PhD Open Days

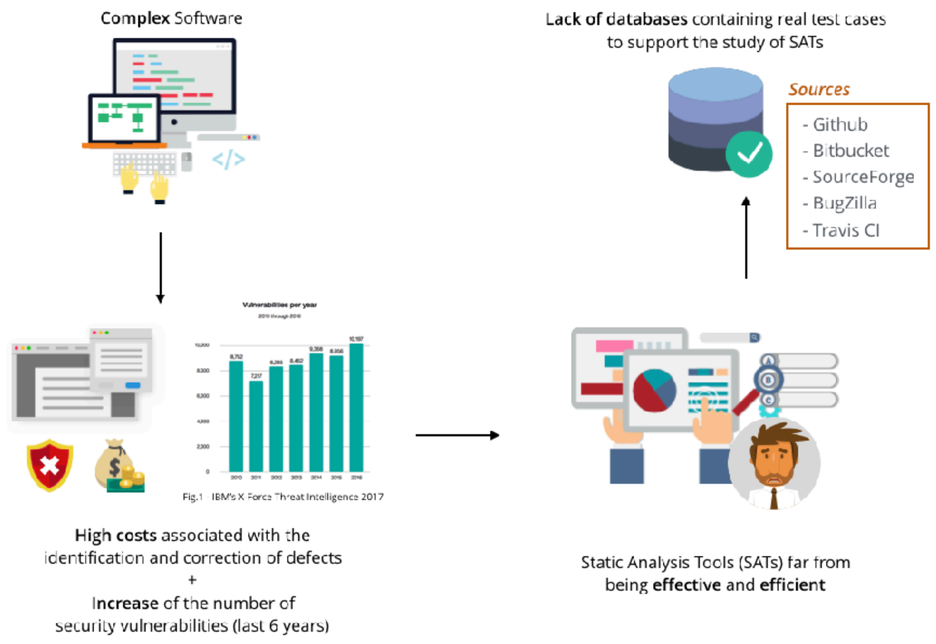
Leveraging Known Vulnerabilities to Modernize Static Analysis Tools

PhD Program in Information Systems and Computer Engineering

Sofia Oliveira Reis (sofia.o.reis@tecnico.ulisboa.pt)

Introduction

Currently, to satisfy the high number of system requirements, complex software is created which makes its development cost-intensive and more susceptible to security vulnerabilities.



These studies are challenging due to the lack of widely accepted and easy-to-use databases of real vulnerabilities [1], as well as the fact it requires human effort and CPU time [2]. Consequently, researchers tend to use databases of hand-seeded vulnerabilities which differ inadvertently from real vulnerabilities, and which according to some researchers may not be a feasible approach under specific conditions [3]. These tools have a high potential of integrating the CI/CD pipeline and other than help to reduce the costs aggregated to software production; they can also help companies making a more confident tools choice.

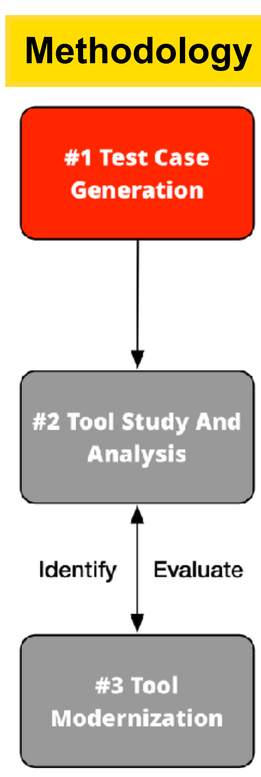
Motivation

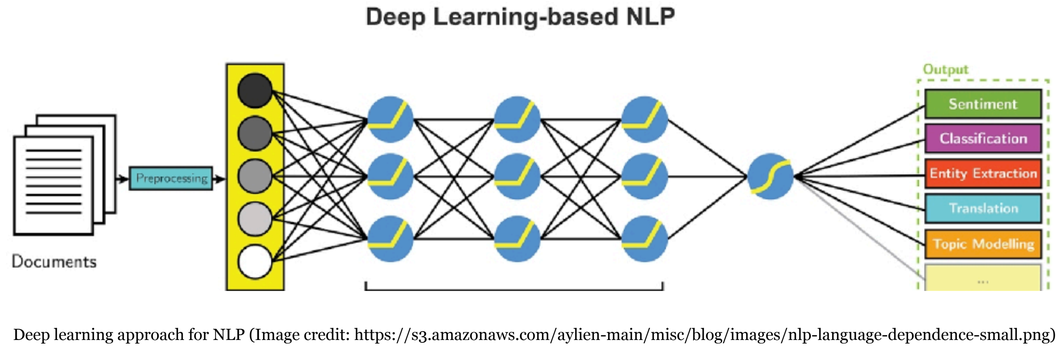
Static Analysis Tools

Focus: Only tools able to identify security vulnerabilities

More than **20 open-source tools**

Methodology





Deep learning approach for NLP (Image credit: https://s3.amazonaws.com/aylien-main/misc/blog/images/nlp-language-dependence-small.png)

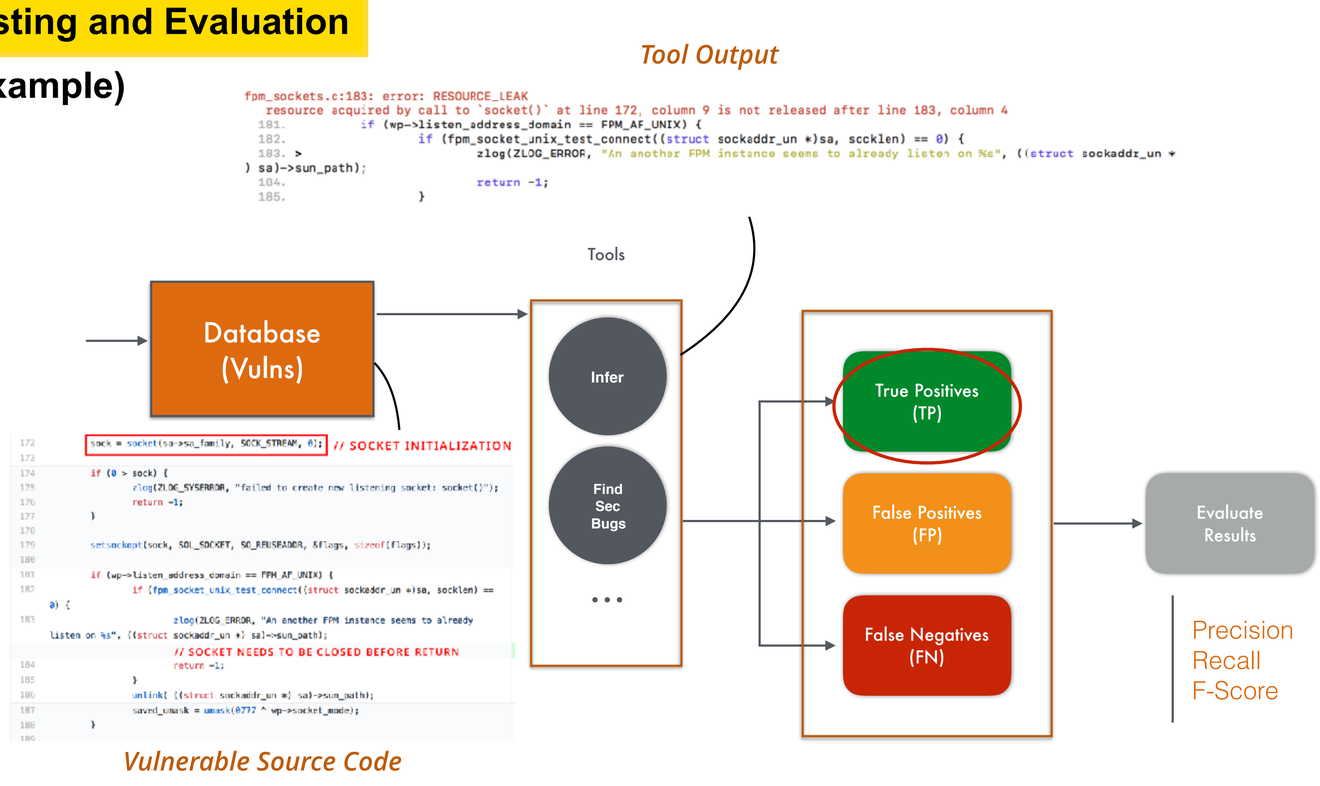
Forward Looking

The next step is to improve the mining approach used to create the secbench database [4]. Now, the focus will be the creation of a deep-learning approach applied to natural processing languages (NLP) using named entity recognition (NER) to detect indications of vulnerabilities in the commits' messages of open-source software repositories and minimize the percentage of false positives (messages containing misleading evidences).

Example of a real vulnerability-fixing commit.

Testing and Evaluation

(Example)



References

[1] René Just, Darioush Jalali, and Michael D. Ernst. “Defects4j: A database of existing faults to enable controlled testing studies for java programs.” (2014).
[2] L. C. Briand. “A critical analysis of empirical research in software testing.” (2007)
[3] René Just, Darioush Jalali, Laura Inozemtseva, Michael D. Ernst, Reid Holmes, and Gordon Fraser. “Are mutants a valid substitute for real faults in software testing?” (2014)
[4] Reis, Sofia, and Rui Abreu. "SECBENCH: A Database of Real Security Vulnerabilities." (2017).

TÉCNICO LISBOA

Supervisor: Rui Abreu

PhD Program in Information Systems and Computer Engineering

phdopendays.tecnico.ulisboa.pt