TEXT-BASED TEST CASE SELECTION EVALUATED BY CODE COVERAGE

This briefing reports scientific evidence on evaluating a tool for regression test case selection using a non-instrumented code coverage analysis based on empirical research experiments.

FINDINGS

- The findings presented in this briefing show the evaluation of the AutoTestPlan tool compared with a test architect using code coverage as a metric

Selector	TCs	Coverage	Failure
Architect	120	51.74%	0
ATP	62	51.74%	0

Table 1: 1st experiment

- Both executions (architects and ATP) had the same code coverage of 51.74%, but the ATP's test plan is about 50% smaller than that proposed by architects

Selector	TCs	Coverage	Failure
Architect	175	4%	9
ATP	166	4%	13

Table 2: 2nd experiment

- ATP has shrunk the 175 test cases into 166 test cases while preserving its code coverage of 4% and increasing the number of failures found. The main reasons associated to this are: (i) the test cases are somewhat vague, leaving the tester free to choose some variations in the smartphone interaction; (ii) the architect's selection was executed in India, and ATP's selection was executed in Brazil.

Selector	TCs	Coverage	Failure
Architect	116	53.45%	0
ATP	108	52.07%	0

Table 3: 3rd experiment

- Once again the coverage was too close: 53.45% from the architects and 52.07% from ATP.

Selector	TCs	Coverage	Failure
Architec	116	53.45%	0
t			
ATP	211	58.57%	7

Table 4: 4th experiment

- We used ATP directly in the whole test database of our industrial partner, related to this product. In this fourth experiment, we did not use a preselected test cases subset (called master plans) as previous experiments
- With 211 selected test cases, ATP got a code coverage of 58.57%. This was 5.12% superior to architect's selection, and it was enough to ATP's test plan reveal 7 new failures

- We noted that we could never have a plan with greater coverage than the master plan because the master plan is always fully executed by our industrial partner.

- There is no further selection from test architects. Even though, ATP has provided a close code coverage and with fewer test cases

- The second experiment was a very particular situation that currently should not occur in practice. We tried to observe what happens if we use a shorter (compared to current practice) time period to perform a regression

- By deciding to avoid having the master plan as a preliminary filter in our fourth experiment:
    - Our attempt to increase code coverage beyond the master plan was not so impressive because we had only a 5.12% gain in code coverage.
    - Our attempt to eventually finding bugs was very interesting because we found 7 new unreported failures with this small improvement

Keywords:

Information Retrieval;
Test Case selection and prioritization;
Code coverage;

Who is this briefing for?

Software engineering practitioners who want to make decisions about regression test case selection based on scientific evidence.

Where the findings come from?

All findings of this briefing were extracted from the empirical research conducted by Claudio Magalhães et al.

What is included in this briefing?

Briefing about evidence and the main findings of the study conducted.

What is not included in this briefing?

Detailed descriptions of what was done and how the studies were done.