Volume 5, Issue 2, February 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

A Systematic Review of Regression Test Selection Techniques: An Empirical Study

Ruchika Sharma
Assistant Professor (CSE),
BUEST, India

Abstract - Regression testing is an important and expensive activity that is undertaken every time a program is modified to ensure that the modifications do not introduce new bugs into previously validated code. An important research problem. Since regression testing is an expensive process, researchers have proposedregression test selection techniques as a way to reduce some of this expense. These techniquesattempt to reduce costs by selecting and running only a subset of the test cases in a program'sexisting test suite. Although there have been some empirical evaluations ofindividual techniques, to our knowledge only one comparative study, focusing on one aspect oftwo of these techniques, has been reported in the literature.

Key Words - Regression testing, selective retest, empirical study

I. INTRODUCTION

As developers maintain a software system, they periodically *regression test*it, hoping to find errors caused by their changes and provide confidencethat their modifications are correct. To support this process, developersoften create an initial test suite, and then reuse it for regression testing.The simplest regression testing strategy, *retest all*, reruns every test casein the initial test suite. This approach, however, can be prohibitivelyexpensive—rerunning all test cases in the test suite may require anunacceptable amount of time. An alternative approach, *regression testselection*, reruns only a subset of the initial test suite. Of course, thisapproach is imperfect as well—regression test selection techniques canhave substantial costs, and can discard test cases that could reveal faults,possibly reducing fault detection effectiveness.This trade-off between the time required to select and run test cases andthe fault detection ability of the test cases that are run is central toregression test selection. Because there are many ways in which to approachthis trade-off, a variety of test selection techniques have beenproposed (e.g., Agrawal et al. [1993], Chen et al. [1994], Harrold and Soffa[1988], Hartmann and Robson [1990], Leung and White [1990], Ostrandand Weyuker [1988], and Rothermel and Harrold [1997]). Although therehave been some analytical and empirical evaluations of individual techniques[Chen et al. 1994; Rosenblum and Weyuker 1997b; Rothermel andHarrold 1997; 1998], to our knowledge only one comparative study, focusingon one aspect of two of these techniques, has been reported in theliterature [Rosenblum and Rothermel 1997].

II. REGRESSION TESTING SUMMARY AND LITERATURE REVIEW

A. Regression Testing

Let *P* be a procedure or program; let *P*' be a modified version of *P*; and let*T* be a test suite for *P*. A typical regression test proceeds as follows:

(1) Select *T*' # *T*, a set of test cases to execute on *P*'.
(2) Test *P*' with *T*', establishing *P*'s correctness with respect to *T*'.
(3) If necessary, create *T*'', a set of new functional or structural test casesfor *P*'.
(4) Test *P*' with *T*'', establishing *P*''s correctness with respect to *T*''.
(5) Create *T*''', a new test suite and test execution profile for *P*', from *T*,*T*', and *T*''.

Each of these steps involves important problems. Step 1 involves the*regression test selection problem*: the problem of selecting a subset *T*' of *T*with which to test *P*'. Step 3 addresses the *coverage identification problem*:the problem of identifying portions of *P*' or its specification that requireadditional testing. Steps 2 and 4 address the *test suite execution problem*:the problem of efficiently executing test suites and checking test results forcorrectness. Step 5 addresses the *test suite maintenance problem*: theproblem of updating and storing test information.

B. Regression Test Selection Techniques

A variety of regression test selection techniques have been described in theresearch literature. A survey by Rothermel and Harrold [1996] describesseveral families of techniques; we consider three such families, along withtwo additional approaches often used in practice. Here we describe thesefamilies and approaches, and provide a representative example of each;Rothermel and Harrold [1996] and the references for the cited techniquesthemselves provide additional details.

C. Minimization Techniques

Minimization-based regression test selectiontechniques (e.g., Fischer et al. [1981] and Hartmann and Robson[1990]), hereafter referred to as *minimization techniques*, attempt to selectminimal sets of test cases from *T* that yield coverage of modified or affectedportions of *P*.

- **Dataflow Techniques.** Dataflow-coverage-based regression test selectiontechniques (e.g., Harrold and Soffa [1988], Ostrand and Weyuker[1988], and Taha et al. [1989]), hereafter referred to as *dataflow techniques*,select test cases that exercise data interactions that have been affected bymodifications.
- **Safe Techniques.** Most regression test selection techniques—minimizationand dataflow techniques among them—are not designed to be*safe*. Techniques that are not safe can fail to select a test case that wouldhave revealed a fault in the modified program. In contrast, when an explicitset of safety conditions can be satisfied, safe regression test selectiontechniques guarantee that the selected subset, $T^{'}$, contains all test cases inthe original test suite *T* that can reveal faults in *P*'
- **Ad Hoc/Random Techniques.** When time constraints prohibit theuse of a retest-all approach, but no test selection tool is available, developersoften select test cases based on "hunches," or loose associations of testcases with functionality. Another simple approach is to randomly select apredetermined number of test cases from *T*.
- **Retest-All Technique.** The retest-all technique simply reuses allexisting test cases. To test *P*', the technique effectively "selects" all testcases in *T*.

III. RESEARCH METHOD

A. Research Questions

This review aims at summarizing the current state of the art inregression test selection research by proposing answers to thefollowing questions:

1) Which techniques for regression test selection in theliterature have been evaluated empirically?
2) Can these techniques be classified, and if so, how?
3) Are there significant differences between thesetechniques that can be established using empiricalevidence?
4) Can technique *A* be shown to be superior to technique*B*, based on empirical evidence?

B.Sources of information

In order to gain a broad perspective,we searched widely in electronicsources. The following databases were covered:

- Inspec(www.theiet.org/publishing/inspec)
- Compendex(www.engineeringvillage2.org)
- ACM Digital Library (portal.acm.org)
- IEEE eXplore (ieeexplore.ieee.org)
- ScienceDirect (www.sciencedirect.com)
- SpringerLNCS (www.springer.com/lncs)
- Webof Science(www.isiknowledge.com)

These databases cover the most relevant journals and conferenceand workshop proceedings within software engineering, asconfirmed by Dyba et al. [8].

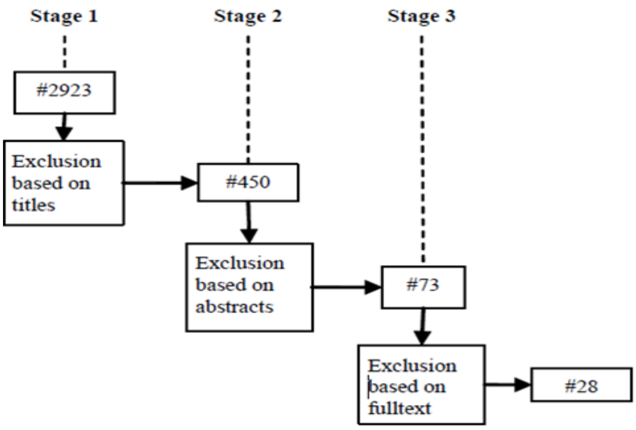

Figure 1: Study selection procedure

C. Search criteria

The initial search criteria were broad in order to include articleswith different uses of terminology. The key words used were<regression> and (<test> or <testing>) and <software>. The startyear was set to 1969 to ensure that all relevant research within thefield would be included, and the last date for inclusion ispublications within 2006. Kitchenham recommends that exclusionbased on languages should be avoided [27]. At least one paper inthe original selection was in Japanese, and since we were not ableto read Japanese, we could not follow this recommendation. Thus,only papers written in English are included. The initial search located 2923 potentially relevant papers.

D. Study Selection

In order to obtain independent assessments, four researchers wereinvolved in a three stage selection process, see Figure 1. In thefirst stage duplicates and irrelevant papers were excludedmanually based on titles. In our case, the share of irrelevantpapers was extremely large since papers on software for *statistical*regression testing could not be distinguished from papers on*software* regression testing in the database search. After the firststage 450 papers remained. In the second stage information inabstracts was analyzed and the papers were classified along twodimensions: research approach and regression testing approach.Papers not presenting an empirical research approach wereexcluded as well as papers not focusing on regression testselection, e.g. papers on test suite maintenance or test automation.In the third stage a full text analysis was performed and theempirical quality of the studies was further assessed. Thefollowing questions were asked in order to form an opinion aboutwhich studies to exclude or include for final data extraction:

- Is a specific regression test selection method evaluated?
- Are the metrics and the results relevant for acomparison of methods?
- Is data collected and analyzed in a sufficiently rigorousmanner?

These questions are derived from a list of questions, used for asimilar purpose, published by Dybå et al. [8]. However in ourreview context, quality requirements for inclusion had to beweaker than suggested by Dybå et al. in order to obtain a usefulset of studies to compare.

E. Threats to validity

Threats to the validity of the systematic review are analyzed according to the following taxonomy; construct validity,reliability, internal validity and external validity.Construct validity reflects to what extent the phenomenon understudy really represents what the researchers have in mind andwhat is investigated according to the research questions. The main threat here is related to terminology. Since the systematic reviewis based on a hierarchical structure of terms – regressiontest/testing consists of the activities modification identification,test selection, test execution and test suite maintenance – wemight miss other relevant studies on test selection. However, thisis a consciously decided limitation which has to be taken intoaccount in the use of the results.Reliability focuses on whether the data is collected and theanalysis is conducted in a way that it can be repeated by otherresearchers with the same results. In a systematic review, theinclusion and exclusion of studies is the major focus here,especially in this case where another domain (statistics) also usesthe term regression testing. Our countermeasures taken were to setup criteria and to use two researchers to classify papers in stages 2and 3. In cases of disagreement, a third opinion is used. One ofthe primary researchers was changed between stages 2 and 3.Still, the uncertainties in the classifications are prevalent and amajor threat to reliability, especially since the quality standardsfor empirical studies in software engineering are not high enough.Research databases is another threat to reliability [8]. The threat isreduced by using multiple databases; still the non-determinism ofsome database searches is a major threat to the reliability of anysystematic review.

IV. CONCLUSIONS

In this article we present initial results of an empirical study of regression test selection techniques. This study examined some of the costs and benefits of several regression test selection techniques. Our study provide an infrastructure for further research by ourselves and others. As we discussed earlier, this experiment, like any other, has several limits to its validity. Keeping this in mind, we draw several observations from this work.

REFERENCES

[1] Agrawal, H., Horgan, J.R., Krauser, E.W., “Incremental regression testing”, CSM-93 (Cat. No.93CH3360-5). IEEE Comput. Soc. Press, 348-57, 1993.

[2] Anjaneyulu, P. and Animesh, B, “An approach for test suite selection to validate applications on deployment of COTS upgrades”, IEEE Computer Society, 7 pp. 2005.

[3] Bible, J., Rothermel, G., “A comparative study of coarse- and fine-grained safe regression test-selection techniques”. 2001.

[4] Binkley, D., “The application of program slicing toregression testing. Information and Software Technology”,40(11-12), 583-94.

[5] Brereton, P., Kitchenham, B., “Lessons from applying the systematicliterature review process within the software engineeringdomain”, 80(4), 571-83, 2007.

[6] Chen, Y.-F., Rosenblum, “Test tube:A system for selective regression testing”, 211-220, 1994

[7] Chengying, M. and Yansheng, L., “Regression testing forcomponent-based software systems by enhancing changeinformation,” IEEE Computer Society, 8 pp,2005.

[8] Dybå, T., Dingsöyr, T., and Hanssen, G.K., “ApplyingSystematic Reviews to Diverse Study Types: An ExperienceReport”, 225-234, 2007.

[9] Elbaum, S., Kallakuri, P., Malishevsky, A., “Understanding the effects of changes on thecost-effectiveness of regression testing techniques. SoftwareTesting, Verification and Reliability”, 13(2), 65-83,2003.

[10] Fischer, K., Raji, F., and Chruscicki, A, “A methodologyfor retesting modified software”,IEEE, 6-3,1981

[11] Graves, T.L., Harrold, M.J., Kim, J.M., Porter, “An empirical study of regression testselection techniques. ACM Transactions on SoftwareEngineering and Methodology”, 10(2), 184-208, 2001.

[12] Gregg, R. and Mary Jean, H., “A Safe, EfficientRegression Test Selection Technique”,2000

[13] Gupta, R., Harrold, M.J, "An approachto regression testing using slicing", IEEE Comput. Soc.Press, 299-308, 1992.

[14] Gupta, R., Harrold, M.J, "Programslicing-based regression testing techniques. Software Testing,Verification and Reliability", 6(2), 83-111,1996.

[15] Haftmann, F., Kossmann, D., "A frameworkfor efficient regression tests on database applications", VLDBJournal. 16(1), 145-64, 2007.

[16] Harrold, M.J., Jones, J.A., Tongyu, L., Donglin, L., "Regression test selection for Java software", ACM, 312-26, 2001

[17] Harrold, M.J. and Souffa, M.L. 1988, "An incremental approachto unit testing during maintenance", (IEEE Cat. No.88CH2615-3). IEEE Comput. Soc. Press, 362-7, 1988.

[18] Hartmann, J. and Robson, D.J., "Approaches to regressiontesting. In Proceedings of the Conference on SoftwareMaintenance, IEEEComput. Soc. Press, 368-72, 1988.

[19] Hartmann, J. and Robson, "Techniques for selectiverevalidation. IEEE Software, 7(1), 31-6, 1990.

[20] Hutchins, M., Foster, H., Goradia, T., "Experiments on the effectiveness of dataflow- and controlflow-based test adequacy criteria", IEEE Comput. Soc. Press, 191-200, 1994

[21] Hyunsook, D., Elbaum, S., and Rothermel, "Infrastructure support for controlled experimentation withsoftware testing and regression testing techniques," IEEE Comput. Soc, 60-70, 2004.

[22] Jiang, Z., Robinson, B., Williams, L., "Applying regression test selection for COTS-basedapplications ACM, 512-21, 2006.

[23] Jiang, Z., Robinson, B., Williams, L., "Alightweight process for change identification and regressiontest selection in using COTS components", IEEE Computer Society, 7 pp, 2006.

[24] Juristo, N., Moreno, A.M., "Insearch of what we experimentally know about unit testing", 2006.

[25] KampenesVigdis, B., Dyba, "A systematic review of effect size in softwareengineering experiments. Information and SoftwareTechnology, 49(11-12), 1073-1073,2007.

[26] Kim, J.-M., Porter, A., "An empiricalstudy of regression test application frequency. SoftwareTesting, Verification and Reliability", 15(4), 257-279, 2005.

