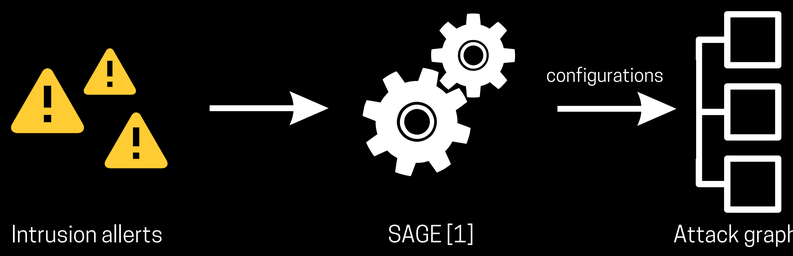
Investigating the modeling assumptions of alert-driven attack graphs

A cognitive load-based quantification approach of interpretability in attack graphs

01 Introduction

Multiple methods to generate attack graphs, but how to choose the easiest to interpret?



02 Methodology

Leverage user-experience-related concepts [2] to quantify interpretability as cognitive load:

Similarity

Similar elements are visually grouped.

Exploit the usage of popular patterns in attacker strategies.
Cluster nodes based on prior probabilities of being in the same attack path.

Continuity

The human eye will follow the smoothest path.

Quantify the ability to easily follow a path.
Intersecting edges results in paths harder to follow.
Solution: Graph maximal planarity metric.

Proximity

Elements tend to be grouped if they are close to each other.

Identify most central nodes using betweeness centrality.
Quantify the connectivity inside a cluster between central nodes and the rest of the nodes.

References:
[1] Azqa Nadeem, Sicco Verwer, and Shanchieh Jay Yang. Sage: Intrusion alert-driven attack graph extractor. In Symposium on Visualization for Cyber Security (Vizec). IEEE, 2021.
[2] Cameron Chapman. Exploring the gestalt principles of design. Toptal.

Author
Vlad-Mihai Constantinescu
Email: V.Constantinescu@student.tudelft.nl
Supervisors
Sicco Verwer
Azqa Nadeem

TU Delft

03 Interpretability metric

Baseline metric:

$$I = \text{[similarity]} + \text{[proximity]} = \frac{1}{N_c} + (1 - I_n)$$

Proposed metric:

$$I = \text{[similarity]} + \text{[proximity]} + \text{[continuity]} = \frac{1}{N_c} + (1 - I_n) + P$$

04 Results

Similar ranking results with the baseline metric

Rankings based on the two metrics shown a high level of similarity when measured using Kendall Corelance Coefficient:

Dataset	Kendall Coefficient
CPTC-2018	0.931
CPTC-2017	0.869

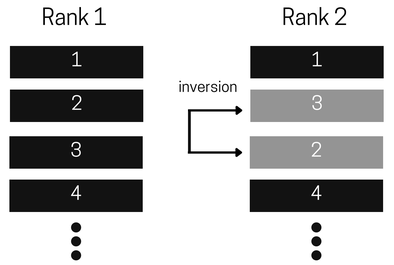


Figure 1. Inversion between two rankings.

However, the majority of rank inversions in the new ranking was considered benefic and accepted by the expert opinion as in Figure 2.

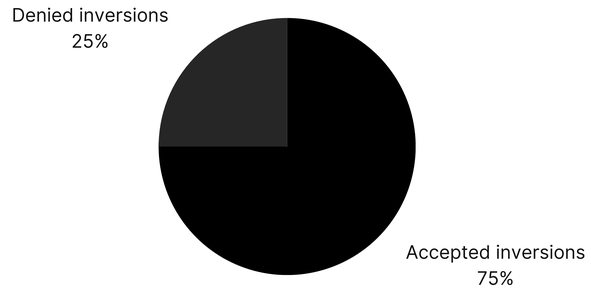


Figure 2. Manual analysis of inversions in CPTC-2018 and CPTC-2017 datasets.

Observation! *All of the denied inversions were denied because of the high similarity of the attack graphs, and not as a result of a wrong ranking.*

Comparable results among among various generation strategies

By measuring the average interpretability between attack graphs resulted from an suffix-based probabilistic deterministic finite automaton (S-PDFA), a Markov chain, and a suffix tree, we receive comparable results:

Dataset	S-PDFA	Markov Chain	Suffix Tree
CPTC-2018	0.727	0.743	0.790
CPTC-2017	0.679	0.669	0.682

Reults shows that attack graphs generated from a suffix tree are the most interpretable among the 3 specified methods.

Important! *This statement is confirmed by the expert's opinion gathered in a group discussion between members of Group 5 CSE3000-Research Project 2023.*

05 Conclusion

- Maximal planarity affect the interpretability.
- The new metric presents similar results with the baseline in terms of pariwise compaison and ranking.
- The new metric addresses discrepancies detected by experts in the baseline ranking.

06 Limitations

- The quantification of interpretability remains a subjective topic.
- The cluster algorithm requires prior knowledge.
- The proposed metric leverages only 3 our of the 7 Gestalt principles.