Dear Prof. Dwyer, Associate Editor, and Reviewers,

RE: Submission to IEEE Transactions on Software Engineering

We hereby wish to submit our journal paper manuscript—entitled "*Network-Clustered Multi-Modal Bug Localization*"—to the IEEE Transactions on Software Engineering (IEEE-TSE). In this manuscript, we present a novel extension of our previous work on the adaptive multi-modal bug localization (AML), which was published as a conference paper at ESEC/FSE 2015 entitled "*Information Retrieval and Spectrum Based Bug Localization: Better Together*" (Digital Object Identifier: https://doi.org/10.1145/2786805.2786880).

The work described in this manuscript—hereafter referred to as NetML—presents several major extensions to our previous AML technique, as summarized below:

- NetML features a novel integrator component that provides two sets of latent parameters characterizing each bug report and method respectively, whereas AML only presents one set of latent parameters for each bug report. The incorporation of the (additional) latent method parameters provides NetML with a higher degree of freedom to model the rich variations of different bug reports and methods more accurately. Details of the extended integrator model can be found in Section 3.1.

- To our best knowledge, NetML is the first bug localization approach with a loss function formulation that aims for a joint optimization of bug localization error and clustering of the latent parameters of both bug reports and methods. In particular, the clustering is achieved via a network Lasso regularization, which enforces similar bug reports or methods to have latent parameters that are close together. By contrast, the latent parameters of each bug report in AML are learned independently of those of other bug reports, without taking advantage of their similarity traits. Details of the loss function formulation are given in Section 3.2.

- NetML extends the adaptive learning procedure of AML by performing a Newton update to simultaneously update the latent parameters of bug reports and methods on a per-feature basis. The procedure is based on the formulation of strict convex loss function, which provides a theoretical guarantee that any minimum found will be globally optimal. The NetML adaptive learning procedure is elaborated in Section 3.3.

In Section 4, we have also included an extensive suite of experiments using 157 real bugs from four software projects to evaluate the efficacy of NetML. Specifically, we have conducted a benchmark study to assess the performance of NetML in comparison to AML as well as several other state-of-the-art bug localization methods. The results show that NetML surpasses the best performing baseline (i.e., AML) by 48.39%, 15.49%, 8.7%, and 13.92%, in terms of the number of bugs successfully localized among the top 1, 5, and 10 methods, and Mean Average Precision (MAP), respectively. Our statistical significance test also shows that NetML significantly improves

AML, in terms of Top 1, Top 5 and MAP scores.

Additionally, we have revised all other sections to improve the exposition of the paper.

Given the above points, we believe that all the major extensions constitute more than 30% improvement, which makes our paper qualified for potential publication in IEEE-TSE.

Thank you for your consideration and attention. We look forward to receiving your review.

Yours Sincerely,

Thong Hoang, Richard J. Oentaryo, Tien-Duy B. Le, and David Lo
School of Information Systems
Singapore Management University