SAND2016-6694PE

Exceptional service in the national interest

Sandia National Laboratories

Evaluate Static Code Analyzers

Open Source vs. COTS

Aretzby Bonilla

Albuquerque Institute of Mathematics and Science

Intended Major: Business Administration, May, 2017

Mentors:

Michael Coram (9511)

Michael McDaniel (6832)

Sandia National Laboratories

July 26th, 2016

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 2016-475863

ABSTRACT

- Background
  - Software testing and code quality improvement are vital, but often overlooked, components of software development. Multiple tools, both free/open source and commercially available, can be utilized to automate testing and improve software quality.
- Objective
  - Determine how the performance of a free open source tool compares to that of commercial tools.
- Description
  - Compare results of three static analyzers—one open source and two commercial—run on the same code to determine open source tool performance based on issues found, overlap and severity.

ABSTRACT, continued

Sandia National Laboratories

- Major Finding
  - The open source tool found roughly the same amount of issues as commercial tool A, but found more issues than commercial tool B.
  - 33% of the issues detected by the open source tool overlap with commercial tool A, while only 18% of the issues detected by the open source tool overlap with the other commercial tool B.
- Major Conclusion and Results
  - Some tools overlapped issues but some were unique and not detected at all.

MM33

  - To find the greatest number of issues, multiple tools may need to be run.

MM34

  - The open source tool found the highest number of issues that required fixing, but didn't detect them all.

Slide 3

MM33 not sure what this means. Do you mean, "unique and only detected by one of the three tools"?
Mcdaniel, Michael, 7/6/2016

MM34 Rewording of above bullet, only keep one of the two.
Mcdaniel, Michael, 7/6/2016

INTRODUCTION

- To improve software, it is important to find a tool that locates issues and has high performance level.
- Tools, like static code analyzers, allow developers to conduct automated testing to find potential errors (bugs).
- Findbugs is an open source tool; Klocwork and Coverity are commercial tools. The experiment compares the performance of FindBugs to that of Klocwork and Coverity.
- The analysis was run on a custom, stand-alone Java program consisting of 24,383 lines of code and 369 classes.
- Hypothesis: commercial tools will find more significant issues than the open source tool.

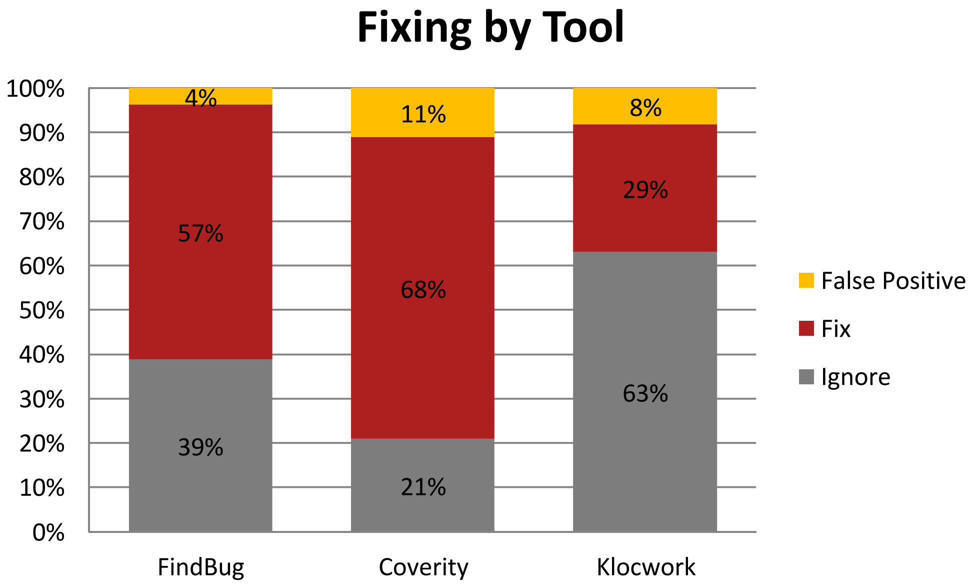
METHODS

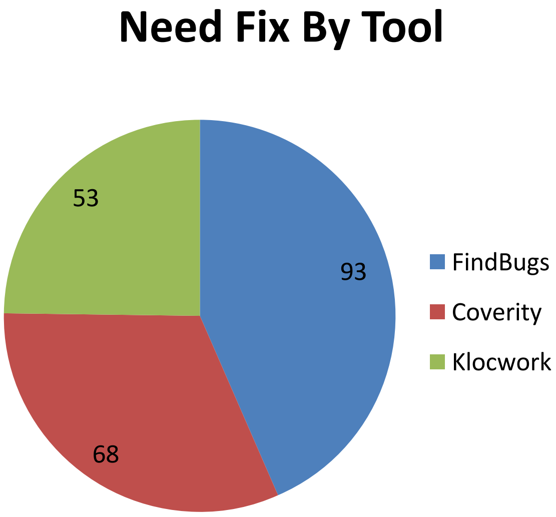
1. Run static code analyzers on software source code
2. Convert the results from XML to Excel
3. Based on the file name, line number, and severity (SME assessment needed) find the overlap between tools, and analyze the required resolution for all issues detected
4. Determine which tool is better based on severity, uniqueness, and amount of overlap

RESULTS

- FindBugs (162) found almost same number of bugs as Klocwork (184), but more than Coverity (100)
  - 9.54% of the total issues between FindBugs and Klocwork overlapped, only 22.14% of those that FindBugs and Coverity found, and 5.28% of those found by Coverity and Klocwork
- FindBugs, the open source tool, was able to find the most issues that needed fixing.

RESULTS

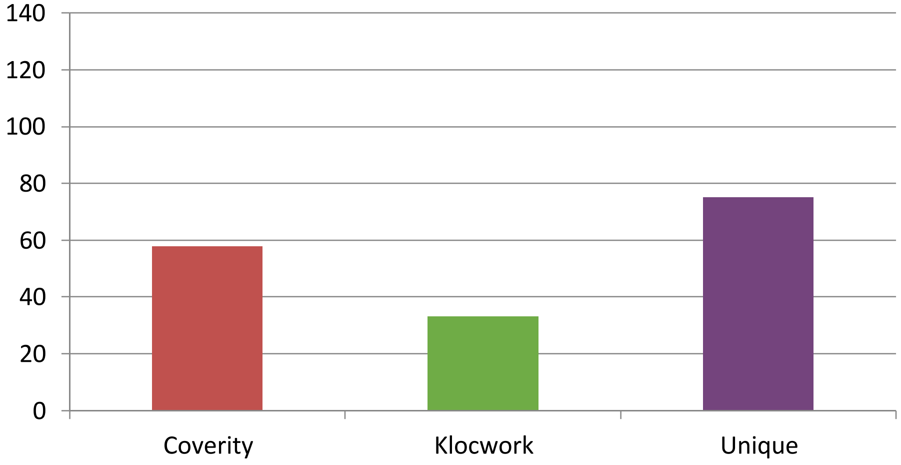




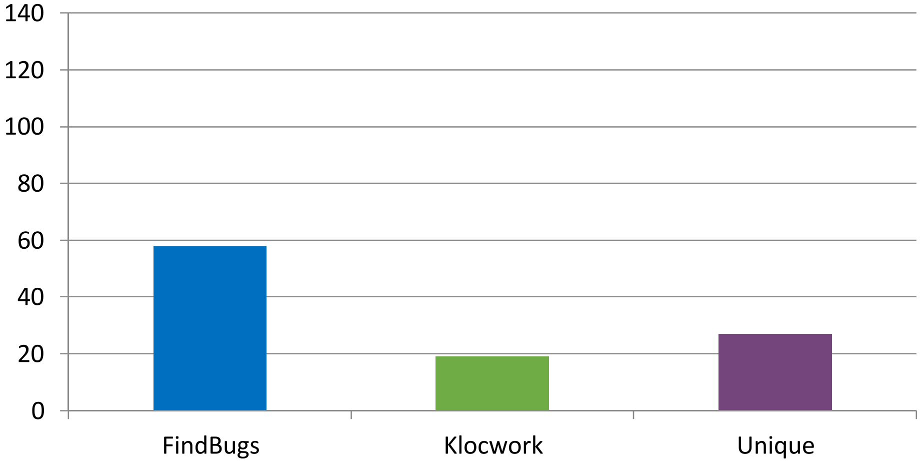
	Ignore	Fix	False Positive	Total
FindBug	63	93	6	162
Coverity	21	68	11	100
Klocwork	116	53	15	184

RESULTS

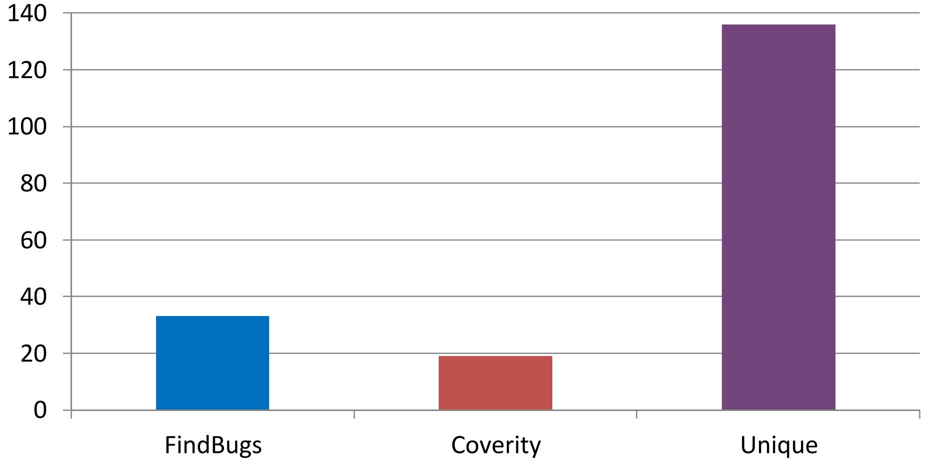
FindBugs Overlapping Issues



Coverity Overlapping Issues



Klocwork Overlapping Issues



DISCUSSION

- Major Findings
  - Due to lack of overlap between the tools, one tool may be insufficient to locate all issues.
  - Based on how severe the tool deemed the issue, FindBugs found more issues that require fixing
- The open source tool, Find Bugs, outperformed the commercial products based on the number of issues and their severity.
- Future Research: further evaluation on other tools to see which tool is better or if multiple tools is best to find all significant issues.

THANK YOU

- This work was funded in part by NNSA NA-243, Office of Nuclear Verification
- This project was funded by the STAR Summer Fellowship Program