# Partial Reproduction of Bug Localization Results from BugLocator, BLUiR and AmaLgam+

Mohammad Masudur Rahman
University of Saskatchewan, Canada
masud.rahman@usask.ca

Chanchal K. Roy
University of Saskatchewan, Canada
chanchal.roy@usask.ca

## ABSTRACT

In this one-page abstract, we discuss about three existing studies that were partially reproduced by us as a part of the experiments conducted for our ESEC/FSE 2018 paper – *"Improving IR-Based Bug Localization with Context-Aware Query Reformulation"*. We discuss our methodology of reproduction, the obtained results, and then also present several non-trivial observations as follows.

**ACM Reference Format:**
Mohammad Masudur Rahman and Chanchal K. Roy. 2018. Partial Reproduction of Bug Localization Results from BugLocator, BLUiR and AmaLgam+. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 1 page. https://doi.org/10.1145/nnnnnnn.nnnnnnn

## PAPERS & AUTHORS

**Reproduced Studies:** Results are partially reproduced from three existing studies on bug localization – BugLocator [4], BLUiR [2] and AmaLgam+ [3] as follows:

(1) J. Zhou, H. Zhang and D. Lo. Where should the bugs be fixed? More accurate information retrieval-based bug localization based on bug reports.
(2) R. K. Saha, M. Lease, S. Khurshid and D. E. Perry. Improving bug localization using structured information retrieval.
(3) S. Wang and D. Lo. AmaLgam+: Composing Rich Information Sources for Accurate Bug Localization.

**Reproducer Study:** We reproduce the above three studies in our recent ESEC/FSE 2018 paper [1] as follows:

- M. M. Rahman and C. K. Roy. Improving IR-Based Bug Localization with Context-Aware Query Reformulation.

## ITEMS REPRODUCED

We partially reproduced three performance measures, *Hit@K*, *Mean Average Precision@K (MAP@K)* and *Mean Reciprocal Rank (MRR)* from each of the three studies on IR-based bug localization.

## INTERESTING ASPECTS OF RESULTS

(1) We choose a different set of subject systems than those the existing studies do. However, our reproduced results, on average, are *surprisingly close* to their reported results. This demonstrates the *external validity* of their findings.
(2) When bug reports were categorized according to their quality aspect or noise level, then we noticed *significant difference* between the published results and the reproduced results especially for low quality (noisy) bug reports. Such caveat was not discovered by any of the earlier investigations.

## STEPS OF REPRODUCTION

(1) We collected the authors' prototype of both BugLocator and BLUiR. Unfortunately, they failed to run for the systems containing non-compilable source code. Hence, we spent 8-10 hours in the manual filtration to make them run successfully.
(2) We implemented AmaLgam+ by consulting with the original authors since its prototype was not available. However, we did not use their weight optimization since it did not help.
(3) We use the same set of metrics as done by the papers.

## LESSONS LEARNED

(1) **Working Prototype Really Matters** since it saves the time and efforts of a researcher. However, it might not guarantee a smooth reproduction if the prototype is not well designed.
(2) **Prototype+Source Code ≫ Prototype**. However, the code should be sanitized and well-documented.
(3) **Potential or Weakness through Reproduction:** The true potential or severe weaknesses of an idea or an approach can be only understood through the rigorous reproduction.
(4) **Reproducing Multiple Studies:** Reproduction of only the state-of-the-art might not be sufficient enough especially when experiments are done with new subject systems.
(5) **Infeasibility for Practical Use:** Experimentally superior approaches do not guarantee a smooth practical use. For example, AmaLgam+ was the state-of-the-art. However, it might suffer from severe usability issues due to its high dependencies on the multiple information sources.
(6) **Replication package is hardly available** even for the very high quality publications at the topmost venues.
(7) **Reproduction** makes you (a) appreciate the merit of other researchers' works and (b) a better researcher in the process.

## REFERENCES

[1] M. M. Rahman and C. K. Roy. 2018. Improving IR-Based Bug Localization with Context-Aware Query Reformulation. In *Proc. ESEC/FSE*. 12 (to appear).
[2] R. K. Saha, M. Lease, S. Khurshid, and D. E. Perry. 2013. Improving bug localization using structured information retrieval. In *Proc. ASE*. 345–355.
[3] S. Wang and D. Lo. 2016. AmaLgam+: Composing Rich Information Sources for Accurate Bug Localization. *JSEP* 28, 10 (2016), 921–942.
[4] J. Zhou, H. Zhang, and D. Lo. 2012. Where should the bugs be fixed? More accurate information retrieval-based bug localization based on bug reports. In *Proc. ICSE*. 14–24.


*Conference'17, July 2017, Washington, DC, USA*

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn