POSTER: ‘False negative - that one is going to kill you’: Understanding Industry Perspectives of Static Analysis based Security Testing

Amit Seal Ami*, Kevin Moran†, Denys Poshyvanyk*, and Adwait Nadkarni*
*William & Mary, Williamsburg, VA, USA; aami@, denys@cs., nadkarni@cs.wm.edu
†University of Central Florida, Orlando, FL, USA; kpmoran@ucf.edu

Title: ‘False negative - that one is going to kill you’: Understanding Industry Perspectives of Static Analysis based Security Testing

Authors: Amit Seal Ami, Kevin Moran, Denys Poshyvanyk, Adwait Nadkarni

Venue: 2024 IEEE Symposium on Security and Privacy (SP), San Francisco, CA, USA, 2024

DOI: 10.1109/SP54263.2024.00019

Full Reference: A. Ami, K. Moran, D. Poshyvanyk and A. Nadkarni, “‘False negative - that one is going to kill you.’ - Understanding Industry Perspectives of Static Analysis based Security Testing,” in 2024 IEEE Symposium on Security and Privacy (SP), San Francisco, CA, USA, 2024 pp. 19-19, available at: https://doi.ieeecomputersociety.org/10.1109/SP54263.2024.00019

Abstract: The demand for automated security analysis techniques, such as static analysis based security testing (SAST) tools continues to increase. To develop SASTs that are effectively leveraged by developers for finding vulnerabilities, researchers and tool designers must understand how developers perceive, select, and use SASTs, what they expect from the tools, whether they know of the limitations of the tools, and how they address those limitations. This paper describes a qualitative study that explores the assumptions, expectations, beliefs, and challenges experienced by developers who use SASTs. We perform in-depth, semi-structured interviews with 20 practitioners who possess a diverse range of software development expertise, as well as a variety of unique security, product, and organizational backgrounds. We identify 17 key findings that shed light on developer perceptions and desires related to SASTs, and also expose gaps in the status quo - challenging long-held beliefs in SAST design priorities. Finally, we provide concrete future directions for researchers and practitioners rooted in an analysis of our findings.

Pre-print: https://arxiv.org/pdf/2307.16325.pdf

WILLIAM & MARY
CHARTERED 1693

"False negative - that one is going to kill you"
Understanding Industry Perspectives of Static Analysis-based Security Testing

Amit Seal Ami[§], Kevin Moran[†], Denys Poshyvanyk[§] and Adwait Nadkarni[§]
[§] William & Mary, [†]University of Central Florida
{aami@, denys@cs, nadkarni@cs.} wm.edu, kpmoran@ucf.edu
https://amitsealami.com

Secure Platforms Lab
https://spl-wm.github.io/

SEMERU
http://www.cs.wm.edu/semeru/

Introduction

- **Increasing Reliance on program analysis for security,** e.g., SASTs
- **SASTs** suffer from **design and implementation flaws**
- We have a **Limited Understanding of the Perceptions, Expectations, and Beliefs** regarding SASTs

Research Questions

- **How are SASTs chosen** at organizations with different business and security needs?
- **What do practitioners know and believe** about the limitations of SASTs?
- How do practitioners navigate, address, or **work around flaws of SASTs?**

Findings

Selecting SASTs

F04 - Participants generally recall selecting SASTs due to factors such as recommendations/reputation, ease of use/integration, corporate pressure, cost, or compliance requirements. Only one participant selected a SAST for their product via exhaustive testing of 10-15 tools using a (custom) benchmark.

> "We didn't evaluate that many tools in terms of static analysis tools. We take what is the industry standard across different companies. Like <tool> is pretty popular, so that is our first choice."
> - P08 Media, Web and Back-end services

F05 - Participants who are aware of benchmarks generally do not trust them for evaluating/selecting SASTs, viewing benchmarks as either too basic to model real problems, or biased towards specific SASTs, given that vendors often contribute to their construction.

> *"Quite a few of these benchmarks are created by tool vendors where their tool finds some specific edge case. No one in the right mind would write an application like this, but their tool finds a specific edge case, so they put it in the benchmark."*
> - P01, Program Analysis for Security

Reducing False Negative vs False Positive

F10 - Nearly all the practitioners expressed a preference for fewer false negatives, i.e., as long as the SAST is able to find valid security vulnerabilities, they would tolerate and even prefer few false negatives at the cost of many false positives.

> *"False negative for sure.I just told you the amount of the price of the bug (in millions), so I don't care if there are 10 false positives. False negative - that one is going to kill you."*
> - P04 Automobile Sensors
>
> *"If you're getting a bunch of false positives, then that typically means your static code analysis tool is doing its job. …I'd rather my security tool be annoying and tell me about every single possible issue over it not telling me anything and just letting security things slide through."*
> - P14 Law Enforcement

F11 - Practitioners are generally more tolerant of false positives than the 20% upper bound proposed in literature, given their preferences and the tools they currently use, with some finding even 80% or more false positives practical.

> *"I wouldn't mind wading through false positives, if I thought there were actully going to be genuine issues there"*
> - P02 OSS Java App Server

Effective False Positives and SAST

F12 - Practitioners are generally against letting developers define "effective" false positives, or letting them decide when to run SASTs. This reservation stems from their prior experience of the adverse cost of leaving a vulnerability in the code, and/or from their knowledge of developers.

> *"It seems familiar, but it may be new. And then you're just going to ignore it because it's close enough to something you've seen in the past, and you just say that it's OK. So we do need to be vigilant on those false positives to make sure that they are truly false positives"*
> - P14 Law Enforcement

Addressing/Reporting flaws to SASTs

F15 - Participants may hesitate to report flaws/false negatives for several reasons, e.g., prior negative experiences with SASTs (including inaction on reported flaws) or issues internal to the organization, such as the need to maintain product confidentiality (without an explicit NDA), red tape, and the lack of incentive to perform the additional effort.

> *"So, <SASTA>, we have a worse experience. They are mostly evasive, so they are not really progressing as <SASTB>. It takes a lot of time to convince them that they are bugs. Even though you have a small example, they still ask you to try different configurations and all that stuff, but we were aware of that before we came to this part, before we selected them. Because simply they (<SASTA>) are, I wouldn't say confident, but they are confident that their solution works."*
> - P04 Automobile Sensors

Exploiting Flaws and Evasive Developers

F15 - The risk of evasive developers is real. That is, while some participants consider the scenario of "evasive developers" as adequately prevented by existing code reviews, this optimism is not universal: others have prior experience of evasive developers in their teams, or have evaded SASTs themselves.

> *There was an extreme pressure because we needed to bypass the SAST tests, otherwise we would not receive green flag from the security team. So it actually happened once. We used to work late night to resolve all those conflicts and red flags."*
> - P06 Software Service

Methodology

Survey

Purpose

1. Gain ideas about the Landscape
   - Prioritizing security by organizations
   - Prioritizing security by Individuals
   - Reliance on Automated techniques
   - Reliance on Manual Techniques
   - Impact due to Unsound SASTs
2. Guiding the design of the Interview Protocol

Recruitments

1. Snowball sampling from Professional Network
2. OSS developers from GitHub who interacted with SASTs via CI/CD

Numbers & Insights

- 39 valid responses from 2000+ invited
- Organizations and their practitioners can have different priorities for security
- Impact of flaws in SASTs taken lightly

Interview

Purpose

Learn about

- Participant, project, and organization
- Security and Organization
- Organizational Context of SASTs
- Limitations and Expectations about SAST
- Impact of Unsound SAST
- Challenges and Improvements relating to SASTs

Participants

20 participants
- Spread across Asia, Europe, the United Kingdom, and North America
- Working in Local or International Projects
- Diverse Industry Contexts
  - Safety critical systems
  - Business critical systems
  - R&D
  - Open-source software
- Experience ranging from entry-level to project manager
- Diverse Security contexts, e.g., compliance

META

- Semi-Structured Interview
- Menlo Report Guideline for Sensitive Questions
- Single Coder Approach
- Lead Interviewer/ Lead-and-Backup Interview
- **187,000** words transcribed
- Reflexive Thematic Analysis with Inductive Coding

Takeaways

Practitioners tend to rely on ad-hoc/subjective criteria.
A solid preference for security is universally expressed.
BUT, that is not reflected in the selection process of SASTs.

Unreasonable Optimism and Trust
in Reputation of SASTs dissuade practitioners from evaluating SASTs

Benchmarks are insufficient
Developers interested in evaluating SASTs lack the means.

Developers **want** (a) ease of use and
(b) SASTs **to detect vulnerabilities**

Paradoxical Takeaway - Industry is not Ready for Flaws of SASTs

1. Practitioners use SASTs to cover blind spots, subjective bias, and knowledge gaps in manual analysis, i.e., *detect vulnerabilities they are unaware of.*
2. Practitioners believe a flawed SAST won't impact them because they can find the remaining vulnerabilities using manual analysis.

References

A. S. Ami, K. Moran, D. Poshyvanyk, and A. Nadkarni, *"'False negative - that one is going to kill you' - Understanding Industry Perspectives of Static Analysis based Security Testing,"* in Proceedings of the **2024 IEEE Symposium on Security and Privacy (S&P),** San Francisco, CA, USA, May 2024.

LEARN MORE

Coastal Virginia

This work is supported in part by the NSF-1815336, NSF CNS-2132281, NSF CNS-2132285, CCF-1955853 grants, and CoVA CCI Dissertation Fellowship.
Any opinions, findings, and conclusions expressed herein are the authors' and do not necessarily reflect those of the sponsors.