

Highlighting code areas, which are most influenced with bugs.

Yelshat Duskaliyev
 Innopolis, Tatarstan
 e.duskaliyev@innopolis.ru

Abstract—In every company, there exists a practice of organizing the work with the code in a way, when one programmer clones the latest version of the repository, bumps a new version of the code, which is called a branch, makes changes in it according to the specification of the requested feature and then raises a pull request. Unfortunately, not all such code modifications are safe to use, so there is a practice of *code reviews* in most companies. However, as this procedure requires the reviewer’s attention to all parts of the code editions, the reviewer sometimes may not pay much attention to the areas that are often tainted with the bug. This paper suggests a method of highlighting the code areas, which often misbehaves.

Index Terms—git, pull request, machine learning, diff

I. INTRODUCTION

When the programmer finishes writing his code, he creates a pull request to suggest merging its changes into the main repository branch. In order to assure the quality of the code, the reviewer checks the code by watching it and looking for the programmer’s mistakes. When the reviewer finds a mistake, he writes a commentary to the specific code line. He wants to highlight to pay the programmer’s attention to the line and change it according to the reviewer’s commentary. It often happens that the programmers eventually make several mistakes in the same area for some reason.

This information might be used to apply the machine learning algorithm. We may detect the area, which is often highlighted during the code reviews, and then changed for proper behavior and in future pull requests check that this area was or was not touched with changes, and, if it was touched, we might notify the reviewer to pay attention to such code area and apply some standard optimizations for it, in order to improve the quality of the code.