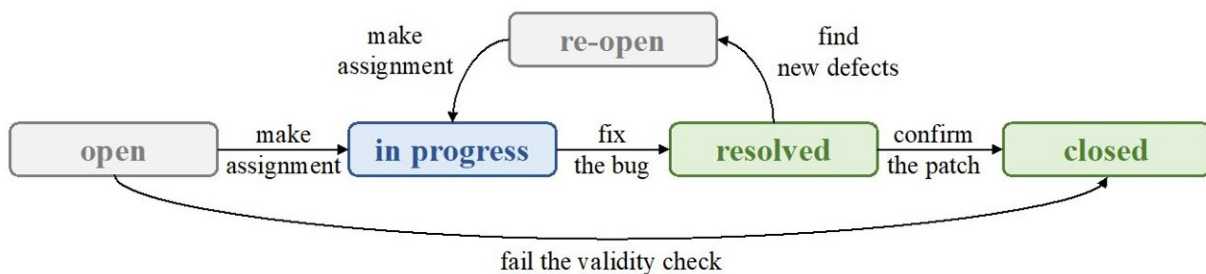


Without texts, automatic bug assignment still works well: Study

August 26 2024



Credit: *Frontiers of Computer Science* (2024). DOI: 10.1007/s11704-024-3299-6

Automatic bug assignment has been well studied in the past decade. As textual bug reports usually describe the buggy phenomena and potential causes, engineers highly depend on these reports to fix bugs. Researchers heavily depend on the textual content in the bug reports to locate the buggy files. However, noises in texts bring adverse impacts to automatic bug assignments unexpectedly, mainly due to insufficiency of classical Natural Language Processing (NLP) techniques.

To acquire a deep understanding on the effects of textual features and nominal features, a research team led by Zexuan Li published their [research](#) in *Frontiers of Computer Science*.

The team reproduced an NLP technique, TextCNN, to learn whether

improved NLP technique can lead to better performance for textual features. The results reveal that textual features do not surpass other features even with the relatively advanced technique. The team further explore the influential features for bug assignment approaches and give an explanation from a statistical perspective.

They find that the influential features selected are all nominal features that indicate the preference of developers. Experimental results show that nominal features can achieve competitive results without using text.

In the research, they made efforts to answer three questions. First, how effective are textual features with deep-learning-based NLP techniques? They reproduce TextCNN and compare the effectiveness of textual features with the group of nominal features.

Second, what are influential features for bug assignment approaches and why are they influential? They employ the wrapper method and the widely-used bidirectional strategy. By repeatedly training a classifier with different groups of features, it judges the importance of features according to the metric. They speculate that nominal features can contribute to reducing the search scope of the classifier and verify the speculation in a statistical method.

Third, to what extent can the selected influential features make improvements on bug assignments? They train models with fixed classifiers on changing groups of features and conduct two popular classifiers (Decision Tree and SVM) on five groups of features.

The experiment used five projects in different sizes and types as datasets. The results demonstrate that improved NLP technique has limited improvement and the selected key features achieve 11–25% accuracy under two popular classifiers.

Future work can focus on introducing source files to build a knowledge graph between these influential features and descriptive words for better embedding of nominal features.

More information: Zexuan Li et al, Automatic bug assignments without texts: a study, *Frontiers of Computer Science* (2024). [DOI: 10.1007/s11704-024-3299-6](https://doi.org/10.1007/s11704-024-3299-6)

Provided by Frontiers Journals

Citation: Without texts, automatic bug assignment still works well: Study (2024, August 26) retrieved 1 September 2024 from <https://techxplore.com/news/2024-08-texts-automatic-bug-assignment.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.