

# Can We Predict the Quality of Spectrum-based Fault Localization?

Submitted by aekwall on Mon, 03/01/2021 - 4:47pm

Title Can We Predict the Quality of Spectrum-based Fault Localization?  
Publication Type Conference Paper  
Year of Publication 2020  
Authors [Golagha, M.](#), [Pretschner, A.](#), [Briand, L. C.](#)  
Conference Name 2020 IEEE 13th International Conference on Software Testing, Validation and Verification (ICST)

Keywords [automated fault localization](#), [Complexity theory](#), [Couplings](#), [Debugging](#), [dynamic test metrics](#), [dynamic test suite](#), [effective fault localization](#), [Electronic mail](#), [fault diagnosis](#), [fault localization effectiveness](#), [fault localization techniques](#), [fault-related metrics](#), [machine learning](#), [Measurement](#), [Metrics](#), [potential effectiveness](#), [prediction model](#), [Predictive models](#), [program debugging](#), [program testing](#), [pubcrawl](#), [Resiliency](#), [Scalability](#), [software fault tolerance](#), [software metrics](#), [spectrum-based fault localization](#), [static test metrics](#), [static test suite](#), [Tools](#), [work factor metrics](#)

Abstract Fault localization and repair are time-consuming and tedious. There is a significant and growing need for automated techniques to support such tasks. Despite significant progress in this area, existing fault localization techniques are not widely applied in practice yet and their effectiveness varies greatly from case to case. Existing work suggests new algorithms and ideas as well as adjustments to the test suites to improve the effectiveness of automated fault localization. However, important questions remain open: Why is the effectiveness of these techniques so unpredictable? What are the factors that influence the effectiveness of fault localization? Can we accurately predict fault localization effectiveness? In this paper, we try to answer these questions by collecting 70 static, dynamic, test suite, and fault-related metrics that we hypothesize are related to effectiveness. Our analysis shows that a combination of only a few static, dynamic, and test metrics enables the construction of a prediction model with excellent discrimination power between levels of effectiveness (eight metrics yielding an AUC of .86; fifteen metrics yielding an AUC of .88). The model hence yields a practically useful confidence factor that can be used to assess the potential effectiveness of fault localization. Given that the metrics are the most influential metrics explaining the effectiveness of fault localization, they can also be used as a guide for corrective actions on code and test suites leading to more effective fault localization.

DOI [10.1109/ICST46399.2020.00012](https://doi.org/10.1109/ICST46399.2020.00012)

## Citation Key golagha\_can\_2020



[Scalability](#) [Electronic mail](#) [Resiliency](#) [pubcrawl](#) [Metrics](#) [tools](#) [program testing](#) [machine learning](#) [Measurement](#) [Complexity theory](#) [Predictive models](#) [program debugging](#) [software metrics](#) [software fault tolerance](#) [Couplings](#) [fault diagnosis](#) [automated fault localization](#) [debugging](#) [dynamic test metrics](#) [dynamic test suite](#) [effective fault localization](#) [fault localization effectiveness](#) [fault localization techniques](#) [fault-related metrics](#) [potential effectiveness](#) [prediction model](#) [spectrum-based fault localization](#) [static test metrics](#) [static test suite](#) [work factor metrics](#)

---