

Improving Run-Time Bug Detection in Aviation Software Using Program Slicing

Abstract

We present an improvement to run-time bug detection by using program slicing instead of a heuristic to identify relevant program variables to monitor. Our results show that selecting variables using the program slice produces higher sensitivity compared to selecting variables heuristically. Moreover, at least in the experiments conducted, program slicing resulted in more consistent detection, with lower variance in sensitivity as compared to the heuristic approach.

Introduction

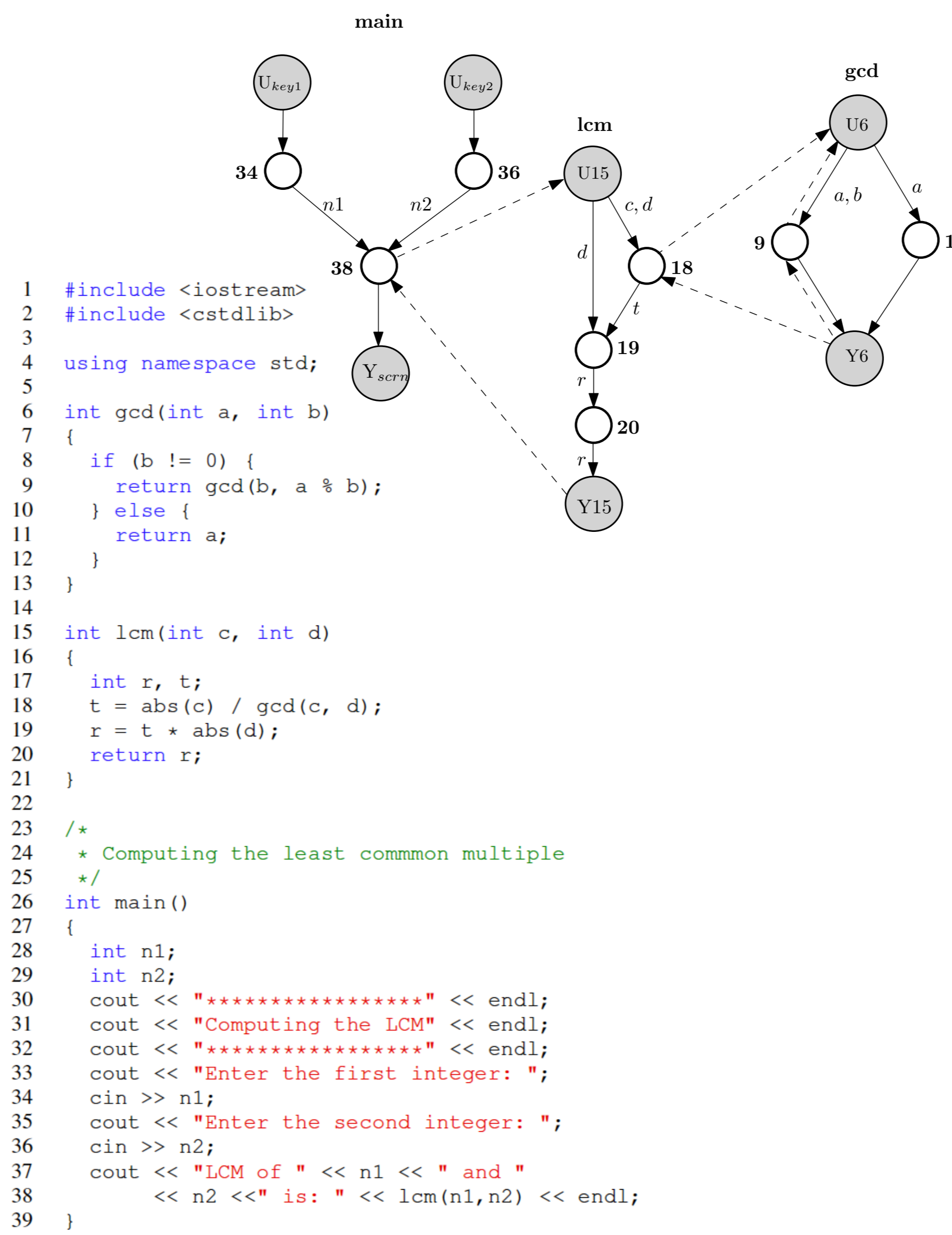
- We address the question how to choose monitoring variables to detect bugs in software by leveraging program slicing, a well-known static analysis technique originally proposed by Weiser¹.
- Our hypothesis is that variables chosen using program slicing result in better bug detector performance than variables chosen through a reasonable heuristic.
- We construct bug detectors using a standard machine learning algorithm, the One-Class SVM. We use the One-Class SVM to conduct two computational experiments in order to test our hypothesis. The experiments, which consider a simplified but representative flight control code, address two questions.

Research Questions

Question 1: For the same level of specificity, do instrumentation variables chosen using a program slice, as opposed to a heuristic, result in higher monitor sensitivity for bugs injected at different locations in the model code?

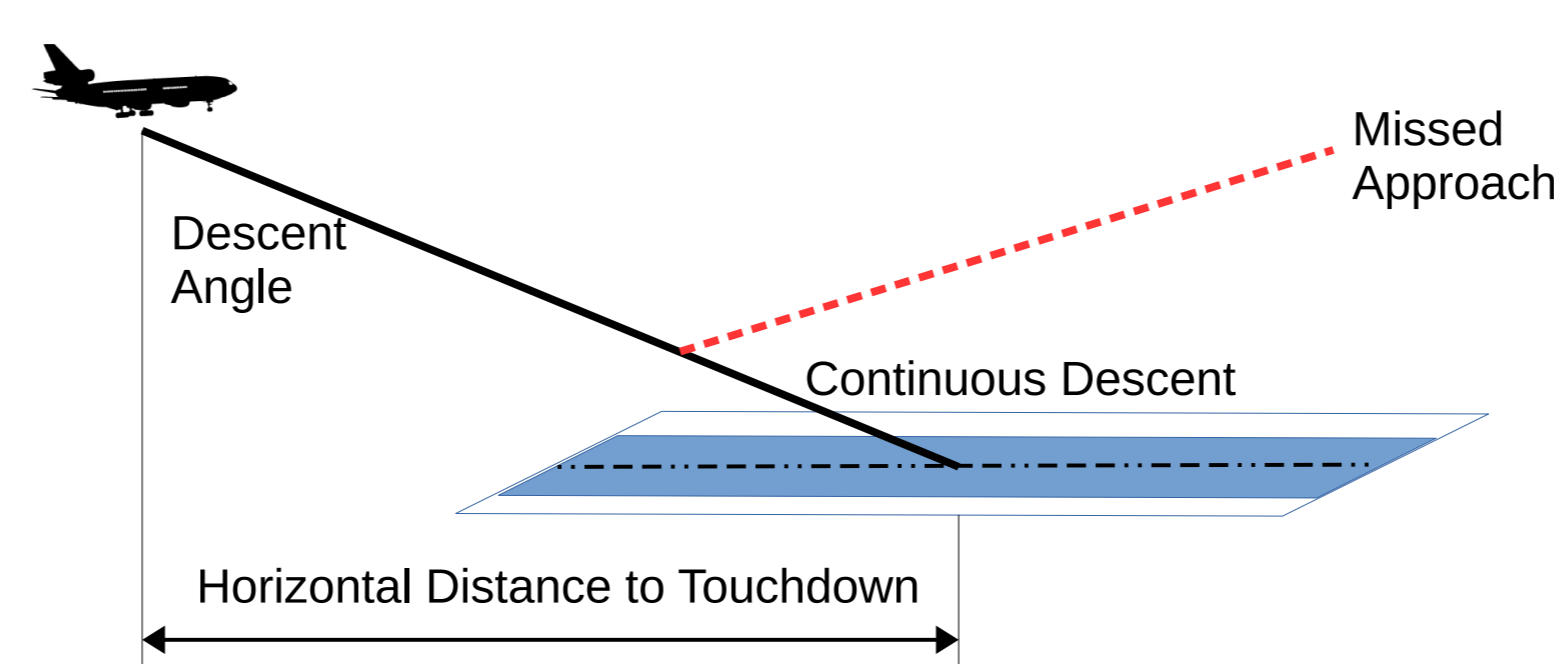
Question 2: Is the benefit of selecting instrumentation variables using program slicing enhanced as bug magnitude gets smaller, indicating greater sensitivity when detection is hardest?

Background



Methodology

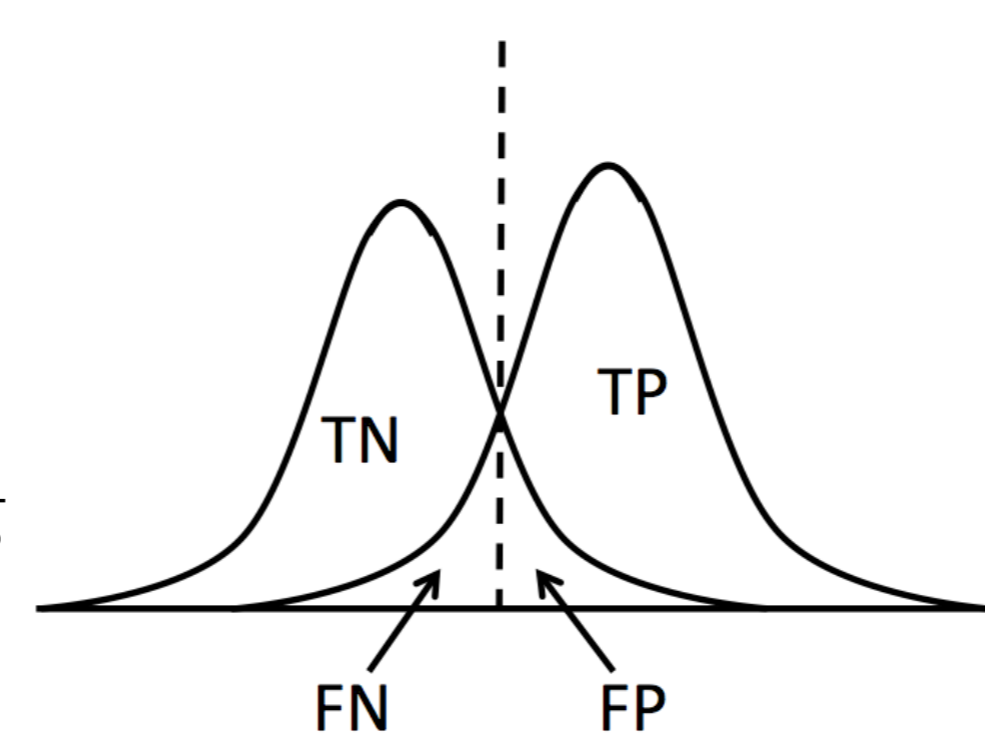
Operational Scenario:



Evaluation Metric:

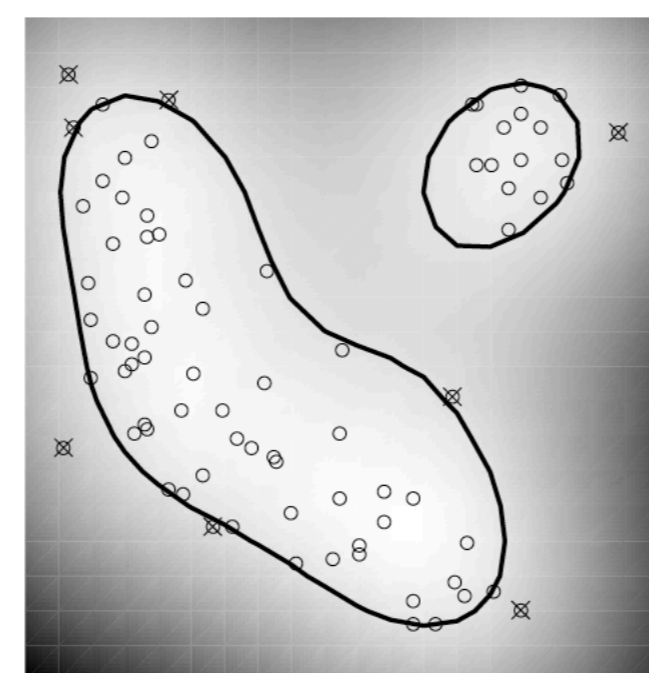
$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

$$\text{Specificity} = \frac{TN}{TN + FP}$$



Anomaly Detection Model:

One-Class SVM



Kernel = RBF

$\gamma = 0.001$

$\nu = 0.2$

Results

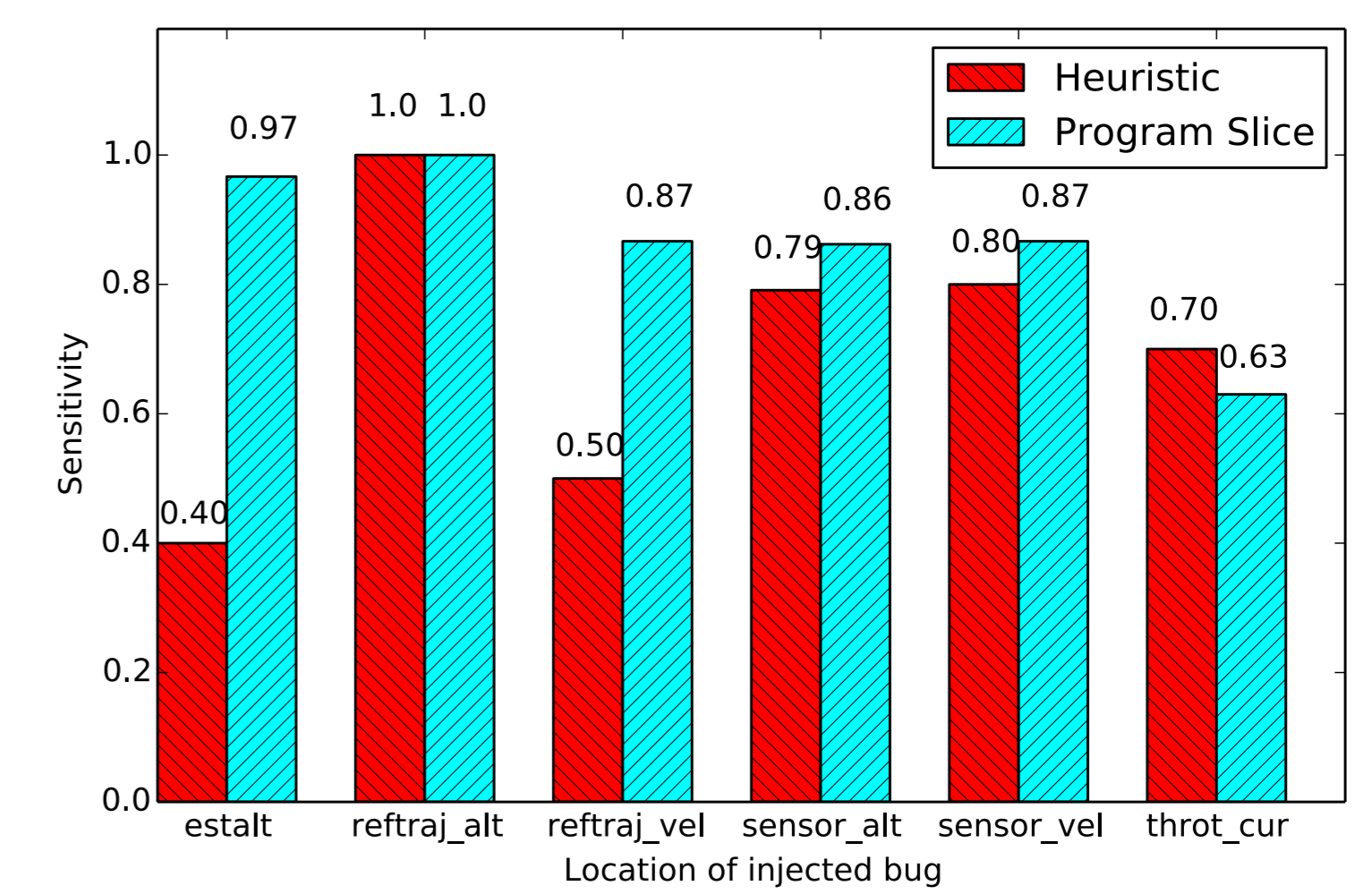


Figure 3. Sensitivities of both heuristic and program slice detectors for the six local variables where bugs were injected (higher is better). While the program slice performs better across four of the six variables, it does not always perform better. The program slice detector actually performs slightly worse for throt_cur and ties for reftraj_alt. For all the bug injection sites, the specificity was set to 0.8.

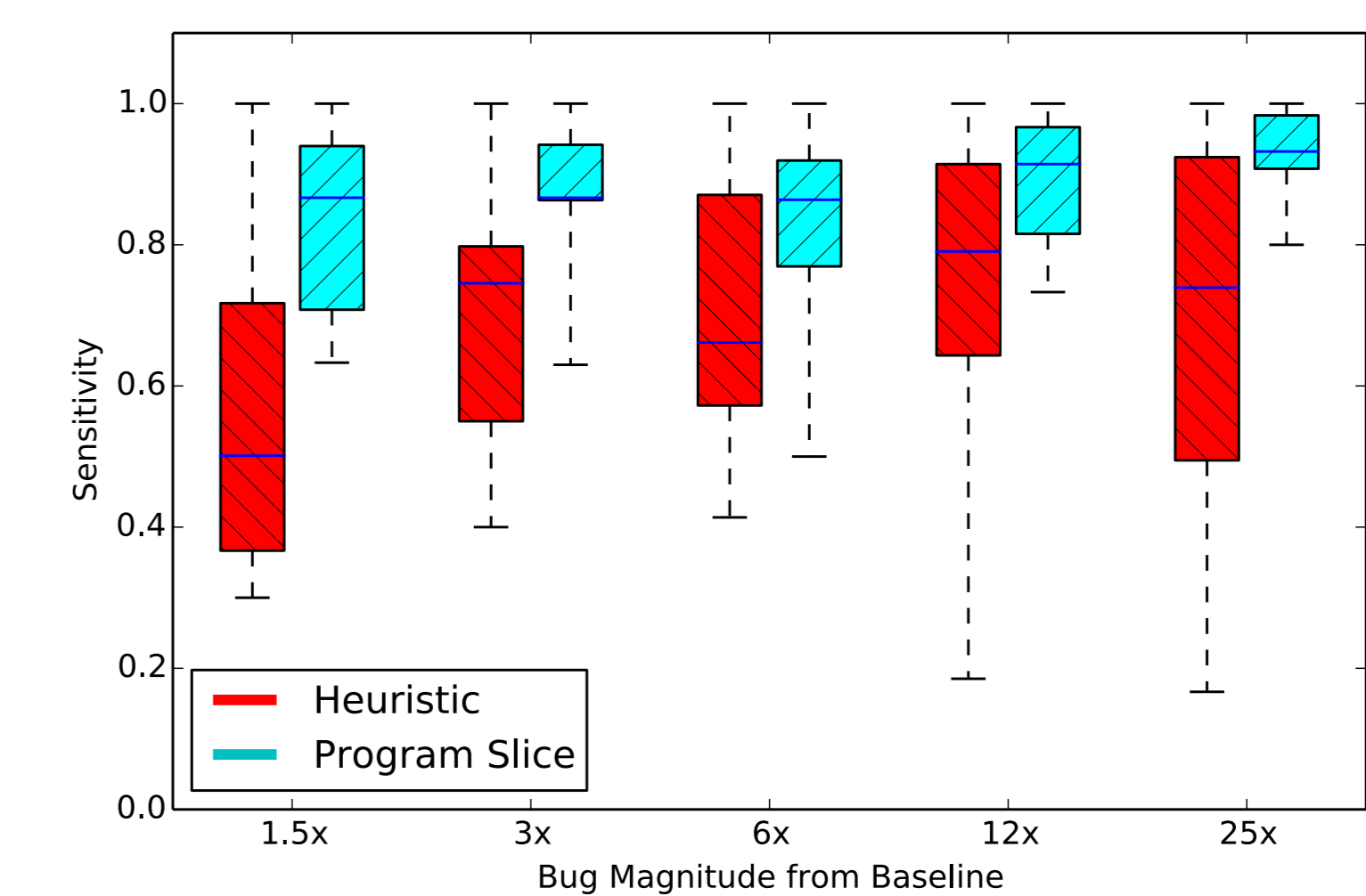


Figure 4. Box plots for Experiment 2, where higher on the vertical axis corresponds to better sensitivity. With larger bug magnitudes, both the heuristic and the program slice detectors have relatively high sensitivities. At lower bug magnitudes, the program slice detectors perform significantly better.

Future Work

We are currently preparing a manuscript to be submitted to TCPS journal that expands on the theoretical notions of the program slice for a CPS.

In addition, we are currently building compiler infrastructure to automate program slicing and will be applying this technique on real-world code bases such as Ardupilot.

Acknowledgement

This research was supported by National Science Foundation under grant CNS-1329341

Contact

Hu Huang
Tufts University
Email: hu.huang@tufts.edu

References

- Mark Weiser. 1984. Program Slicing. IEEE Transactions on Software Engineering 4, SE-10 (1984), 352–357.
-