USENIX Security '24 Artifact Appendix: How does Endpoint Detection use the MITRE ATT&CK Framework?

Apurva Virkud, Muhammad Adil Inam, Andy Riddle, Jason Liu, Gang Wang, Adam Bates
University of Illinois Urbana-Champaign
{avirkud2, mainam2, rriddle2, jdliu2, gangw, batesa}@illinois.edu

A Artifact Appendix

A.1 Abstract

We make available our code and analysis from our study on how MITRE ATT&CK is used by endpoint detection products to support future work in the area. This analysis investigates the extent to which ATT&CK coverage is suitable to serve as a security metric– *Does ATT&CK coverage vary meaningfully across different products? Do endpoint products that detect the same attack behaviors even claim to cover the same ATT&CK techniques?* We attempt to answer these questions by analyzing 4 ATT&CK-annotated endpoint detection rulesets; we provide the specific snapshots of the open-source rulesets (Splunk, Elastic, Sigma) that were used for analysis for reproduction of our results. Notably, our work shows that coverage of an ATT&CK technique does not consistently imply coverage of the same real-world threats.

A.2 Description & Requirements

Our GitHub repository contains our analysis code and the data used in our paper. In this work, we analyze endpoint detection rulesets' usage of the MITRE ATT&CK framework. The data consists of snapshots of 3 open-source rulesets (Splunk, Elastic, Sigma) that we used for analysis. Note that while we analyze an additional commercial ruleset in the paper (Carbon Black), we do not include it in the artifact evaluation due to publishing restrictions. We provide Jupyter notebooks containing the analysis for two of three main research questions in the paper (RQ1, RQ3). This is because RQ2 is purely qualitative analysis and thus has no associated code.

A.2.1 Security, privacy, and ethical concerns

There is no risk to evaluators while executing our artifact. The provided data is already open-source.

A.2.2 How to access

Our GitHub repository containing the code and data is found at https://github.com/avirkud/endpoint-detection-mitreattack/releases/tag/sec24-ae-final.

A.2.3 Hardware dependencies

None.

A.2.4 Software dependencies

We have provided two Jupyter notebooks (Python) for our research questions. Instructions for setting up the Python environment and required packages can be found in the repository README.

A.2.5 Benchmarks

All required data is hosted in the repository. A description of each file can be found in the repository README.

A.3 Set-up

The analysis notebooks require a Python environment (3.8.10). The required packages are specified in a requirements.txt file and can be installed with pip. See the repository README for the exact commands.

A.3.1 Installation

After setting up a Python environment and installing the required Python packages with pip, the notebooks should be ready to run.

A.3.2 Basic Test

After completing the setup, the notebook for research question 1 (code/RQ1 Analysis.ipynb) can be run in its entirety to confirm functionality. This should take less than 5 minutes to complete.

A.4 Evaluation workflow

A.4.1 Major Claims

(C1): *Certain ATT&CK tactics are highly covered by the commercial products, while others have little to no coverage. This is presented in Figures 1 and 7 in the paper. The corresponding analysis in code/RQ1 Analysis.ipynb is labeled* `Implemented Techniques Per Ruleset Across Tactics (Figures 1 & 7)`.

(C2): *There are differences in the frequency of rules implemented per technique across the products (Figure 2, 3, 8, 9). The corresponding analysis in code/RQ1 Analysis.ipynb is labeled* `Rules Per Technique (Figure 2, 3, 8, & 9)`. *However, there is a statistically significant similarity in which techniques are covered between all pairs of products (Section 4, paragraph ATT&CK Technique Density). The corresponding analysis in code/RQ1 Analysis.ipynb is labeled* `Consistency in Technique Ranking (Spearman coefficient)`.

(C3): *A minority of rules have multiple technique annotations (Figure 10). The corresponding analysis in code/RQ1 Analysis.ipynb is labeled* `Techniques Per Rule (Figure 10)`.

(C4): *Filtering out rules with lower operational values (e.g., lower risk, severity, confidence) affects the overall MITRE ATT&CK coverage. The corresponding analysis in code/RQ1 Analysis.ipynb is labeled* `Confidence, Risk, and Severity (Figures 4 & Supplementary Materials)`.

(C5): *There exist inconsistencies in technique labeling for similar rules both within the same product and across products. The corresponding analysis in code/RQ3 Analysis.ipynb is labeled* `Case Studies`. *There are four notebook subheadings for the case studies in this section:* `CVE-2021-4034`, `Meterpreter`, `Tactic Disagreement Example`, *and* `Inconsistent Technique Labels within the same product`. *These correspond to the paragraphs with the same name in Section 6.2 of the paper respectively.*

A.4.2 Experiments

(E1-E4): *[RQ1] [30 human-minutes + 5 compute-minutes]: The notebook code/RQ1 Analysis.ipynb contains all of the code for C1-4.*
Preparation: *See Section A.3 for set up instructions.*
Execution: *After completing the setup instructions, run code/RQ1 Analysis.ipynb in its entirety. The notebook contains headings for each claim's analyses.*
Results: *The corresponding figures and text in the paper for each claim are enumerated in Section A.4.1. For example, compare the notebook analysis for C1 to Figures 1 and 7 in the paper.*

(E5): *[RQ3] [30 human-minutes + 30 compute-minutes]: The notebook code/RQ3 Analysis.ipynb contains the code for C5.*
Preparation: *See Section A.3 for set up instructions.*
Execution: *After completing the setup instructions, run code/RQ3 Analysis.ipynb in its entirety. The analysis for this claim consists of several case studies under the notebook heading* `Case Studies`.
Results: *The corresponding text in the paper for each case study are enumerated in Section A.4.1. In the notebook, we caption the case study output with the specific rule indexes we are analyzing in the paper, for ease of reference. For example, compare the notebook output for C5,* `CVE-2021-4034` *to Section 6.2, paragraph CVE-2021-4034 in the paper - this case study is comparing rules e69 and s489.*

A.5 Version

Based on the LaTeX template for Artifact Evaluation V20231005. Submission, reviewing and badging methodology followed for the evaluation of this artifact can be found at https://secartifacts.github.io/usenixsec2024/.