



A Survey on Test Case Selection and Prioritization

Presitha Aarthi. M, Nandini. V

Sona College of technology,
India

Abstract— *Software testing is traditionally considered a process of executing test cases, which are carefully designed using test case design techniques. Test case design techniques aim at ensuring systematic Coverage, detection of typical error types and reduction of redundant Testing. Every time it is not possible to perform each and every test case. Hence it is important to select the test case and prioritize it. The test case selection aims at selecting the test case from large test suite. The major goal of Test case prioritization is to prioritize the test case sequence and find faults as early as possible to improve the efficiency. Larger system consist of huge number of test cases, in such a case prioritizing the test case is very efficient. Hence the test case selection and prioritization plays a vital role in software testing. In this paper many test case selection and prioritization papers were analysed and methods used in those papers were discussed.*

Keywords— *Fault-Exposing-Potential (FEP), Genetic Algorithm (GA), Average Percentage rate of Faults Detection (APFD), Regression Test Selection (RTS), Regression Test Prioritization (RTP)*

I. INTRODUCTION

In software production the software maintenance activity is an expansive phase which cost nearly 60% of total cost. Regression testing is an important phase in software maintenance activity to ensure the modification caused by debugging. It provides confidence that changes do not harm the existing behaviour of the software. Test case minimization, selection and prioritization are the approaches in regression testing. The test case selection identifies the test cases that are relevant to some set of recent changes. Test case prioritization orders test cases in such a way that early fault detection is maximized. Regression test selection technique will help in selecting test cases from the test suite. Even a properly developed and optimized regression test suite can grow extremely large if the system under test is also large and complex. The consequence is that full regression testing can become prohibitively expensive. Hence various selection and prioritization methods were analysed to find which method is efficient based on both cost and performance.

II. LITERATURE REVIEW

In last few years there were many publications discussed the concept of test case selection and prioritization. In this section various prioritization methods and related works are discussed.

(A) An Efficient Algorithm for Reducing the Test Cases:

This paper deals with testing. Testing includes planning, creation and execution of test cases. Regression is the process of repeating the testing with same test cases or additional test cases. The priority can be given to the test cases depending on the code coverage. An algorithm is used to improve the testing process by finding the failure in early stages.

Regression test case selection algorithm is used to select the test cases from the test suite. When the original code is modified the code coverage executes the test cases for the modified lines of code. Test case prioritization algorithm is used for prioritization and priority is given based on the comparison between the test cases.

The test case prioritization algorithm works for both requirement and testing phase. Each test case is given a value and based on its value the priority can be set. This algorithm is effective for the early fault detection. This algorithm is suitable for smaller applications. Large applications consist of huge number of test cases and it is difficult to assign value for all the test cases. Hence it is necessary to validate this algorithm with large set of test cases.

(B) Test Case Selection for Black-box Regression Testing of Database Application:

This paper focuses on Black-box approach such as specification based approach. This approach aims at selecting regression test case for large scale applications. The Classification tree model is used to select the test cases. The large scale application consists of plenty of test cases, in order to select the test cases the similarity measurement is combined with partition based test case selection.

Regression testing is very expensive activity so the test cases must be carefully selected and prioritized. The classification tree model is used to partition the input domain of the system being tested. CTE-XL tool is used to partition the test domain and generate a test specification. Random partition-based test case selection is used to partition the test cases and then select a random test case from that partition. The most similar and most diverse test cases are identified and encoded as Boolean matrix. The similar functions such Euclidian, Manhattan, Mahalanibis and NCD are used to generate similarity matrix for entire test suite.

Combining Mahalanibis similarity function and 1+1 EA algorithm proved to be most efficient. In classification tree model the diversity is often low within partitions thus the potential benefit is low.

(C) Effective test Selection for the End User:

This paper focuses on test case prioritization by analysing various prioritization techniques. The techniques analysed are No prioritization, Random prioritization, Total branch coverage prioritization, Additional branch coverage prioritization, Total fault-exposing –potential (FEP) prioritization, Additional fault-exposing-potential (FEP) prioritization

The result shows that the prioritization techniques can increase the fault detection rate of test cases. MOST-RELEVANT coverage, CANONICAL prioritization and RANDOM selection methods are considered as baseline methods and proposed CONFIDENCE Prioritization, COS-DIST prioritization and LEAST-RELEVANT prioritization especially to find surprise failure.

The three proposed methods use different techniques to order the test case and prioritize it. These method increases the test case selection efficiency. The false positive and false negative test cases are given equal cost.

(D) Search Based Constrained Test Case Selection:

In this paper Luciano S. deSouza et al. worked with test case selection. Binary Constraints Particle Swarm Optimization (BCPSO) algorithm is used for test case selection. Testing takes 40% of final project development cost .Literature survey techniques are black box, white box, test case generation and execution. Task of reducing test suites based on given selection criteria called TCS.

The PSO algorithm defines the random population of particle, each one having the position in the search space. Particles are evaluated by the fitness function. The algorithm stops when the particle reaches the better position. Hybrid implementation of BCPSO was developed by combining it with forward selection and hill-climbing algorithms.

Hybrid search technique were developed and validated. The quality of the test case selection process depends upon the effectiveness of the used search technique, constraints imposed by the user and the feature of the test suite at hand. The quality of the process can be increased. Experiment was performed on limited test suites only. The test cases are selected only based on the requirements.

(E) A Uniform Representation of Hybrid Criteria for Regression Testing:

In this paper Sreedeevisampath et al. worked with Regression testing. The hybrid combinations are prioritized using rank, merge and choice formulations. In regression test selection a subset of test cases is selected from large test suite. In this method Multiple criteria is used which is more beneficial than individual criteria. In functional paradigm the individual criteria is used as an operands and their combinations are used as an operators.

Genetic algorithm is used for prioritization. Individual criteria are combined to form hybrid criteria using the functional paradigm. A random choice of test case is made when multiple test cases cover the same number of statements. In this method three hybrids are used, Rank gives the precedence to different criteria, Merge executes multiple criteria simultaneously and Choice selects the best from multiple criteria. The relationship of multiple criteria to different techniques such as test case prioritization and selection to be examined.

(F) Code Coverage-Based Regression Test Selection and Prioritization in Web Kit:

This paper focuses on regression testing. Selective retesting can be applied to reduce the testing cost while maintaining similar defect detection capability. To enhance the internal quality of the system including the reliability and efficiency of the regression testing can be conducted. The quality of the regression test suite is indicated by defect detection capability. The regression testing is prohibitively expensive because a large number of test cases have to be rerun many times. The fundamental regression technique used is regression test selection and prioritization based on code coverage.

Various properties are tested in this technique. Whether the reduced test set identifies the same failure and how much reduction can be achieved on an average. The procedure level code coverage of changes are applied first. The data is taken from the WebKit version control system, test execution and defect data base covering fixed interval of system's evolution. The initial coverage information for each procedure is determined and the set of changed procedure for a specific version is identified. The database is looked up to check if there is any set of test cases for the changed procedure. The prioritization is set for too large selection and the selected tests are tested with regression testing.

The change based selection method has been successfully implemented in live environment of the project. The selective test provides significant improvement in total testing time. Test case prioritization strategy is useful to reduce the size of selection set. As the priority is set for each test case only necessary test cases can be run after modification. The change based test selection method is beneficial only when the small portion of test is relevant for the change. The overall quality of the system is investigated only based on the code coverage.

(G) Size Constrained Regression Test Case Selection Using Multicriteria Optimization:

In this paper SiavashMirarab et al. Regression test case selection using multicriteria optimization. This is a novel approach for selecting and ordering a predetermined number of testcase from an existing test suite. Constraints relaxation is used to find many close-to-optimal solution points and the final solution is obtained using voting mechanism.

The regression testing is most cost effective because it involves huge number of test cases. Regression Test Selection (RTS) and Regression Test Prioritization (RTP) are the approaches emerged from regression testing. RTS technique select sub set of test cases for re-execution, reducing the overall cost. RTP reorder test cases based on their importance, enabling the faster detection of faults. Size constrained Regression Test Selection problem was introduced, which is to select a subset from the given number of test cases. The max-min criterion has been used in this approach. An approximate and suboptimal mathematical solution to the Integer linear programming was proposed and the greedy algorithm is used for prioritizing the selected subset of test cases.

The fault detection capability is a positive scalar quantity denoted by D . D_{\min} function is used as a promising optimization criterion. The D -function maximizes the minimum coverage across all software elements. The MIN-Greedy algorithm takes two set of inputs: Coverage data and Priority measure for software element. The regression test selection takes number of test cases to be selected as input. Multicriteria approach helps to use those limited resource more effectively. The proposed algorithm does not find the exact Pareto front. The weights are assigned to each objective but these weights do not take all possible values. The set of solution found is not complete.

(H) Code Coverage Based Test Case Selection and Prioritization:

In this paper R.Beena et al. discussed about the code coverage based test case selection. Regression testing is executed to guarantee the desirable functionality of existing software. The maintenance of test suite is enormous which needs a high investment of time and money. Test case minimization, test case selection and test case prioritization are regression testing techniques. Test case minimization aims to eliminate the redundant test cases, Test case selection technique are performed to reduce the size of test suite and Test case prioritization techniques are concerned with ordering of test cases.

The test cases are grouped into three clusters namely out-dated, required, and surplus. The out-dated cluster contains test case required by both original and modified program. The required test case group consists of test cases required to execute the modified version of software. The surplus group consists of test cases required for latest version but not for the modified version. The test cases are represented as binary matrix. Test case prioritization consists of much iteration and those test cases are reduced by iterations. The size of the test suite is reduced and hence the cost and time for execution of test cases can be minimized. A new technique for test case selection and test case prioritization for regression technique has been proposed. The proposed technique is very effective in terms of cost and time.

(I) A Regression Test Case Selection and Prioritization:

This paper focuses on regression testing. An evolutionary regression test case prioritization based on dependence graph model using genetic algorithm (GA). The approach is based on optimization of selected test case from test suite. The test cases are selected based on the affected statement and ordered them based on their fitness by using GA. Average Percentage of rate of Faults Detection (APFD) metric is used to verify the goodness of ordering.

The prioritization of test case by identifying and selecting of best test case from the selected test cases are the challenge in regression testing. Regression testing approach is based on source code. Evolutionary prioritization approach selects best test case from the existing test suite using Dependence Graph. Dependence graph acts as intermediate to identify the changes and the selected test cases will be prioritized by using genetic algorithm. Genetic algorithm will reduce the cost of regression testing by increasing the rate of fault detection. Many activities are involved in proposed test case selection and prioritization: Identify changes, adding objects, deleting objects, Test coverage generation, ESDG model constructor and ESDG update, affected statement identification, test case selection and test case prioritization. The effectiveness and efficiency was evaluated using APFD metric. The approach provides the better results for rate of fault detection. Every time the changes are introduced into the program the ESDG have to be updated in timely manner.

(J) A New Code Based Test Case Prioritization Technique:

In this paper Muhammad Shahid et al. discussed about prioritization technique. Test cases are important part of software testing activity. The prioritization of test suites is very beneficial. The test cases that cover maximum part of the code are considered more valuable. Regression test selection and prioritization technique attempts to save time and decrease the overall cost. Regression testing can be involved in different levels such as unit, integration or system level testing.

New regression test cases priority technique that arranges the selected test cases based on their code coverage percentage. The algorithm is based on the information of total methods covered by test case. Algorithm determines the percentage of methods covered by each case and then uses it to sort the test cases in descending order. Prioritization algorithm reduces the high cost and time. This algorithm is applied for only smaller system.

III. CONCLUSIONS

For regression testing there are many algorithms which present a new approach for test case selection and prioritization. The studies show that regression testing is expensive approach. It consists of large number of test cases. If the original program is modified all the test cases cannot be rerun, so the selection of test cases is necessary. Those selected test cases are prioritized based on certain criteria. This approach is mainly applicable for large system. Each paper uses certain methods to select and prioritize the test cases. In future the analysis will be conducted on metrics to overcome limitations of above mentioned methods.

REFERENCES

- [1] SiavashMirarab, SoroushAkhlaghi, and LadanTahvildari, "Size-Constrained Regression Test Case Selection Using Multicriteria Optimization," IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, VOL. 38, NO. 4, JULY/AUGUST 2012.
- [2] A.Pravin and Dr. S. Srinivasan,"An Efficient Algorithm for Reducing the Test Cases which is Used for Performing Regression Testing," 2nd International Conference on Computational Techniques and Artificial Intelligence (ICCTAI'2013) March 17-18, 2013
- [3] Erik Rogstad , Lionel Briand, Richard Torkar , " Test case selection for black-box regression testing of database applications" ELSEVIER, Information and Software Technology 55 (2013) 1781–1795.
- [4] Arpad Beszedes, TamasGergely, LajosSchrettner, JuditJasz, Laszlo Lango, TiborGyimothy, " Code Coverage-Based Regression Test Selection and Prioritization in WebKit," 2012 28th IEEE International Conference on Software Maintenance (ICSM)
- [5] R.Beena, Dr.S.Sarala, "CODE COVERAGE BASED TEST CASE SELECTION
- [6] AND PRIORITIZATION," International Journal of Software Engineering & Applications (IJSEA), Vol.4, No.6, November 2013.
- [7] Alex Groce, Todd Kulesza, Chaoqiang Zhang, ShaliniShamasunder, Margaret Burnett, Weng-Keen Wong,SimoneStumpf, Shubhomoy Das, Amber Shinsel, Forrest Bice, and Kevin McIntosh," You Are the Only Possible Oracle: Effective Test Selection for End Users of Interactive Machine Learning Systems," IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, VOL. 40, NO. 3, MARCH 2014
- [8] Luciano S. deSouza , Ricardo B.C.Prudêncio, FlaviadeA.Barros , Eduardo H.daS.Aranha," Search based constrained test case selection using execution effort," ELSEVIER, Expert Systems with Applications 40 (2013) 4887–4896
- [9] Samaila Musa, Abu BakarMd Sultan, Abdul Azim Bin AbdGhani, SalmiBaharom,"A Regression Test Case Selection and Prioritization for Object-Oriented Programs using Dependency Graph and Genetic Algorithm" *Research Inveny: International Journal of Engineering And Science Vol.4, Issue 7 (July 2014), PP 54-64 Issn (e): 2278-4721, Issn (p):2319-6483.*
- [10] Muhammad Shahid and Suhaimi Ibrahim," A New Code Based Test Case Prioritization Technique", International Journal of Software Engineering and Its Applications Vol.8, No.6 (2014), pp.31-38.
- [11] IzzatAlsmadi,SaschaAlda," Test Cases Reductionand Selection Optimization in Testing Web Services"*I.J. Information Engineering and Electronic Business*, 2012, 5, 1-8
- [12] Navdeep Koundal1, Ankur Shukla2, Mohsin Rashid Mir,"Test Case Selection Using Bee Colony Optimization", International Journal of Science and Research (IJSR),Impact Factor (2012): 3.358.
- [13] Sujata1 , G.N.Purohit, " TOOL SUPPORT FOR TEST CASE SELECTION IN REGRESSION TESTING" International Journal of Software Engineering & Applications (IJSEA), Vol.2, No.4, October 2011.