

Critique on

Baishakhi Ray, Vincent Hellendoorn, Saheel Godhane, Zhaopeng Tu, Alberto Bacchelli, and Premkumar Devanbu. 2016. On the "naturalness" of buggy code

Astha Sharma
asharma6@uno.edu

SUMMARY

According to the given research paper, softwares that solves real world problems are natural - highly repetitive and predictable. It also lives by a hypothesis that a code that is unlikely or surprising to a good statistical language - language used to define naturalness of softwares - is unnatural, and therefore doubtful. To verify this hypothesis, an experiment was conducted with large corpus of bug fix commits from 10 different java projects and had focused on the language statistics for evaluating naturalness of buggy codes and corresponding fixes. The findings that were derived from the study are as listed below.

- Codes with bugs are more unnatural (entropic)
- Entropy lessens as the bugs are fixed
- The defect prediction method can be identified from the cost-effectiveness score that is obtained by ordering files based on their average entropy.
- Cost-effectiveness is similar in both case:
 - ❑ using known bug finders and ordering warnings using entropy measures
 - ❑ focusing on highly entropic lines of code.

STRENGTHS

1. The research consist of large set of real-time data giving more precise findings.
2. The research was more focused on providing more fine and granular prediction than a typical statistical defect prediction method.
3. The projects that were used in the research included varying domains from github to Apache Software Foundation.
4. Evaluation of the bug classification tool to ensure whether or not the keyword based search identifies the buggy commits.

WEAKNESSES

1. Larger data set should have been tedious and time consuming.
2. This research has been based on Java programming language but does not verify if the results hold true for every language, in general.
3. The paper does not tell us how the developers can actually use any measures to aid their bug finding process.