Developers often spend much effort and resources to debug a program. To help the developers debug, numerous information retrieval (IR)-based and spectrum-based bug localization techniques have been devised. IR-based techniques process textual information in bug reports, while spectrum-based techniques process program spectra (i.e., a record of which program elements are executed for each test case). While both techniques ultimately generate a ranked list of program elements that likely contain a bug, they only consider one source of information—either bug reports or program spectra—which is not optimal. In light of this deficiency, this paper presents a new approach dubbed Network-clustered Multi-modal Bug Localization (NetML), which utilizes multi-modal information from both bug reports and program spectra to localize bugs. NetML facilitates an effective bug localization by carrying out a joint optimization of bug localization error and clustering of both bug reports and program elements (i.e., methods). The clustering is achieved through the incorporation of network Lasso regularization, which incentivizes the latent parameters of similar bug reports and similar program elements to be close together. To estimate the latent parameters of both bug reports and methods, NetML features an adaptive learning procedure based on Newton method that updates the parameters on a per-feature basis. Extensive experiments on 157 real bugs from four software systems have been conducted to evaluate NetML against various state-of-the-art localization methods. The results show that NetML surpasses the best-performing baseline by 48.39%, 15.49%, 8.7%, and 13.92%, in terms of the number of bugs successfully localized when a developer inspects the top 1, 5, and 10 methods and Mean Average Precision (MAP), respectively.

Developers often spend much effort and resources to debug a program. To help the developers debug, numerous information retrieval (IR)-based and spectrum-based bug localization techniques have been devised. IR-based techniques process textual information in bug reports, while spectrum-based techniques process program spectra. While both techniques ultimately generate a ranked list of program elements that likely contain a bug, they only consider one source of information—either bug reports or program spectra—which is not optimal. In light of this deficiency, this paper presents a new approach dubbed Network-clustered Multi-modal Bug Localization (NetML), which utilizes multi-modal information from both bug reports and program spectra to localize bugs. NetML facilitates an effective bug localization by carrying out a joint optimization of bug localization error and clustering of both bug reports and program elements. Extensive experiments on 157 real bugs from four software systems have been conducted to evaluate NetML against various state-of-the-art localization methods. The results show that NetML surpasses the best-performing baseline by 48.39%, 15.49%, 8.7%, and 13.92%, in terms of the number of bugs successfully localized when a developer inspects the top 1, 5, and 10 methods and Mean Average Precision (MAP), respectively.