

Detecting Unsafe Updates in Software Ecosystems

Yao-Wen Chang

September 22, 2023

Supervised by: Christoph Treude

Co-supervised by: Behnaz Hassanshahi (Oracle Labs, Australia)

BACKGROUND

- ▶ **SolarWind's Orion** platform is polluted.
- ▶ Malicious action against **Eslime-Scope**.

OBJECTIVE

- ▶ Introduce a new framework.
- ▶ Evaluate the effectiveness of the framework.
- ▶ Automate the unsafe update explore process.

Research Questions

1. *What is the scope of the impact of the risks that exist within our target Python and Java repositories?*
2. *What are the results if these suspicious updates from contributors in open source projects compromise the target?*
3. *To what extent does this work enhance the security of CI/CD pipelines based on the findings and recommendations from our research?*

OUTLINE

- ▶ Related Works
- ▶ Research Methods
 - ▶ Data Source
 - ▶ Framework
 - ▶ Pipeline
 - ▶ Metric
- ▶ Summary
- ▶ QA

RELATED WORKS

Regex

Introduce papers....

Machine Learning

Introduce papers....

CI / CD Based

Introduce papers....

RESEARCH METHODS I

Data

► Fetch Data

► Clean Data

Contributors 401



+ 390 contributors

Deployments 198

✓ github-pages 3 years ago

+ 197 deployments

Languages



► Third item

vinta Merge pull request #2498 from arunachalamev/add-lighting		c532ba9 on Jul 14	1,644 commits
github	cleanup		3 years ago
docs	Removed dead css		4 years ago
gitignore	Sort readme and add to docs build		3 years ago
.travis.yml	Sort readme and add to docs build		3 years ago
CONTRIBUTING.md	Update CONTRIBUTING.md		2 years ago
LICENSE	add LICENSE Fixes #328		8 years ago
Makefile	update Makefile		4 years ago
README.md	added deep learning framework Lighting		2 months ago
mkdocs.yml	update mkdocs.yml		4 years ago
requirements.txt	add requirements.txt		4 years ago
sort.py	Sort readme and add to docs build		3 years ago

RESEARCH METHODS II

Framework

Introduce papers....

Pipeline

Introduce papers....

RESEARCH METHODS III

Metric

$$SM = (W_p \cdot P) * (W_{tf} \cdot TF) * \left(\frac{W_{tc}}{TC}\right) * (W_s \cdot S) \quad (1)$$

SM = Security Scanner Metric

P = Precision (as a decimal)

W_p = Weight for Precision

TF = Total Findings (TP + FP)

W_{tf} = Weight for Total Findings

TC = Time Cost

W_{tc} = Weight for Time Cost

S = Normalized Severity Score

W_s = Weight for Severity

SUMMARY

This is the first slide of your presentation.

REFERENCES I