Prioritizing Dissimilar Test Cases in Regression Testing using Historical Failure Data

Test case prioritization assigns new order of test cases for detecting regression faults at early. In regression testing when new version is released, all the test cases of both previous and current versions are executed to ensure the desired functionality. This process increases the volume of test cases in regression testing, which is expensive and time consuming. That is why the test cases are needed to be reordered for exploring maximum faults in minimum test cases execution. Usually test case prioritization techniques are designed based on source code coverage, requirements clustering, etc. Most of these techniques contain the similarity relationship among the test cases. However, similarity based technique may stuck in local minima. To overcome the limitation of similarity based prioritization, this paper proposed the dissimilar clustering based approach using historical data analysis to detect maximum faults. In this approach, dissimilar test cases placed in the top of the test suites and executed earlier than similar test cases. Proposed scheme is evaluated using well established Defects4j dataset, and it has reported that proposed strategy performs 54.95%, 41.83% and 7.00% better than untreated (normal ordering), random and similarity cluster based prioritization methods.
respectively.

References


Index Terms

Computer Science

Information Sciences

Keywords

Prioritization, Dissimilarity, Clustering, Historical Data