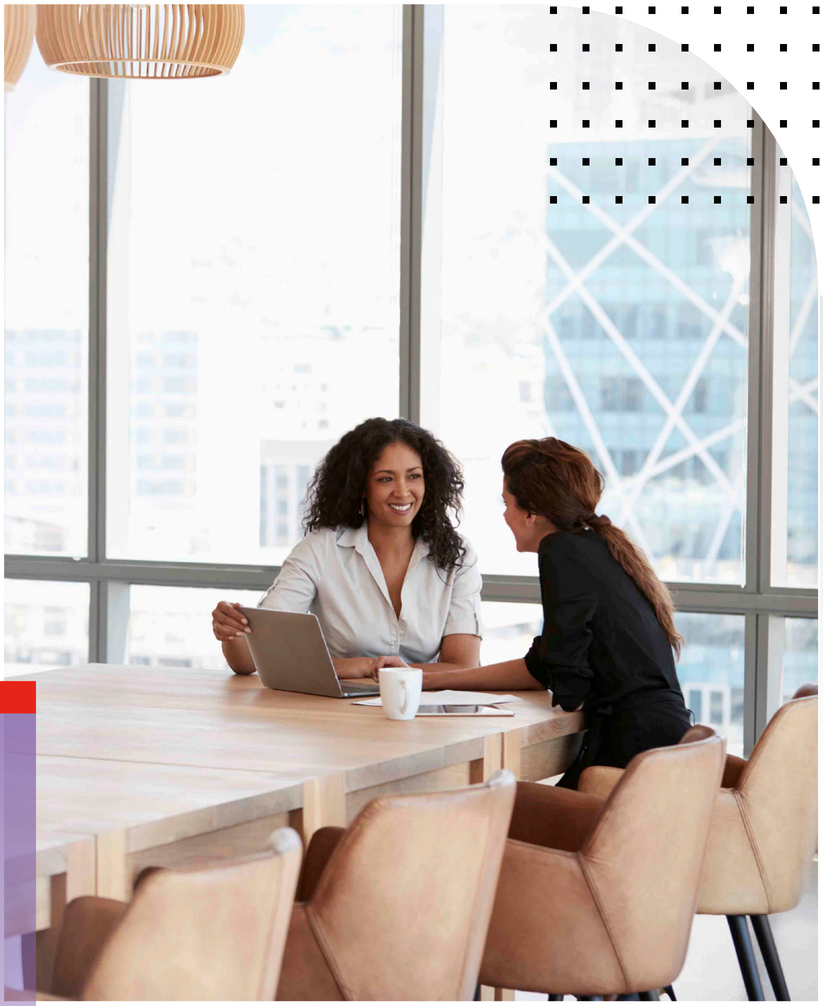
WHITE PAPER

Assess Your Endpoint Security

Better Understand Capabilities Using MITRE Engenuity ATT&CK Evaluations



Executive Summary

Making a decision on endpoint security tools is complex, so access to objective vendor-neutral information is key. Organizations can use the results from the MITRE Engenuity ATT&CK evaluation to evaluate the efficacy of endpoint solutions. They also can gain a better understanding of where their current security posture stands by using MITRE tools to evaluate their existing detection coverage and gaps. Then they can compare their results against MITRE's extensive list of tactics and techniques that adversaries use in real-world attacks.

Endpoint security is all about reducing risks, so before evaluating solutions, security professionals need to make sure they have the basics in place. Best practices in security hygiene can dramatically lower threat exposure, and security professionals should have a strategy for improving their security posture and visibility. Once they have these fundamentals in place, they can use the results of third-party tests to evaluate and select solutions that best fit their needs.

Endpoint detection and response (EDR) was the most often-cited priority when organizations were asked what their biggest endpoint security investment priorities are for the next 12-18 months.[1]

Objective Product Comparisons

The constantly evolving threat landscape is characterized by substantively new versions of existing attacks as well as brand new cyber threats. It's critical that an organizations' security posture and the individual security controls on which it is built remain strong. However, cybersecurity products can be difficult to assess because they often take different approaches or use different terms to describe how they do what they do. Third-party testing that evaluates products can help organizations compare "apples to apples" so they can make more informed choices and find the right solution for their current situation and their desired security posture. However, it can be challenging to determine the value of testing as it relates to the overall security infrastructure when testing takes a "black box" approach to reporting that only shows the outcome (blocked or missed, detected or undetected) rather than the operations.

MITRE ATT&CK Evaluations

Starting in 2019, the MITRE evaluations have acted as objective and detailed tests of the capabilities of endpoint security solutions by emulating real-world cyber campaigns and their techniques and tactics. 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. The results not only show the technical ability of a solution to detect a specific campaign but more importantly also show the techniques and tactics that are common among many of today's cyber threats.

The MITRE Engenuity ATT&CK evaluations are powerful because they are based on the MITRE ATT&CK Framework, which is a knowledge base of adversarial techniques. It provides a breakdown and classification of offensive actions taken by attackers that can be used against particular platforms, such as Windows. Unlike prior work in this area, the focus isn't on the tools and malware that adversaries use but on how they interact with systems during an operation.

To provide context, the ATT&CK Framework organizes techniques into a set of tactics. Each technique includes information that's relevant to defenders to help them understand the context surrounding events or artifacts generated by a technique in use. The relationship between tactics and techniques can be visualized in the ATT&CK Matrix, which spans 14 discrete techniques. The Matrix offers a robust and granular mapping of the activity of potentially utilized cyberattacks. Each area has 7 or more tactics and spans from reconnaissance through impact.

The 2022 round of MITRE Engenuity ATT&CK tests focused on two threat actors, Wizard Spider and Sandworm. Wizard Spider is a financially motivated criminal group that has been conducting ransomware campaigns since August 2018 against a variety of organizations, ranging from major corporations to hospitals. Sandworm is a destructive Russian threat group that is known for carrying out notable attacks such as the 2015 and 2016 targeting of Ukrainian electrical companies and 2017's NotPetya attacks. These two threat actors were chosen based on their complexity, relevancy to the market, and how well MITRE Engenuity's staff can fittingly emulate the adversary.

In the tests, MITRE first runs the detection test to see what the EDR solution will catalog followed by the protection test to see if or when it will block the attack.

The Detection Test

The 2022 test was composed of 19 steps with multiple stages called *sub-steps*. The evaluations used five terms to express how the product performed for each test and also noted the data source for the detection.

- **None.** No data indicating the test behavior was detected could be seen within the product.
- **Telemetry.** The behavior could be seen but was minimally processed.
- **General.** The behavior was processed and flagged, but without detail as to why (tactic) or how (technique) the action was performed.
- **Tactic.** The behavior was processed and designated malicious, as well as enriched with the tactic or other notation about why it was performed by the cyberattack.
- **Technique.** The behavior was processed and designated malicious, as well as enriched with the technique or other notation about how it was performed by the cyberattack. Detecting each sub-step via the technique is ideal.

In cases where the tested behavior is often associated with potentially legitimate operations, as well as malicious ones, a designation of None may not be a negative result, particularly if the risk of letting the behavior pass is also None. An example is when the cybercriminal has not yet achieved their intended outcome.

Telemetry or General are the first level of detection where behavior can be identified and recorded in the solution, but with limited detail about why it was identified. This level is sufficient for organizations that lack the time or expertise to dig into the detail of exactly how the cybercriminal is attempting to conduct their mission. Detections tagged with MITRE Tactic or Technique are most valuable to seasoned security professionals who want to understand the detailed activity of the cybercriminals.

The Protection Test

The protection tests only use two terms

- **None.** There was no evidence that the technique was blocked or otherwise unsuccessful because of the product.
- **Blocked.** The technique was blocked and the user was informed that the attack was unsuccessful.

Although there are fewer terms, the interpretation is more complex because the timing of the blocking is important and can vary based on the test case. Testing for false positives is not part of the scope of the evaluation, but consider that blocking early may increase the possibilities of false positives. In contrast, blocking too late may expose the organization to a degree of risk even if the end objective is not achieved. Here is an example of each scenario.

Suppose that Test 1 (steps 1-3 in this test) was blocked at the first step 1.a.1. The result sounds great; the cyberattack was stopped at the earliest stage possible. But what if it was actually the user execution that was blocked? In that case, you would want to know the basis on which the user was prevented from accessing a file. Was there a high confidence malicious indicator or were policies set too aggressively? Alternatively, suppose blocking occurred at the end at step 3.a.5, which was email collection. In that case, the product stopped the intended data breach, but it did allow step 1.A.8 remote file copy to occur, which means the attack had a malicious impact. Often the impact of an attack can have several negative consequences.

In this particular case, arguably the safest time to block the attack to minimize the risk of false positives given the "proof" gathered and the malicious impact given the intended action would be at step 1.a.3. This step occurs when a script attempts the first malicious file manipulation. But this information is something that can only be determined after understanding each sub-step of each stage. The success or failure is based on an organization's concern about impeded legitimate user activity vs. the risk of malicious impact from a cyberattack.

An Important Reminder About the Evaluation Results

MITRE emphasizes the fact that the evaluations are not a competitive analysis and there are no scores, rankings, ratings, or "winners." Instead, they show the detections observed and how each vendor approaches threat detection through the language and structure of ATT&CK. The evaluations can help organizations answer questions such as:

- Does this tool detect known threats to your organization?
- How does the tool present the data to your analysts?
- Can it strike the right balance between aggressive detection/protection and the potential risk of compromise?

The evaluations can tell you which vendors provide the most visibility across adversary techniques and the vendors that best address the techniques that the threats use. An evaluation can show you the insights you may receive and how often a solution is updated to cover new adversary techniques. The evaluations also can tell you if a tool uses a graphical user interface (GUI) or offers turnkey options for less experienced analysts or provides raw data that more experienced analysts may need.

However, the evaluations can't answer questions like:

- What is the impact on systems and users?
- What is the volume of alerts and the amount of manual research and investigation that is required?
- How does the tool fit within your broader security posture? Is the tool additive or duplicative?
- Does the system erroneously block a legitimate action?
- How does the tool integrate with your other tools?
- How much does the tool cost?

Getting answers to questions like these requires additional research, testing, and considering the bigger picture in your organization.

Takeaways From the Evaluations

The greatest value of MITRE testing is that it demonstrates a product's ability to defend against tactics and techniques that are represented by but not limited to the attack samples. Using the evaluations, you can assess your exposure to previously unknown attacks based on its tactics and techniques, as opposed to basing your exposure level on existing one-to-one or one-to-many threat intelligence or models.

CISOs can use the results of the MITRE Engenuity ATT&CK evaluations to assess gaps in their security coverage. No single solution can detect every attack or technique that may exist, but you can learn which products can detect a given type of attack. Organizations need an integrated approach so that they can:

- Detect and block threats as early in the attack as possible
- Balance confidence in an event before a block vs. early blocking
- Find out how a solution contains a threat. Does it block specific malicious actions in real time and defuse them using micro containment and process isolation, or does the system rely on network isolation to prevent lateral movement?
- Reserve "heavy-handed" containment tools such as process termination and endpoint isolation for situations where it's appropriate
- Streamline security operations

Retaining Perspective

Endpoint security is more critical than ever. Fast-moving attacks like ransomware can take minutes, if not seconds, to inflict damage across multiple networks. The manual responses in first-generation EDR tools aren't enough anymore. The objective of a robust cybersecurity infrastructure (including endpoint security) is to reduce the overall risk exposure. Effective security policies and continuous monitoring need to be in place to discover and predict, prevent, detect and respond, and remediate attacks.

Discovery, hardening, and prevention are the foundations of security hygiene. Doing these basics correctly can dramatically reduce risk. However, every CISO will agree that while essential, prevention is never 100%.

As a result, in addition to prevention, enterprises need to have the ability to effectively detect threats early, respond and contain threats rapidly to stop a breach, and recover to return to a known good state. Ultimately, the goal is to minimize business disruption and help ensure enterprise resiliency.

MITRE ATT&CK evaluations can help organizations evaluate endpoint security solutions based on their efficacy of detection and some degree of prevention with the recent protection test. However, when selecting endpoint security to meet their needs, in addition to looking at MITRE results, organizations should:

- Practice good security hygiene with discovery and prediction to reduce the attack surface with visibility and preemptive control (such as virtual patching or application control) until patches can be applied
- Improve accuracy to reduce the impact of false positives and alert fatigue
- Focus on responses and minimizing impact with precise and automated responses for more effective risk mitigation
- Maintain availability and system stability even in the midst of an attack, particularly for operational technology (OT) and executive systems
- Use technology to its full potential; no tool will deploy, run, and maintain itself

Conclusion

MITRE ATT&CK evaluations look beyond simple effectiveness scoring of a security product to explore its operation. The primary benefit of this approach is a better understanding of the capabilities you are deploying. In conjunction with other assessments, you can go beyond the specific sample set used for testing. MITRE ATT&CK evaluations are an excellent resource to help select the endpoint security product that will reduce your cybersecurity risk in the areas of greatest concern while balancing the impact on people, process, and systems. The evaluations also can help you determine a solution's fit within your overall security infrastructure and posture.

However, no security solution should exist in a vacuum. Certain techniques cannot be detected in the endpoint alone. At the same time, certain behaviors can be picked up by multiple solutions. Although overlapping coverage is better than security gaps, lack of integration can cause its own challenges as well.

[1] David Gruber, "ESG Master Survey Results: Trends in Endpoint Security," ESG, March 5, 2020.

FORTINET

www.fortinet.com

March 23, 2022 10:20 AM

968201-A-0-EN