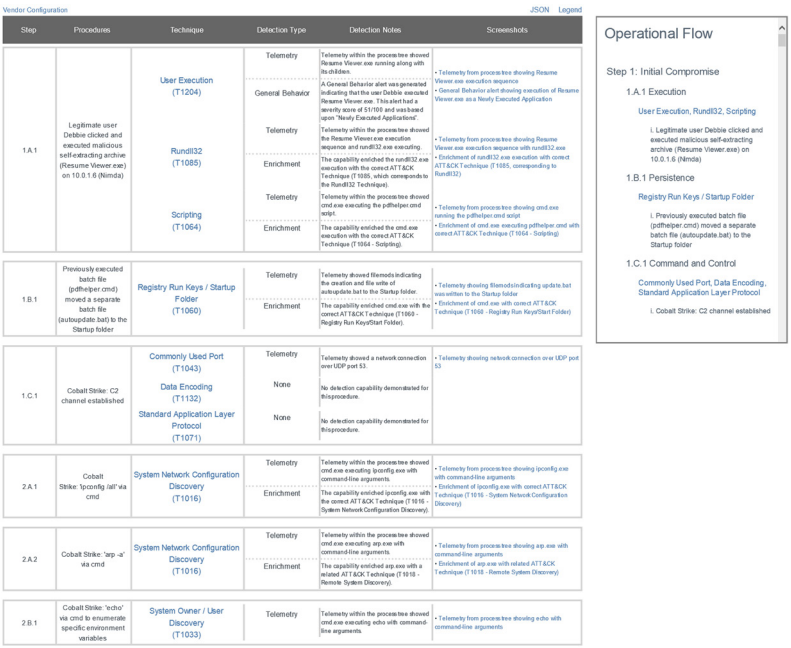
MITRE does not assign scores, rankings, or ratings. The evaluation results are available to the public, so other organizations may provide their own analysis and interpretation - these are not endorsed or validated by MITRE.

Vendor Configuration — All Results JSON Legend

Procedures	Step	Detection Type	Detection Notes	Screenshots
Empire: 'copy' via PowerShell to overwrite magnify.exe with cmd.exe	17.C.1	Telemetry	Telemetry showed filemod events overwriting magnify.exe in the system directory.	• Telemetry showing creation and file write replacing magnify.exe in the system directory • Specific Behavior alert on powershell.exe when it replaced magnify.exe (mapped to correct ATT&CK Technique, T1015 - Accessibility Features)
		Specific Behavior	A Specific Behavior alert was generated for powershell.exe with a severity score of 51/100 when magnify.exe was replaced. The alert was also mapped to the correct ATT&CK Technique (T1015 - Accessibility Features).	
magnifer.exe previously overwritten by cmd.exe launched through RDP connection made to Creeper (10.0.0.4)	20.A.1	Telemetry	Telemetry within the processtree that showed magnify.exe executing from utilman.exe.	• Telemetry from processtree telemetry showing magnify.exe execution • Three alerts (one Specific Behavior and two General Behavior alerts) from execution of magnify.exe showing red severity scores
		Specific Behavior	A Specific Behavior alert was generated on execution of magnify.exe named "Suspicious screen magnifier process" with a 76/100 severity score.	
		General Behavior	A General Behavior alert was generated named "Suspicious renamed cmd process" with a 72/100 severity score.	
		General Behavior	A General Behavior alert was generated named "Execution of cmd from non-standard path" with a 60/100 severity score.	

FIGURE 2: DETECTION RESULTS ASSOCIATED WITH A SPECIFIC TECHNIQUE (T1015)

2. The ***All Results*** tabular view shows all the techniques in scope for a particular round of evaluations. The view displays techniques in the order in which they were executed and can be navigated using the Operational Flow menu on the right. You can view the same results as clicking on a technique within the Matrix Summary view, but with the added context of seeing how the detection of each technique corresponds to adversary behavior and detections within the scenario.



Vendor Configuration JSON Legend

Step	Procedures	Technique	Detection Type	Detection Notes	Screenshots
1.A.1	Legitimate user Debbie clicked and executed malicious self-extracting archive (Resume Viewer.exe) on 10.0.1.6 (Nimda)	User Execution (T1204)	Telemetry	Telemetry within the process tree showed Resume Viewer.exe running along with its children.	• Telemetry from process tree showing Resume Viewer.exe execution sequence • General Behavior alert showing execution of Resume Viewer.exe as a Newly Executed Application
			General Behavior	A General Behavior alert was generated indicating that the user Debbie executed Resume Viewer.exe. This alert had a severity score of 51/100 and was based upon "Newly Executed Applications".	
		Rundll32 (T1085)	Telemetry	Telemetry within the process tree showed the Resume Viewer.exe execution sequence and rundll32.exe executing.	• Telemetry from process tree showing Resume Viewer.exe execution sequence with rundll32.exe • Enrichment of rundll32.exe execution with correct ATT&CK Technique (T1085, corresponding to Rundll32)
			Enrichment	The capability enriched the rundll32.exe execution with the correct ATT&CK Technique (T1085, which corresponds to the Rundll32 Technique).	
		Scripting (T1064)	Telemetry	Telemetry within the process tree showed cmd.exe executing the pdfhelper.cmd script.	• Telemetry from process tree showing cmd.exe running the pdfhelper.cmd script • Enrichment of cmd.exe executing pdfhelper.cmd with correct ATT&CK Technique (T1064 - Scripting)
			Enrichment	The capability enriched the cmd.exe execution with the correct ATT&CK Technique (T1064 - Scripting).	
1.B.1	Previously executed batch file (pdfhelper.cmd) moved a separate batch file (autoupdate.bat) to the Startup folder	Registry Run Keys / Startup Folder (T1060)	Telemetry	Telemetry showed filemods indicating the creation and file write of autoupdate.bat to the Startup folder.	• Telemetry showing filemods indicating update.bat was written to the Startup folder • Enrichment of cmd.exe with correct ATT&CK Technique (T1060 - Registry Run Keys/Start Folder)
			Enrichment	The capability enriched cmd.exe with the correct ATT&CK Technique (T1060 - Registry Run Keys/Start Folder).	
1.C.1	Cobalt Strike: C2 channel established	Commonly Used Port (T1043)	Telemetry	Telemetry showed a network connection over UDP port 53.	• Telemetry showing network connection over UDP port 53
		Data Encoding (T1132)	None	No detection capability demonstrated for this procedure.	
		Standard Application Layer Protocol (T1071)	None	No detection capability demonstrated for this procedure.	
2.A.1	Cobalt Strike: 'ipconfig /all' via cmd	System Network Configuration Discovery (T1016)	Telemetry	Telemetry within the process tree showed cmd.exe executing ipconfig.exe with command-line arguments.	• Telemetry from process tree showing ipconfig.exe with command-line arguments • Enrichment of ipconfig.exe with correct ATT&CK Technique (T1016 - System Network Configuration Discovery)
			Enrichment	The capability enriched ipconfig.exe with the correct ATT&CK Technique (T1016 - System Network Configuration Discovery).	
2.A.2	Cobalt Strike: 'arp -a' via cmd	System Network Configuration Discovery (T1016)	Telemetry	Telemetry within the process tree showed cmd.exe executing arp.exe with command-line arguments.	• Telemetry from process tree showing arp.exe with command-line arguments • Enrichment of arp.exe with related ATT&CK Technique (T1016 - Remote System Discovery)
			Enrichment	The capability enriched arp.exe with a related ATT&CK Technique (T1016 - Remote System Discovery).	
2.B.1	Cobalt Strike: 'echo' via cmd to enumerate specific environment variables	System Owner / User Discovery (T1033)	Telemetry	Telemetry within the process tree showed cmd.exe executing echo with command-line arguments.	• Telemetry from process tree showing echo with command-line arguments

Operational Flow

Step 1: Initial Compromise

1.A.1 Execution

User Execution, Rundll32, Scripting

i. Legitimate user Debbie clicked and executed malicious self-extracting archive (Resume Viewer.exe) on 10.0.1.6 (Nimda)

1.B.1 Persistence

Registry Run Keys / Startup Folder

i. Previously executed batch file (pdfhelper.cmd) moved a separate batch file (autoupdate.bat) to the Startup folder

1.C.1 Command and Control

Commonly Used Port, Data Encoding, Standard Application Layer Protocol

i. Cobalt Strike: C2 channel established

FIGURE 3: "ALL RESULTS" VIEW, SHOWING THE OPERATIONAL FLOW

For more programmatic access to the data, you can also download each vendor's raw results for analysis from their overview page. To access this information, go to a vendor's results page at **https://attackevals.mitre.org/evaluations.html** and click on the JSON link, highlighted in red in the image below.

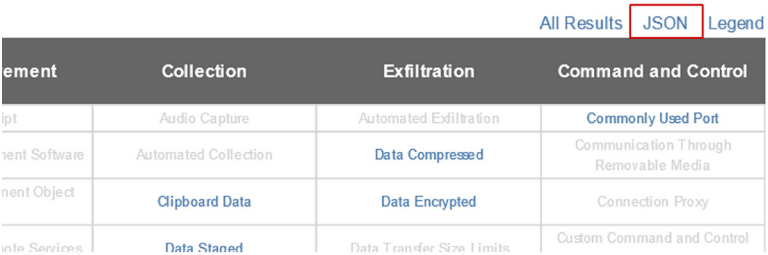


FIGURE 4: HOW TO ACCESS ALL RESULTS IN A JSON FILE, AS A LINK ON THE VENDOR'S SUMMARY PAGE.

This JSON can be analyzed to examine detection categories, technique or tactic coverage, and other deep dives. We have released a data analysis tool, Joystick,[2] accessible via our website, that will allow you to more easily interact with vendor results via a user interface and create analytics to explore the data. We'll provide a short overview in the following section of how to use this tool, but please visit the website for more details on how to leverage this capability. Additionally, while we don't endorse any specific external tooling/methods, other third-party resources and analysis are available.

2. *https://attackevals.mitre.org/tools/data_analysis.html*

TECHNIQUE COVERAGE

Broad Coverage

For broad coverage, check to see if there's a detection for all or most steps (e.g., how many of the steps and/or techniques had a detection?).

To do this, look at the number of detections against number of steps. The easiest way to do this is with the recently released Joystick ATT&CK Evaluations data analysis tool.

Getting started with Joystick is simple; it comes prepackaged with all ATT&CK Evaluations data that has been released and is ready for analysis. Joystick is a local web application, meant for anyone to spin up on their own computer and view in their browser. Get the tool here: **https://attackevals.mitre.org/tools/data_analysis.html**.

Once you've set up the tool, navigate to Joystick's Results page, and select the vendor you want to analyze. This will redirect you to the vendor's summary page, which shows all detections across all steps for the given evaluation. At this point you would be able to observe basic visibility—on what steps did they have some form of a detection?

You can also choose to look at coverage, by only considering specific detection categories. For instance, do you only care about telemetry (Figure 5) so that your analysts will be able to utilize the data, or do you care about pre-built analytics across the entire attack chain? To restrict the visualization, simply select the desired detection category, and Joystick's interface will automatically update.

For visibility, start with any detection capability, but then dive deeper. For example, if the visibility is just telemetry, would that be enough for your analysts?

Additional questions to ask about a security vendor's product include:

- Is data enriched by annotating telemetry with relevant facts that help you make decisions?
- Is data correlated by tying telemetry to previously detected badness?
- Are there a lot of alerts?
- How much relevant detail do the alerts provide?
- Are the alerts actionable?

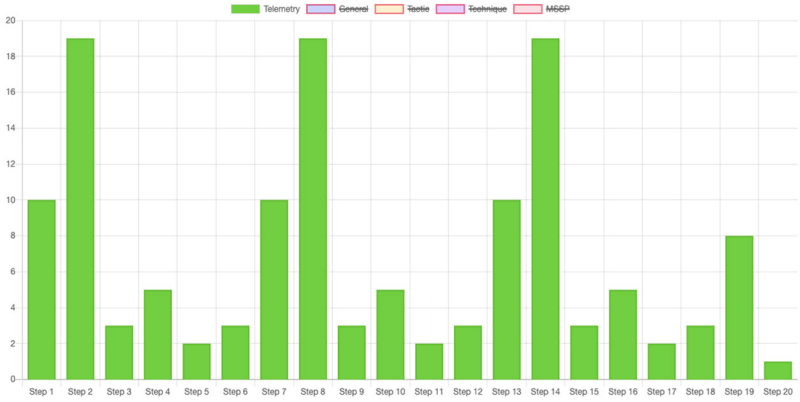


FIGURE 5: EXAMPLE ANALYSIS SHOWING THE STEPS THAT HAD TELEMETRY DETECTIONS.

Targeted Coverage

For targeted coverage, you need to identify techniques that are specifically of interest to your organization. These techniques may be ones you've identified as used by specific threats you're concerned with, ones you've identified as a gap in your current security stack, or techniques that are particularly impactful to your organization.

Next, select a subset of these of techniques to perform a deep dive analysis to see how these solutions satisfy your top needs. We will use T1003 Credential Dumping as an example.[3]

You can do more detailed analysis of the JSON results to look at a subset of techniques. Alternatively, you can also use the Technique Comparison Tool to focus on specific technique results and perform a side-by-side comparison of all vendor results for that technique. To access the tool, navigate to the Technique Comparison Tool under Tools in the ATT&CK Evaluations website's menu, as shown on the following page.

3 ***https://attack.mitre.org/techniques/T1003/***

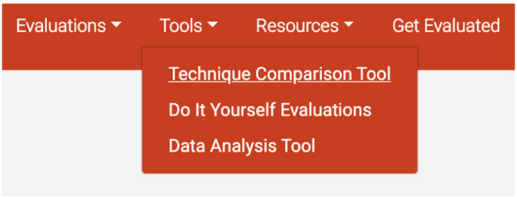


FIGURE 6: HOW TO ACCESS THE TECHNIQUE COMPARISON TOOL ON THE ATT&CK EVALUATIONS WEBSITE

Along the left side of the tool, you can select any technique from the evaluation in the Operational Flow. Simply scroll through the Operational Flow and select one of the techniques identified for additional analysis (e.g., "5.A.2—Credential Dumping").

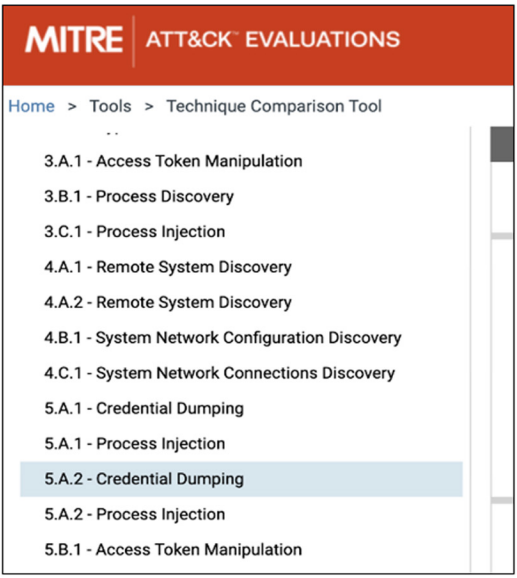


FIGURE 7: SCROLLABLE OPERATIONAL FLOW FOR THE TECHNIQUE COMPARISON TOOL

You will see the results for that technique, for all vendors who participated in the round of evaluations. This allows you to easily see how many detections each vendor had, what type of detections they were, and the details of those detections. Additionally, you can read any footnotes and see applicable screenshots. Selecting 5.A.2 Credential Dumping in the Operational Flow shows the following results:

Detection Types	Detection Notes
Telemetry	Telemetry showed an open handle to a thread into lsass.exe, which is indicative of process injection for credential dumping.
Specific Behavior (Tainted)	A second Specific Behavior alert was generated for Credential Dumping, which indicated that \"a remote thread in LSASS accessed credential registry keys.\" The alert was mapped to the correct ATT&CK Technique (Credential Dumping) and Tactic (Credential Access). The process tree view showed the alert as tainted by a parent detection.
Telemetry	Telemetry for the lsass remote thread and DLL loading would be available in a separate view.
General Behavior (Delayed, Tainted)	OverWatch also generated a General Behavior alert indicating the Credential Dumping activity was suspicious. The process tree view showed the alert as tainted by a previous detection.
Telemetry (Tainted)	Telemetry showed svchost.exe injecting into lsass.exe. The telemetry was tainted by the parent "injected (svchost.exe > lsass.exe)" alert. The hashdumpx64.dll was also seen loaded as a floating executable code.
Specific Behavior	A Specific Behavior alert was generated for the correct ATT&CK Technique (Credential Dumping).
None	No detection capability demonstrated for this procedure.
Enrichment	The capability enriched svchost.exe injecting a thread into lsass.exe with a tag identifying credential dumping.

FIGURE 8: EXAMPLE TECHNIQUE COMPARISON TOOL RESULTS FOR CREDENTIAL DUMPING

You should also consider whether the detection data makes sense for your organization. Ask yourself:

- Is this the right type of visibility for this technique?
  - e.g., an alert for Credential Dumping could mean more than an alert for Process Discovery.
- Will this detection have false positives in your environment?
  - e.g., an alert for PowerShell executing whenever your system administrators use PowerShell may cause too many false positives.
- Will it provide something actionable?
  - e.g., does it tell us:
    - What process dumped credentials?
    - Whether the credentials were plaintext or hashes? What user accounts were compromised?
    - If a known or unknown credential dumper was used?
- Would this complement your other defenses?
  - e.g., if you have honeytoken credentials in place and you detect use of them, is an alert for credential dumping as important as other alerts where you have no other coverage?

GRAPHICAL USER INTERFACE

Before analyzing any GUI, you need to first think about your analysts' workflow. Do they intend to go to another dashboard to analyze the tool's data, or do they intend to push the events to a centralized repository (ex. SIEM or orchestrator) and never look directly at the tool's GUI?

Assuming analysts do plan on using the GUI, you need to think about the level of sophistication and the use cases that are their primary drivers. Different tier analysts will have different needs. For example:

- Tier 1: Alert Triage
  - Quick analysis and summaries
  - Meaningful alerts
  - Low number false positives
- Tier 2: Incident Response
  - Correlation of related events
  - Detailed context on alerts/events
- Tier 3: Advanced IR and Hunting
  - Visibility is paramount
  - Ability to create new capabilities and reconfigure

For the sake of discussion, let's say your organization employs all three tiers of analysts and plans on using the user interface that the tool provides.

As you did when you looked at specific technique results, you start by selecting a step of interest within the Technique Comparison Tool. As your analysts walk through the screenshots provided for each of the vendors, a short survey could help standardize their results. Examples of survey questions include:

- Was there an alert for the behaviors?
  - Was the behavior clearly defined as noteworthy?
  - Is it easy to determine how noteworthy the event was?

- Were there adequate details on the offending behavior?
  - Was there supporting telemetry?
  - Was there a human readable description of what the event meant?
  - Was it correctly mapped to ATT&CK?
- Was the detection correlated to other suspicious activity?
  - Is it easy to follow the correlation to determine what else happened?
  - Does it help explain the full scope of the behavior or is more investigation needed to uncover additional details?
- Was there unnecessary/incorrect information?
  - Did the analyst not agree with the severity level assigned?
  - Was the ATT&CK technique mapping incorrect?
- Would this detection have been noisy in your environment?
- Does the tool's GUI have the features the analysts need, and will it make the analysis process easier?
  - Are there support tools analysts need to use?
  - Are the various interfaces in the tool easy to understand?
  - How much clicking around is required to get to the information analysts need to draw a conclusion?

Analysts from across the tiers should complete the survey. You can assign values/weights to the potential answers or take a more subjective analysis of their responses. In either case, it's important to look across a sizable set of analysts to avoid making a decision purely based on a single analyst's preference.

NON-ATT&CK EVALUATION INPUTS

Other vendor evaluation criteria are beyond the scope of ATT&CK Evaluations, but include:

Cost

Reach out to vendor for quotes for specifically what is and is not needed based on what the vendor offers and what products/configurations were used during the ATT&CK Evaluation.

Integration

Evaluating integration can be hard without actually deploying the solution and seeing how it integrates into the security stack.

One option is to look for advertised partnerships. Vendors who are partners with existing solution vendors may have experience or pre-built plugins that can be leveraged. Homegrown custom integrations may be the most difficult to deploy.

Making a Decision

DOWN SELECT

Finally, you bring together the information gained from your ATT&CK Evaluations analysis with the other non-ATT&CK Evaluation information gathered. In the example below, we've identified three solutions, each with their own pros and cons. The red arrows indicate where ATT&CK Evaluations provided input to the decision.

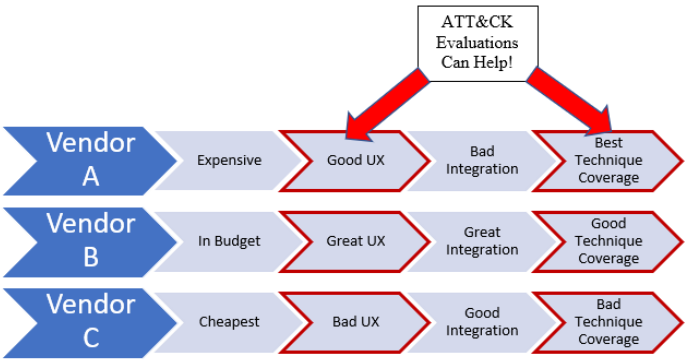


FIGURE 9: HYPOTHETICAL ANALYSIS OF THREE VENDORS

GET A SECOND OPINION

While every organization should make their own decisions based on their own needs, other opinions do matter. Talk to friends, colleagues, and industry peers who may have experience with the tools you're considering.

Additionally, read other analysis of the tools you're considering and of the market as a whole. Different testing organizations will come to different conclusions based on their methodologies. Some organizations even have their own interpretation of ATT&CK Evaluation results. All these data points can help reinforce or refine your selection.

TALK TO THE VENDOR

You've likely built up a list of questions based on information available about the products you're interested in pursuing. The next step is to talk to the vendors about your questions and any concerns you have. They should be able to provide additional insight into your requirements and what their solutions can offer.

GET HANDS-ON EXPERIENCE (PROOF OF VALUE/PROOF OF CONCEPT)

There's no replacement for testing solutions in your own environment. It gives legitimacy to the evaluation results and helps identify the specific configuration that would work best for your organization. The next step is to run a proof-of-value test with the vendors that have the best match to your requirements.

Why run an internal test? Benefits include:

- Alert performance (i.e., false positives) based on users' activities
- Analysts' hands-on impressions
- Ground truth for ease of deployment and use
- Seeing what an adversary looks like in your environment through the tool

USE ADDITIONAL MITRE TOOLS FOR YOUR OWN TESTING

To help with internal testing, MITRE released the ATT&CK Evaluations Operational Flow that we use for each round of testing. The goal was to make it as easy as possible for our Evaluations methodology to be used by security teams to perform similar tests in their own environments.

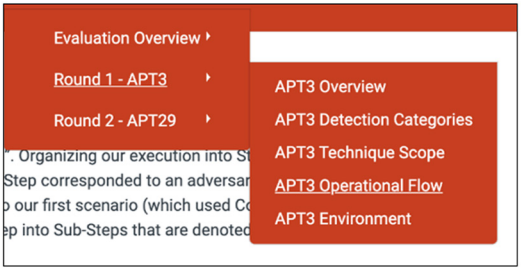


FIGURE 10: HOW TO ACCESS APT3'S OPERATIONAL FLOW - THE STEP-BY-STEP ACTIONS RED TEAM USES DURING THE EVALUATION - ON THE ATT&CK EVALUATIONS WEBSITE

MITRE also released the Do It Yourself version of ATT&CK Evaluations to create an automated adversary emulation option using CALDERA, an open-source automated adversary emulation system developed by MITRE. This is useful for organizations that lack the resources to do hands-on manual assessment or that want continual testing.

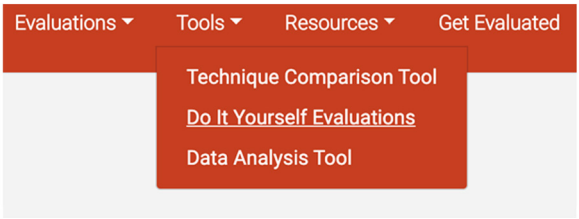


FIGURE 11: HOW TO ACCESS THE DO IT YOURSELF EVALUATIONS WITH CALDERA ON THE ATT&CK EVALUATIONS WEBSITE

You're also free to customize or develop your own evaluation methodology based on the adversaries or techniques of most concern to your organization.

MAKE THE DECISION

Refine your initial analysis, taking into account the personal experience gained from your internal evaluation. Once you decide what tool may work best for what you need, use the analysis to document and justify the decision to those approving the purchase.

To help convey your analysis, the following suggestions may be useful:

- Visualize the coverage of ATT&CK techniques with your current security stack and the expected coverage with the addition of the tool, to show improvement
- Document useful metrics such as:
  - Estimated percentage reduction in analyst fatigue
  - Estimated percentage increase in incidents detected
  - Estimated time reduction in analyzing/responding to incidents
  - Estimated analyst time saved that can be put towards other initiatives

Empowered by the analysis you have performed, we hope your organization can feel confident in the solution selected and that the solution selected provides the added security you're seeking. Your decision will be threat-informed and take into account your analysts' needs and your organization's use case. In a world full of options to secure your organization, this tailored decision is what you should strive for.

What's Next?

As your organization moves forward with the procurement and begins deployment and operationalization of the solution, remember that your analysis was done at a single point in time. Solutions evolve, as do the threats to your organization. You should continually reassess your security needs and operations, based on the threats to your organization and your evolving capabilities. The adversary will evolve, and as defenders you must as well.

ATT&CK Evaluations can provide guidance to defenders on how to leverage their tools as well as to decision makers on where to devote resources. Consider looking at the results of the solution you bought but also at other solutions to understand whether there are ways to improve your deployment, be it with additional data collection or with analytics using the data you have.

A number of open source projects also dive into ATT&CK analytics, and a vibrant community centers around improving defenses by better utilizing ATT&CK. Ask questions of the community, and share what you are learning, so together we can improve cybersecurity for everyone.

Stay Informed

attackevals.mitre.org

twitter.com/MITREattack

linkedin.com/showcase/mitre-att&ck/

medium.com/mitre-attack

Appendix—Key Things to Know

ATT&CK Evaluations are a starting point. We use an open-book and minimally sized environment to understand baseline capabilities of solutions. Operationalization of these solutions is important to consider in the context of your organization, including false positive generation.

There are no scores or winners. The goal of ATT&CK Evaluations is to show the different capabilities of each vendor.

Categories are neither good nor bad. Despite the name, "Tainted" isn't meant to be a bad thing—it just describes a detection put in the context of other malicious behavior. This term was changed to "Correlated" in the APT29 Evaluations.

Not all techniques are created equal. A technique detection for Credential Dumping may not have same value as a technique for Process Discovery, due to the severity of the action. The category gives you a general idea, but you should dive into the details to understand the technique and detection.

Not all procedures are created equal. Process Discovery (T1057) via Command-Line Interface (T1059) can be detected with most process monitoring. Process Discovery via API (T1106) would need API monitoring. A vendor could have a detection for one but not the other.

Counting has limitations. We don't think any single way to count is right for everyone. As we've described, you should consider your own needs and then consider counting based on those.

About the Contributors

FRANK DUFF

Frank Duff is a Principal Cyber Operations Engineer for MITRE and is the ATT&CK Evaluations Lead. Frank previously co-developed MITRE's Leveraging External Transformational Solutions research and development effort that works with commercial cybersecurity vendors to accelerate their adoption by the government community. His interests focus on endpoint security and adversary emulation, as well as leveraging these interests in public-private partnerships to drive organizational security and product improvement.

CONNOR MAGEE

Connor Magee is a Cybersecurity Engineer for MITRE, where he works on a variety of projects involving security operations and development. His interests span the realm of threat-based defense, supporting roles in threat detection, red team operations, and white team tool development. He is the lead developer for ATT&CK Evaluations, as well as a contributor on the MITRE ATT&CK project.

BLAKE STROM

Blake is a Principal Cybersecurity Engineer and Adversary Emulation Capability Area Lead at MITRE. He is the co-creator and lead of MITRE's ATT&CK project. Blake also helped stand up the ATT&CK Evaluations project to provide open and objective insights to detection capabilities on the market as well as CALDERA to automate adversary emulation. His interests span the security domain including security operations, red teaming, adversary emulation, hunting, and automation. Prior to MITRE, Blake worked for the National Security Agency as a cyber operations lead.

JAMIE WILLIAMS

Jamie Williams is a Lead Adversary Emulation Engineer for MITRE, where he works on various efforts involving security operations and research. Jamie specializes in threat-informed red team operations and the analysis of behavior-based detections. He is the Red Team Lead for ATT&CK Evaluations as well as a key member of the MITRE ATT&CK team.

SARAH YODER

Sarah Yoder is a Cybersecurity Engineer for MITRE. She enjoys furthering her red team skills and applying cyber threat intelligence to MITRE ATT&CK. She is the White Team Lead for ATT&CK Evaluations and the creator behind the Threat Report ATT&CK Mapper (TRAM). Prior to joining MITRE, Sarah worked as an Exploit Analyst with the Department of Defense.

About MITRE ATT&CK

MITRE ATT&CK® is a globally-accessible knowledge base of adversary tactics and techniques based on real-world observations. The ATT&CK knowledge base is used as a foundation for the development of specific threat models and methodologies in the private sector, in government, and in the cybersecurity product and service community.

With the creation of ATT&CK, MITRE is fulfilling its mission to solve problems for a safer world—by bringing communities together to develop more effective cybersecurity. ATT&CK is open and available to any person or organization for use at no charge.

Learn more at attack.mitre.org.

About MITRE

MITRE's mission-driven teams are dedicated to solving problems for a safer world. Through our public-private partnerships and federally funded R&D centers, we work across government and in partnership with industry to tackle challenges to the safety, stability, and well-being of our nation.

Learn more at www.mitre.org.

Notes:

SOLVING PROBLEMS
FOR A SAFER WORLD