AUTHORS

Ioan-Cristian Oprea: i.c.oprea@student.tudelft.nl

Investigating the impact of PDFA implementation on alert-driven attack graphs.

A comparison between the Suffix-based PDFA and PDFA models

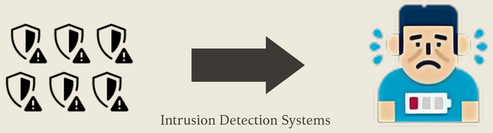
AFFILIATIONS

Responsible Professor: Sicco Verwer [1]
Supervisor: Azqa Nadeem [1]
[1] Delft University of Technology

TU Delft

01. Background knowledge

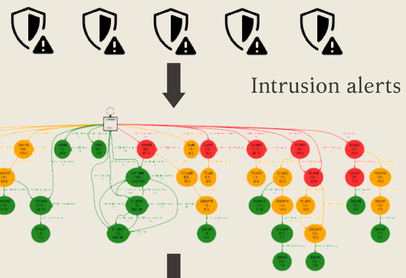
Current Problem



1 million alerts/day -> alert fatigue [1]

Current Solution [2]

- network topology knowledge
- zero-day attacks vulnerability
- expert knowledge



Intrusion alerts

S-PDFA

SAGE: [3]

Attack graphs

Current Limitations

Difficult to validate as no clear ground truth exists.

Cannot monitor attacks real-time, thus is used only for forensics

02. Aim of the research

- Try to validate SAGE by training a PDFA model and comparing it to the baseline:
  - size
  - complexity
  - completeness
  - interpretability
- Explore and give suggestions for improving the PDFA model allowing SAGE to generate attack graphs in real-time

03. Experimental Workflow

Define metrics

↓

Filter analysis candidates using size and complexity

↓

Compare attack graphs head-to-head

Datasets: CPTC-2017 & CPTC-2018

04. Methodology

Size number of nodes

Complexity

$$complexity(G) = \begin{cases} 0 & \text{if } |V| < \min \\ \frac{|E|}{|V|} & \text{if } \min \leq |V| \leq \max \\ 1 & \text{if } |V| > \max \end{cases}$$

$$isComplex(G) = \begin{cases} \text{No} & \text{if complexity(G)} < t_s \\ \text{Yes} & \text{if complexity(G)} > t_s \end{cases}$$

Completeness [4]

- Use alerts as ground truth

$$popComp(G) = \frac{\#\text{unique objectives present in AG}}{\#\text{total unique objectives}}$$

$$schemaComp(G) = \frac{\#\text{paths present in AG}}{\#\text{total individual paths}}$$

Interpretability

- Readability: time to perform protocol [5]
- Compliance with properties of SAGE

05a. Results

Size 16.5% increase over baseline (i.e. S-PDFA)

Complexity

Negligible difference => could possibly use PDFA implementation without compromising complexity

Completeness

Dataset	schemaComp	populationComp
CPTC-2017	87.33%	82.44%
CPTC-2018	88.01%	76.53%

Cause: length(episode) < 3 => discarded

No discarding

Completeness improved, but:

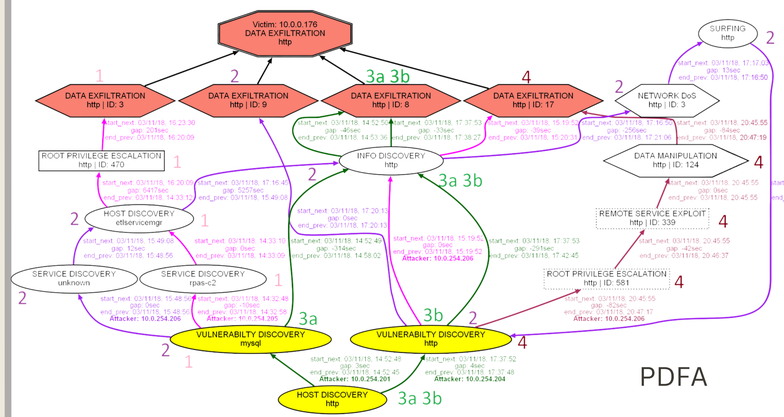
- Attack paths CPTC-2017: from 393 to 538
- Attack paths CPTC-2018: from 323 to 386

Cause: models many false positives, i.e. sub-paths of paths that already exist in attach graphs

05b. Results

Readability

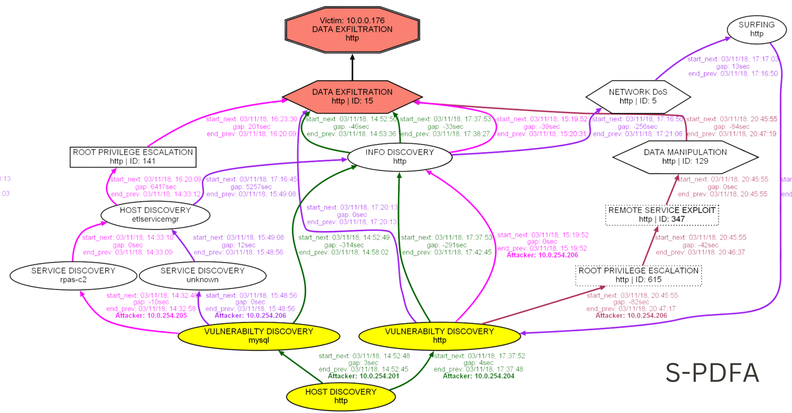
PDFA paths are more spread out => easier traversal of paths => overall increase in readability



PDFA

Compliance with properties

PDFA is better at capturing different attack strategies, but still needs work



S-PDFA

06. Limitations and Future Work

Limitations

- Manual analysis => possible human error
- Intrusion alerts used as ground truth
- Possible SAGE bugs affecting evaluation

Future Work

- Optimize PDFA to reduce the number of false positive attack strategies
- Improve alert fusion => better completeness and less false positive attack paths

07. Conclusion

Attack graphs generated with the PDFA are **more interpretable**, slightly **bigger**, and yet maintain the **same completeness**. However, it still **needs work** at capturing attack strategies

08. References

- [1] Wajih Ul Hassan, Shengjian Guo, Ding Li, Zhengzhang Chen, Kangkook Jee, Zhichun Li, and Adam Bates. Nodoze: Combatting threat alert fatigue with automated provenance triage. Network and Distributed Systems Security Symposium.
- [2] Steven Noel, Matthew Elder, Sushil Jajodia, Pramod Kalapa, Scott O'Hare, and Kenneth Prole. Advances in topological vulnerability analysis. In 2009 Cybersecurity Applications Technology Conference for Homeland Security, pages 124–129, 2009.
- [3] Azqa Nadeem, Sicco Verwer, Stephen Moskal, and Shanchieh Jay Yang. Alert-driven attack graph generation using s-pdfa. IEEE Transactions on Dependable and Secure Computing, 19(2):731–746, 2022.
- [4] Subhi Issa, Onaopepo Adekunle, Fayçal Hamdi, Samira Si-Said Cherfi, Michel Dumontier, and Amrapali Zaveri. Knowledge graph completeness: A systematic literature review. IEEE Access, 9:31322–31339, 2021.
- [5] M. Ghoniem, J.-D. Fekete, and P. Castagliola. A comparison of the readability of graphs using node-link and matrix-based representations. In IEEE Symposium on Information Visualization, pages 17–24, 200

All icons used in this poster were collected from https://icons8.com/