Poster: How do Endpoint Detection Products Make Use of MITRE ATT&CK?

Apurva Virkud, Muhammad Adil Inam, Andy Riddle, Gang Wang, and Adam Bates
University of Illinois Urbana-Champaign
{*avirkud2,mainam2,rriddle2,gangw,batesa*}*@illinois.edu*

***Abstract*—MITRE ATT&CK is an open source taxonomy of adversary tactics, techniques, and procedures based on real-world observations. Increasingly, organizations leverage the ATT&CK as the basis for evaluating their security posture, while Endpoint Detection & Response (EDR) products have integrated ATT&CK into their design and marketing. However, the extent to which this integration has improved real-world security remains unclear –** *Does increasing your organization's coverage of ATT&CK improve its security?* **In this work, we attempt to answer this question by conducting a comparative analysis of EDR products' use of the MITRE ATT&CK knowledge base. We begin by evaluating 3 ATT&CK-annotated EDR detection rule sets from major commercial providers (Splunk, Carbon Black, Elastic) to identify commonalities and underutilized regions of the ATT&CK matrix. We continue by performing a complete qualitative analysis of ATT&CK techniques to determine their feasibility as detection rules. Our initial findings indicate potential limitations of using ATT&CK coverage as an evaluation metric for EDR tools, as we identify several techniques that do not have viable endpoint detection strategies.**

I. Introduction

Since its introduction in 2013, the MITRE ATT&CK Framework [1] has expanded beyond its initial purpose of documenting attacker behavior. The framework draws from real-world observations to define a hierarchy of common adversarial *tactics* representing high-level goals, *techniques* describing actions associated with one or more tactics, *procedures* specifying an implementation of a technique, and *detections* explaining how a technique can be detected. In recent years, ATT&CK has become intertwined with the assessment of enterprise security; a 2020 survey of security professionals found that 57% of respondents also use ATT&CK to evaluate the efficacy of deployed security products [2]. In fact, digital forensics and incident response consultants now regularly conduct audits of their clients coverage of the ATT&CK framework (e.g., [3]). Unsurprisingly, security product vendors have followed suit, such as Splunk advertising ATT&CK coverage for their EDR tool [4]. Despite its prevalence, the effectiveness of ATT&CK coverage as an evaluation metric is not well understood.

In this work, we aim to explore this relationship between MITRE ATT&CK and enterprise security products. We focus on Endpoint Detection and Response (EDR) products due to their ubiquity in enterprise environments. EDR is used as an end-to-end solution for threat detection and remediation, with four main stages: detection, isolation, investigation, and removal (of threats). During detection, the EDR system logs system-level behavior (e.g., process executions), processes the logged events with its rule set containing signatures and heuristics for potential threats, and fires alerts when an event matches. Intuitively, EDR rules will be closely linked to ATT&CK detections and procedures and can be mapped to higher levels of the framework hierarchy. Therefore, we look to analyze rule sets where individual rules are annotated with a corresponding tactic and technique to understand EDR products' use of MITRE ATT&CK.

	Carbon Black	Splunk	Elastic
# Tagged Rules	895	911	473
# Unique Techniques	105	100	92
% Technique Coverage	55%	52%	48%

TABLE I
Dataset Overview: Tagged rules are labeled with a MITRE ATT&CK technique ID.

II. Data Characterization

We consider three popular industry endpoint detection and response systems: VMware Carbon Black [5], Splunk Security Content [6], and Elastic Detection Rules [7]. Splunk and Elastic are both open-source rule sets published on GitHub, while Carbon Black is proprietary. We take a snapshot of each rule set in October 2022 and perform initial filtering before analysis. We omit Carbon Black rules on the CB Tor and Community watchlists, and omit Elastic rules that are deprecated or in development. We determine the number of rules tagged with MITRE ATT&CK technique IDs for analysis, as shown in Table I. The rules are also tagged with other metadata information, such as a text description, false positive scenarios, and a severity score.

As seen in Figure 1, we observe that *defense evasion*, *discovery*, and *persistence* are among the most frequently appearing tactics associated with the rules. *Resource development* has the lowest coverage, as some of the corresponding techniques require offline activity that is not viable to implement. Overall, the rule sets have the highest coverage of the *persistence* and *privilege escalation* tactics. Further, we also see that combining the rule sets would largely improve technique coverage for most tactics.

III. Qualitative Analysis of MITRE ATT&CK Techniques

We aim to understand the relationship between MITRE ATT&CK techniques and rule implementations in the selected

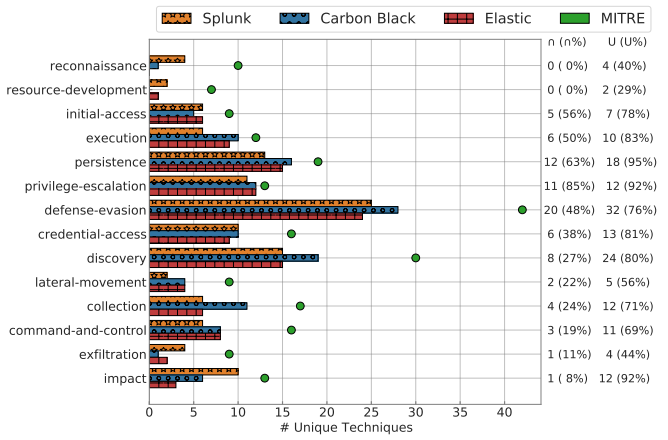


Fig. 1. **Implemented Techniques Per Ruleset Across ATT&CK Tactics:** The y-axis shows the 14 MITRE ATT&CK tactics ordered approximately by phase of attack (e.g., reconnaissance is typically the earliest goal). The dots show the total number of ATT&CK techniques associated with each tactic, while the bars indicate the number of implemented techniques in each ruleset. The size of the intersection and union of implemented techniques is shown on the right, with the percent coverage in parenthesis.

EDR rule sets. We pose three questions to categorize the techniques, regarding (1) the viability of implementing a technique with an endpoint detection strategy, (2) the effectiveness and context of a technique's detection strategies, and (3) the part of the host infrastructure a technique targets. Two authors independently coded these questions for all techniques and collaboratively came to a consensus on all codes.

We assess 26 techniques (14%) to not have viable endpoint detection strategies. Notably, half of the non-viable techniques occur during the early stages of attack development and are associated with the resource development (7) and reconnaissance (6) tactics. Further, all of these techniques target third party infrastructure (e.g., cloud infrastructure, remote repositories) or other non-endpoint targets (e.g., social media). 8 of the non-viable techniques are implemented by at least one of the three rulesets, primarily Splunk, which has 29 associated rules. The main reason for this inconsistency is due to rules using popular cloud and container services (e.g., AWS CloudTrail). The remaining rules look for system indicators that may be from malware performing reconnaissance or resource development.

We also observe 35 viable techniques that have no implemented rules. For 10 of these techniques, our codes indicated that the majority of their detection strategies were not effective or did not fit the context of the technique. Other techniques may not have been implemented for a variety of reasons, such as needing client-specific parameters in the detection strategies - for example, DS0016 monitors for unexpected file access and DS0029 monitors for unexpected surges in network traffic. Other techniques rely on human factors, such as T1598 (Phishing for Information) and T1534 (Internal Spearphishing). Additionally, techniques like T1030 (Data Transfer Size Limits) and T1041 (Exfiltration Over C2 Channel) require time series analysis that is difficult to define within a rule.

IV. Conclusion

In this work, we discuss our preliminary investigation of how EDR tools use MITRE ATT&CK. We find ATT&CK techniques that are not viable to implement as endpoint detection strategies, as well as techniques that are viable, but difficult to implement. This suggests that ATT&CK coverage may have limitations as an evaluation metric for EDR tools. We plan to conduct further analysis to understand triggers for EDR rule creation and if different products that detect the same techniques are detecting the same attack behaviors.

References

[1] The MITRE Corporation, "MITRE ATT&CK®." https://attack.mitre.org, 2023.
[2] J. Basra and T. Kaushik, "MITRE ATT&CK as a Framework for Cloud Threat Investigation." https://cltc.berkeley.edu/publication/mitre-attck/, Oct 2020.
[3] Reality Net System Solutions, "attack-coverage: An excel-centric approach for managing the MITRE ATT&CK tactics and techniques.." https://github.com/RealityNet/attack-coverage, Nov 2020.
[4] Splunk Inc, "ATT&CK® Navigator." https://mitremap.splunkresearch.com/, 2023.
[5] VMware, Inc, "VMware Carbon Black EDR." https://www.vmware.com/products/endpoint-detection-and-response.html, 2023.
[6] Splunk Inc, "Splunk Security Content." https://github.com/splunk/security_content, 2023.
[7] Elasticsearch B.V., "Detection Rules." https://github.com/elastic/detection-rules, 2023.

How do Endpoint Detection Products Make Use of MITRE ATT&CK?

Apurva Virkud, Muhammad Adil Inam, Andy Riddle, Gang Wang, and Adam Bates, *University of Illinois Urbana-Champaign*
{avirkud2, mainam2, rriddle2, gangw, batesa}@illinois.edu

Introduction

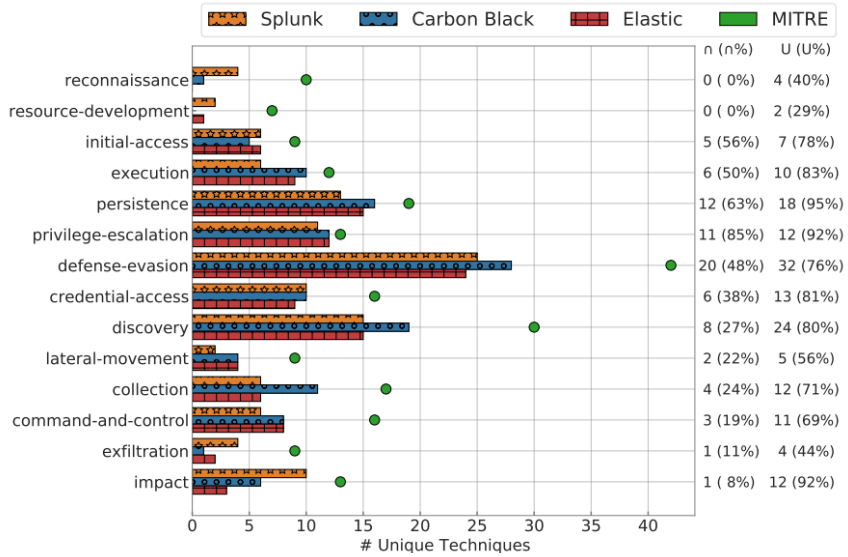
- We aim to understand the relationship between MITRE ATT&CK [1] and enterprise security tools. Recently, ATT&CK has been used to evaluate the efficacy of deployed security products, but there is **little understanding of the impact of MITRE ATT&CK coverage on security.**
- **MITRE ATT&CK** defines common adversarial *tactics* (high-level goals), *techniques* (actions), *procedures* (implementation of a technique), and *detections* (how a technique can be detected).
- We focus on **endpoint detection and response** (EDR) products due to their popularity in enterprise environments. Using annotated EDR rule sets, we analyze **how EDR products use MITRE ATT&CK.**

Methods and Results

- We characterize ATT&CK coverage of three popular commercial tools: Carbon Black [2], Splunk [3], and Elastic [4].
  - Combining the rule sets would improve technique coverage for most tactics.
- For all 191 ATT&CK techniques, we qualitatively assess (1) the viability of implementing a technique with an endpoint detection strategy, (2) the effectiveness and context of a technique's detection strategies, and (3) the part of the host infrastructure a technique targets.
  - The majority of techniques target either the endpoint system or network.
  - 26 techniques (14%) do not have viable endpoint detection strategies.
    - 8 of these were implemented by at least one rule set, primarily related to rules using popular cloud and container services.
  - We observe 35 viable techniques that have no implemented rules.
    - For 10 techniques, we assess the majority of detection strategies to be ineffective or not fit the context of the technique. Other reasons for not implementing techniques include requiring client-specific parameters, involving human factors and incorporating time series analysis that is difficult to define within a rule.
- We investigate clusters of rules with a common trigger for rule creation.
  - Reactivity ranges from immediately after a trigger occurs (e.g., Elastic introduced 4 new rules the day after the 2020 SolarWinds attack) to years later (e.g., Splunk created new rules for Mimikatz throughout 2022).

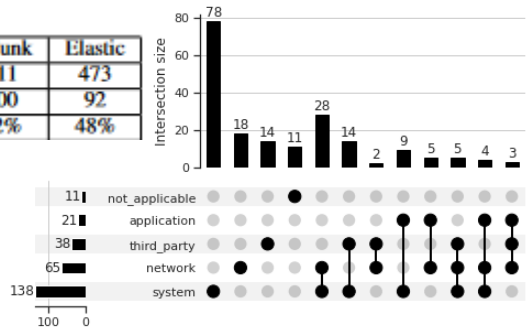
Next Steps

- We conduct comprehensive pairwise rule comparisons to understand if different products that detect the same techniques detect the same attack behaviors.



	Carbon Black	Splunk	Elastic
# Tagged Rules	895	911	473
# Unique Techniques	105	100	92
% Technique Coverage	55%	52%	48%

	CB	S	E
Name of Attack	✓	✓	✓
Description	✓	✓	✓
Known False Positive Scenarios	✓*	✓	✓
Confidence	✓*	✓	
Risk		✓	✓
Severity	✓		✓
Keywords	✓		✓



[1] https://attack.mitre.org
[2] https://www.vmware.com/products/endpoint-detection-and-response.html
[3] https://github.com/splunk/security_content
[4] https://github.com/elastic/detection-rules

ILLINOIS