Learning the Associations of MITRE ATT&CK Adversarial Techniques - Notes

Cameron Noakes

November 2021

1 Sentence Overview

Statistical machine learning analysis proposition for advanced persistent threats (APTs) from software attacks reported by MITRE corporation to help cluster techniques for significant correlation. 500 APT attacks per year to close to 1500 APT attacks as of 2016.

2 Introduction - Notes

appreciate how cyber security defenses are slacking behind that of adversarial tactics. numerical data on advanced persistent threats is detailed, with 500 APT attacks per year to close to 1500 APT attacks as of 2016. With SANS disclosing how 68 percent of attacks went undiscovered, the ones that were discovered often took days to remediate the threats and the damages caused.

Breakdown of Mitre Attack Framework and how as of February 2020 is a total of 440 attack techniques belonging to 27 different tactics. Explanation of the definition of tactic is given to be "A tactic is a behavior that supports a strategic goal".

Described the optimal solution proposed, For these reasons, develop an approach using hierarchical clustering to infer technique associations that represent various techniques in a TTP chain.

The researchers then then developed hierarchical clustering, which is of a more suitable approach due to complexity to represent the sophisticated attack patterns APTs use.

Outlines three key challenges for the research and study to obtain clear pathways to the solution.

1. a suitable distance metric for clustering for technique correlations.
2. address the multi-dimensional relationships exhibited by attack techniques.
3. validate the stability and significance of the clustering from step 1 using a statistical hypothesis test.

3 Datasets

The researchers analysed two different datasets from Mitre Attack which referenced 270 total attack instances which were made up of 209 techniques. The research referenced was APT threats and attacks and software attacks. Each dataset took into account all the post-exploit techniques as well, from privilege escalation to data exfiltration, all techniques were analysed and covered.

The initial dataset was said to have contained 66 APT attack instances which can be seen in the Mitre Attack Framework, these APT attacks were mapped by Mitre from publicly available exploits. The second dataset contains 204 Software attack instances which also originated from the same origin domain as the first dataset.

4 Clustering Techniques

This section managed to cover more information about clustering and how it was conducted to develop the first stage challenge of the overall proposed solution. This was mainly an intellectual specific topic for creating clustering through clustering algorithms, which seems a bit out of scope for the final survey paper.

A complex clustering algorithm was presented in which the resulting variable is titled W to denote the 'within cluster' integer. Where K is the number of clusters.

Further complex mathematical algorithmic formulas are used which denotes information leading to the development of clusters to help aid in the final solution that was proposed by the researchers.

5 Experimentation and Results

Due to the continuous stream of analytical data from the clustering tree the researchers concluded on having a cut off value for the dataset of APT attacks and software attacks. The researchers evaluated the success to be at 95 percent confidence level for the analytical results.

6 Research Gaps and Questions

Minimal information based on the high level of clustering to inform uneducated-topic readers about what they are and how they can help the cyber security industry. Within the conclusion an observation on numerical data was given about how clustering aids in mapping out adversarial behaviour, this should have been reiterated in the conclusion but said near the beginning.

where did the thought of this proposed solution originate from?

7 What problem does this solve?

Statistical machine learning analysis was the result of Mitre Attack datasets being used for cluster algorithms and clustering trees to overall understand APT and software attack instances better, and also help with developing clustering techniques.

8 How do they solve it?

By proposing a solution using clustering algorithms and cluster trees on two real world datasets provided by Mitre Attack Framework with information on APT and software attack instances. This provided numerical, analytical data to conclude findings on whether the proposed solution was a success or not, the researchers provided figure images to show the results ascertained.

9 How do they assess it?

Statistical Hypothesis Testing was processed to determine the overall success of the proposed solution with respect to the initial APT and software attack datasets used at the beginning from the Mitre corporation.

10 Evaluation - Notes

An overall evaluation was considered within the paper to provide more clarity, it spoke about how the statistical hypothesis test provided validity, two methods are also proposed for a successful evaluation assessment of the research and proposition.

The evaluation visually describes the correlations between analytical measurements with regards to the clustering, showing patterns of recognition and similarity. The study explains how they used the normalized mutual information (NMI) for an evaluation then continues to outline the given steps for such an evaluation.

11 Discussion - Notes

This section discusses any findings from the researchers within their study and research to better conclude 1. a better solution and 2. better research.

"Hierarchical clustering reflects the complexities of APT and Software attacks"

was reviewed literature of that of the researcher which outlines the concise details for why we should depend on clustering for knowledge about adversarial threat actors and APTs.

Acknowledgement of the software attack instances numerical data being smaller than the APT attack data was concluded to be about how software attacks typically on average go over a small number of tactics whereas APT threats and attacks tend to (on average) span across complexity and several tactics in the Mitre Attack Framework.

12 Conclusion - Notes

The conclusion briefly explains the success of the work and the research by stating how it was able to accurately learn how to fine-grain technique associations based on real life numerical datasets from Mitre corporation.

It continues to state how APT and software instance attacks are important due to enabling the prediction of future and current adversarial behaviour to better predict outcomes and potential threats.