Technical Report
An Empirical Study on Factors impacting the Commits' Bug-proneness

Michele Tufano

December 10, 2013

1 Introduction

The aim of the work is to identify commits' characteristics impacting the likelihood of introducing bugs. We mined versioning systems and issue trackers of three open source systems comparing the following characteristics between buggy and not buggy commits:

- *Size*: the size of the commit expressed in terms of LOC and Number of Modified Files.
- *Quality metrics*: the Coupling Between Objects (CBO - Coupling), the Lack of COhesion of Methods (LCOM - Cohesion), the Number of Methods (NOM - Size), and the Weighted Methods per Class (WMC - Complexity) of the code components object of the commit.
- *Focus*: how similar are (both textually and structurally) the code components object of the commits. The conjecture is that similar code components involved in commits are easier to manage for the developer (and thus decrease the likelihood of introducing a bug).
- *Developers knowledge of the committed files*: we measured the experience on the committed files of the developer performing the commit. This has been done by computing (i) the textual similarity between the files object of the commit and the files modified in the past by the developer performing the commit and (ii) the number of times the developer already touched the files object of the commits in the past. The conjecture is that the higher the developer's knowledge of the files object of the commit, the lower the likelihood of introducing bugs.
- *Developer's interferences*: let D_i and D_j be two developers, c_0 and c_1 two commits performed by developer D_i on file f. Commit c_2 performed by developer D_j is considered an interference if it involves the file f and was committed in the period of time between the commit c_0 and c_1. The conjecture is that the developer D_i's mental model of file f could be negatively affected by the commit performed by developer D_j, causing the introduction of a bug. We expect that the higher the number of interferences before the commit, the higher the likelihood of introducing bugs.

Note that we distinguish between buggy and not buggy commits by using the SZZ algorithm.
The detailed results are reported in the following pages. In summary, we found that:

- *Size*: commits introducing bugs involve more files and a larger number of LOCs than commits do not introducing bugs.
- *Quality metrics*: commits introducing bugs, with respect to those do not introducing bugs, are performed on code components having: (i) a lower

cohesion, (ii) more methods, and (iii) more complexity. We do not have clear results about coupling.

- *Focus*: commits introducing bugs, with respect to those do not introducing bugs, involve files with a lower focus, i.e., the files in the commits do not introducing bugs are more similar (note that currently we only have data for the textual metric CCBC).
- *Developers knowledge of the committed files*: surprisingly, the developers knowledge is higher for commits introducing bugs.
- *Developer's interferences*: there are much more interferences before the commits introducing bugs.

2 Commit's size

2.1 Number of Modified Files

Descriptive Statistics

System	Commit Type	Median	Mean
Apache Ant	BUG	3.00	4.54
	NO BUG	1.00	2.59
JMeter	BUG	2.00	4.36
	NO BUG	1.00	2.02
Xerces-J	BUG	2.00	6.52
	NO BUG	1.00	2.40

BoxPlots

Figures 1, 2, 3 show the related boxplots.

Wilcoxon Test

System	p-value	Cliff's Delta
Ant	$< 2.2e-16$	0.35 (medium)
JMeter	$< 2.2e-16$	0.36 (medium)
Xerces-J	$5.046e-08$	0.31 (small)

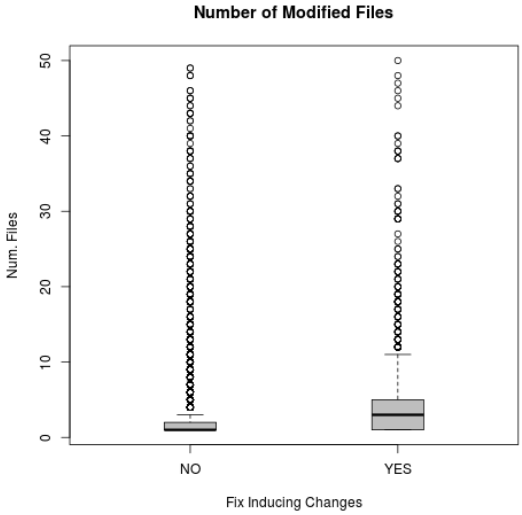

Figure 1: Number of Modified Files - Ant

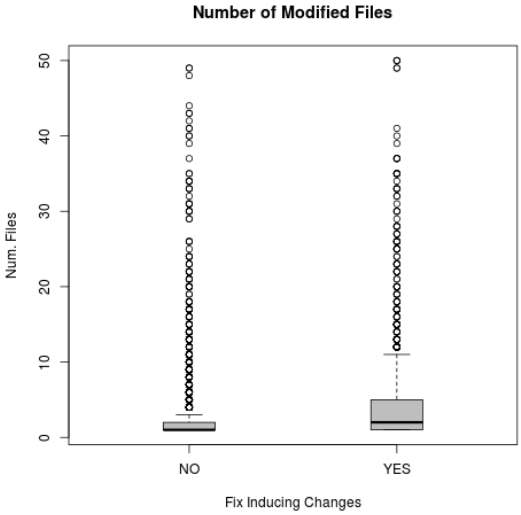

Figure 2: Number of Modified Files - JMeter

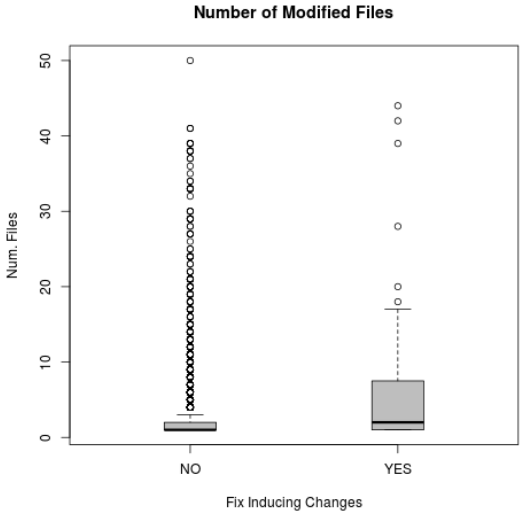

Figure 3: Number of Modified Files - Xerces-J

2.2 LOC

Descriptive Statistics

System	Commit Type	Median	Mean
Apache Ant	BUG	55.42	70.87
	NO BUG	48.00	65.76
JMeter	BUG	12.00	29.78
	NO BUG	10.00	24.71
Xerces-J	BUG	2.00	18.45
	NO BUG	5.41	15.43

BoxPlots

Figures 4, 5, 6 show the related boxplots.

Wilcoxon Test

System	p-value	Cliff's Delta
Ant	$7.427e-07$	0.08 (small)
JMeter	$1.428e-08$	0.10 (small)
Xerces-J	0.2025	0.06 (small)

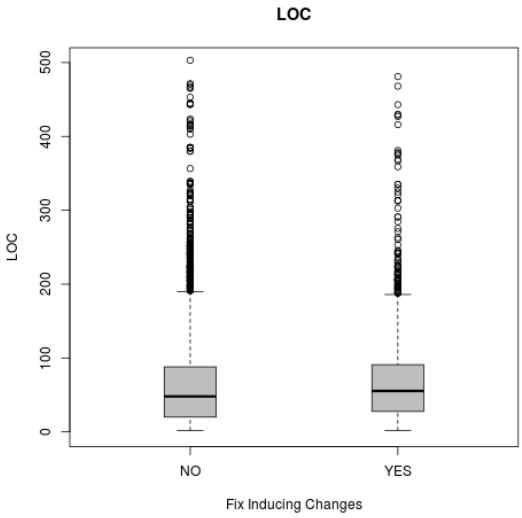

Figure 4: LOC - Ant

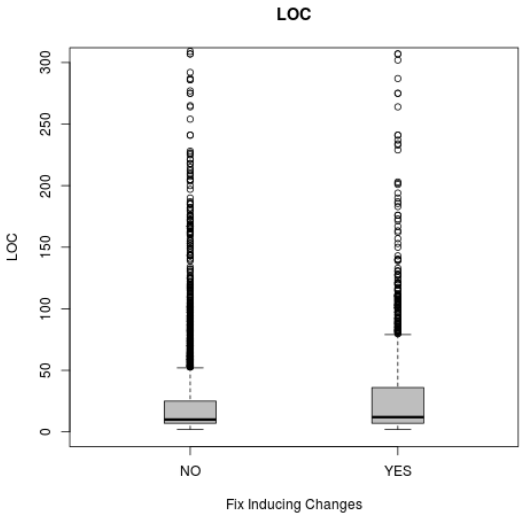

Figure 5: LOC - JMeter

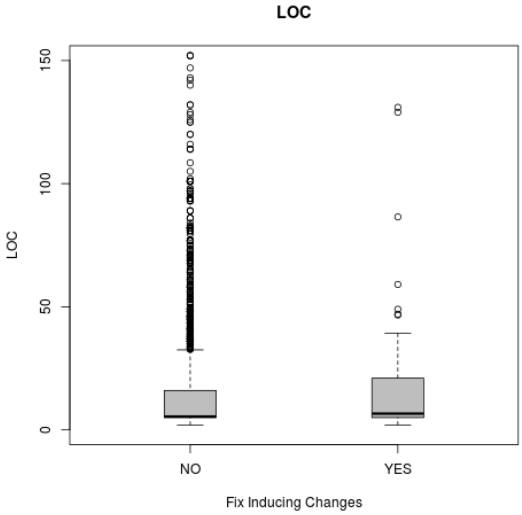

Figure 6: LOC - Xerces-J

3 Commit's Quality Metrics

3.1 CBO

Descriptive Statistics

System	Commit Type	Median	Mean
Apache Ant	BUG	10.26	12.20
	NO BUG	9.00	11.12
JMeter	BUG	16.50	20.75
	NO BUG	15.12	19.64
Xerces-J	BUG	7.22	15.26
	NO BUG	12.00	18.62

BoxPlots

Figures 7, 8, 9 show the related boxplots.

Wilcoxon Test

System	p-value	Cliff's Delta
Ant	$3.344e-09$	0.10 (small)
JMeter	$5.625e-05$	0.07 (small)
Xerces-J	0.9351	−0.11 (small)

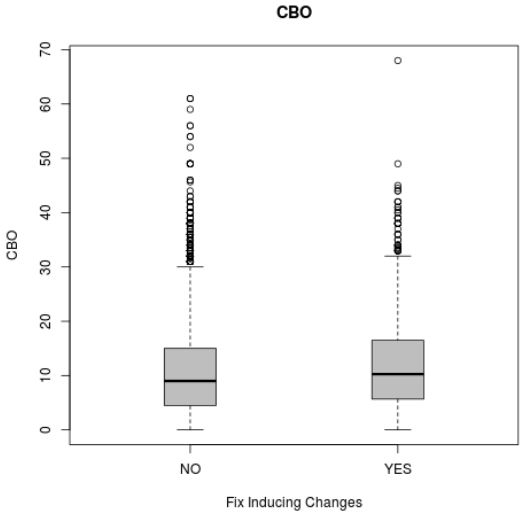

Figure 7: CBO - Ant

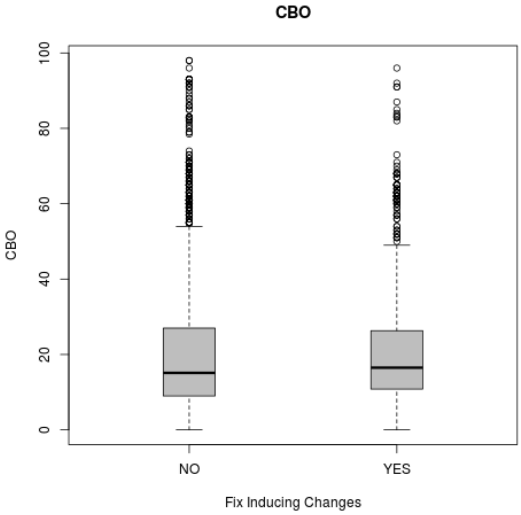

Figure 8: CBO - JMeter

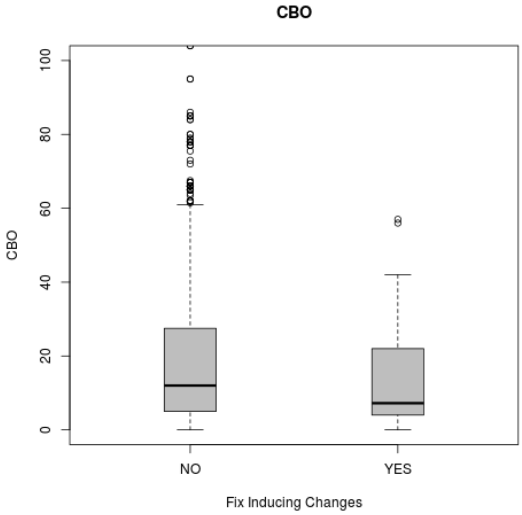

Figure 9: CBO - Xerces-J

3.2 LCOM

Descriptive Statistics

System	Commit Type	Median	Mean
Apache Ant	BUG	6.50	41.81
	NO BUG	3.00	32.62
JMeter	BUG	0.00	11.71
	NO BUG	0.00	8.17
Xerces-J	BUG	0.00	3.18
	NO BUG	0.00	3.06

BoxPlots

Figures 10, 11, 12 show the related boxplots.

Wilcoxon Test

System	p-value	Cliff's Delta
Ant	$3.947e-11$	0.11 (small)
JMeter	$3.285e-16$	0.12 (small)
Xerces-J	0.009621	0.12 (small)

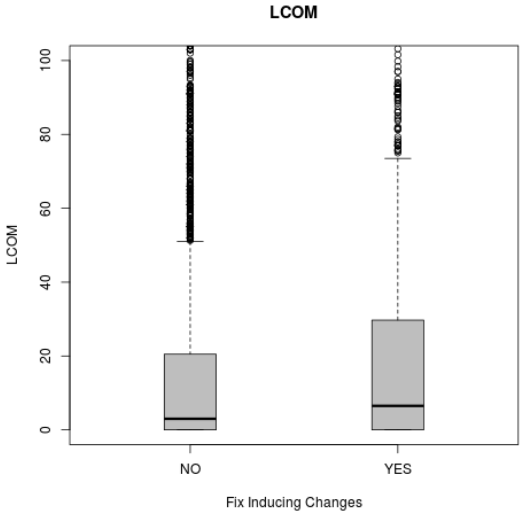

Figure 10: LCOM - Ant

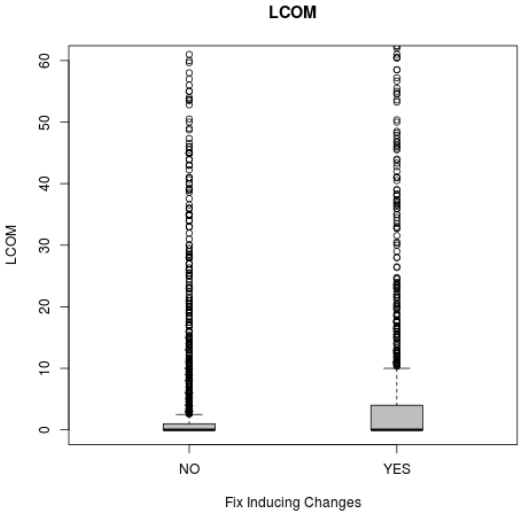

Figure 11: LCOM - JMeter

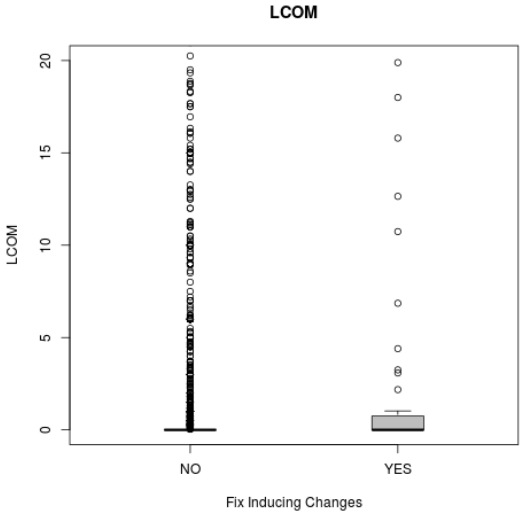

Figure 12: LCOM - Xerces-J

3.3 NOM

Descriptive Statistics

System	Commit Type	Median	Mean
Apache Ant	BUG	4.00	5.79
	NO BUG	3.50	5.54
JMeter	BUG	0.43	2.57
	NO BUG	0.00	2.06
Xerces-J	BUG	0.00	0.89
	NO BUG	0.00	0.89

BoxPlots

Figures 13, 14, 15 show the related boxplots.

Wilcoxon Test

System	p-value	Cliff's Delta
Ant	0.01199	0.04 (small)
JMeter	$2.723e-08$	0.09 (small)
Xerces-J	0.01245	0.14 (small)

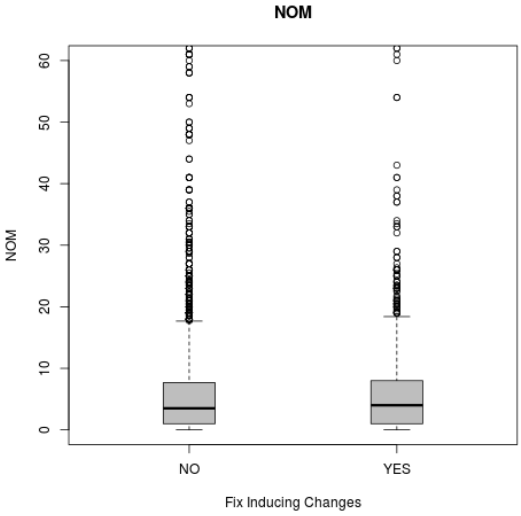

Figure 13: NOM - Ant

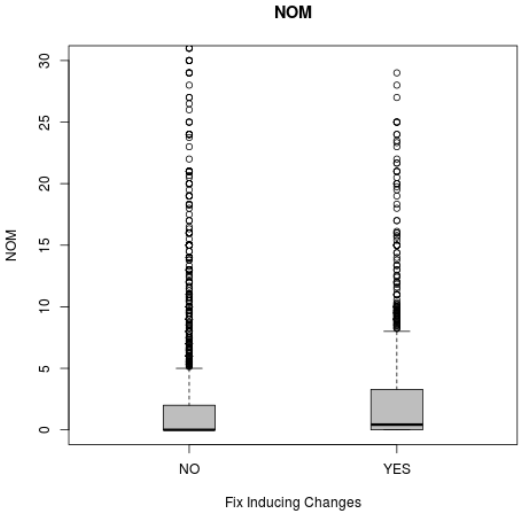

Figure 14: NOM - JMeter

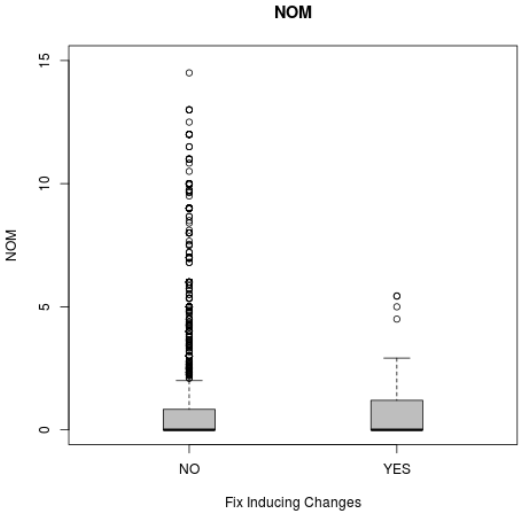

Figure 15: NOM - Xerces-J

3.4 WMC

Descriptive Statistics

System	Commit Type	Median	Mean
Apache Ant	BUG	8.66	12.60
	NO BUG	7.54	12.32
JMeter	BUG	0.50	5.30
	NO BUG	0.00	4.21
Xerces-J	BUG	0.00	2.29
	NO BUG	0.00	2.08

BoxPlots

Figures 16, 17, 18 show the related boxplots.

Wilcoxon Test

System	p-value	Cliff's Delta
Ant	0.03056	0.03 (small)
JMeter	$1.608e-09$	0.10 (small)
Xerces-J	0.008445	0.15 (small)

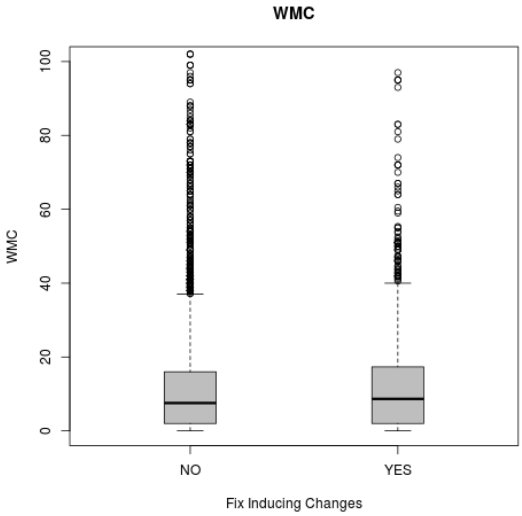

Figure 16: WMC - Ant

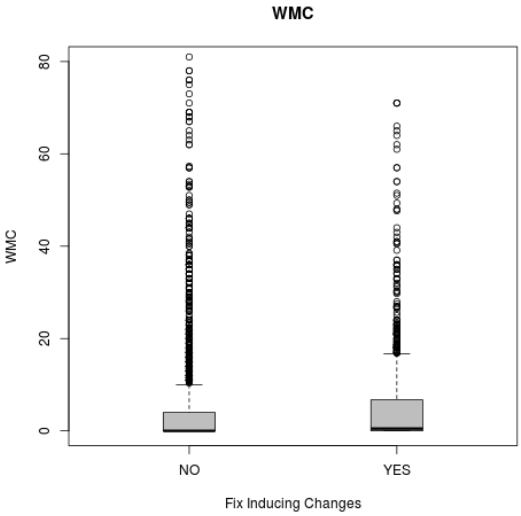

Figure 17: WMC - JMeter

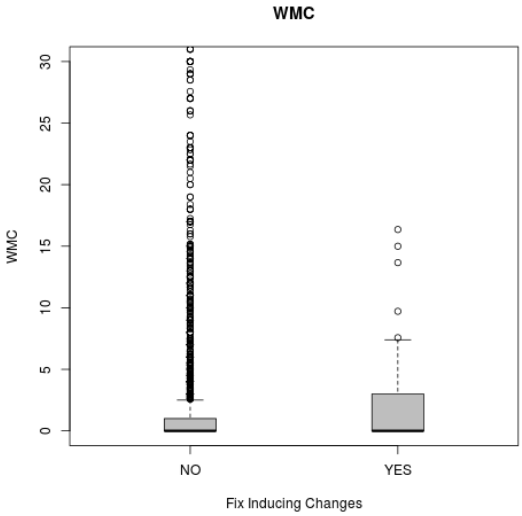

Figure 18: WMC - Xerces-J

4 Commit's Focus

Textual similarity between the files involved in the commits.

Descriptive Statistics

System	Commit Type	Median	Mean
Apache Ant	BUG	0.47	0.58
	NO BUG	1.00	0.78
JMeter	BUG	0.48	0.61
	NO BUG	1.00	0.83
Xerces-J	BUG	0.59	0.69
	NO BUG	1.00	0.83

BoxPlots

The figures 19, 20, 21 show the related boxplots.

Wilcoxon Test

System	p-value	Cliff's Delta
Ant	$< 2.2e-16$	0.33 (medium)
JMeter	$< 2.2e-16$	0.34 (medium)
Xerces-J	$7.151e-06$	0.26 (small)

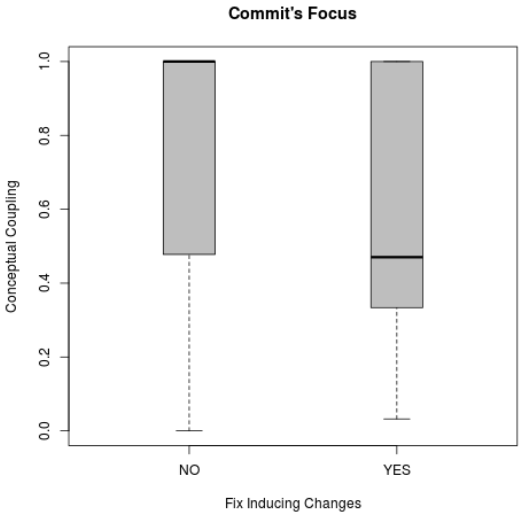

Figure 19: Commit's Focus - Ant

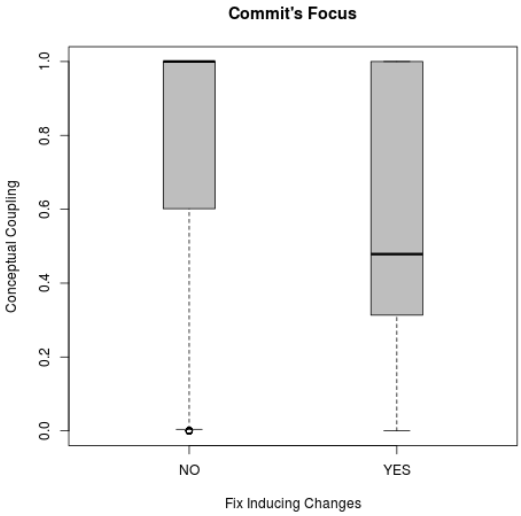

Figure 20: Commit's Focus - JMeter

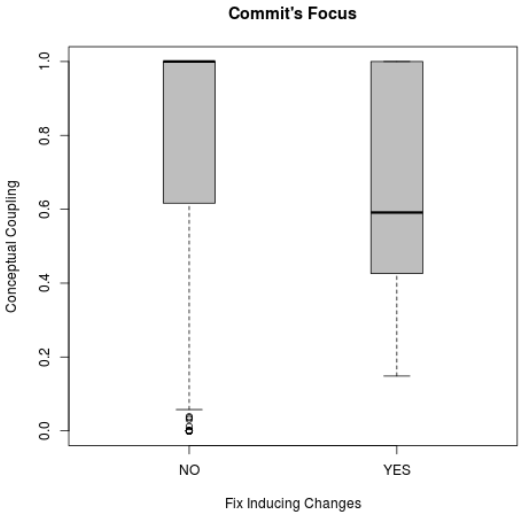

Figure 21: Commit's Focus - Xerces-J

5 Developer's Knowledge

5.1 Conceptual Coupling

Textual similarity between the files involved in the commit and the files modified in the past by the developer (i.e., developer's background). In particular, all files modified in the past by a developer are put in a single textual file, representing its background (note that if a file has been modified n times in the past, it is added n times to the background file).

5.1.1 Complete Background

We consider as developer's background all files modified in the past by her.

Descriptive Statistics

System	Commit Type	Median	Mean
Apache Ant	BUG	0.35	0.39
	NO BUG	0.29	0.33
JMeter	BUG	0.26	0.28
	NO BUG	0.21	0.24
Xerces-J	BUG	0.43	0.42
	NO BUG	0.30	0.33

BoxPlots

The figures 22, 23, 24 show the related boxplots.

Wilcoxon Test

System	p-value	Cliff's Delta
Ant	$< 2.2e-16$	0.26 (small)
JMeter	$< 2.2e-16$	0.24 (small)
Xerces-J	$1.48e-05$	0.31 (small)

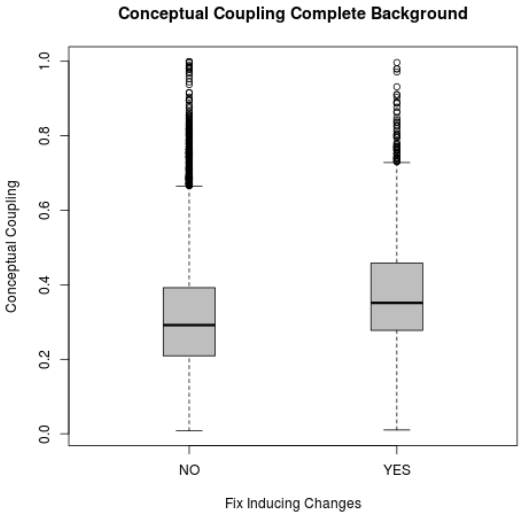

Figure 22: Conceptual Coupling CB - Ant

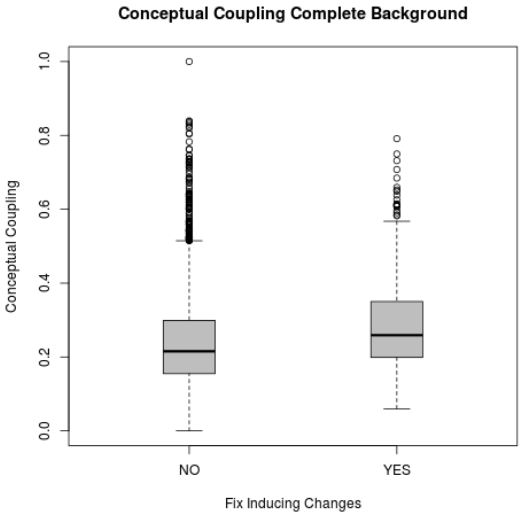

Figure 23: Conceptual Coupling CB - JMeter

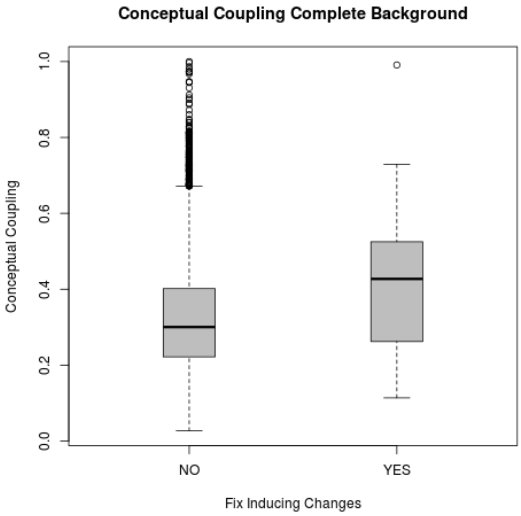

Figure 24: Conceptual Coupling CB - Xerces-J

5.1.2 Progressive Loss Of Memory Background

We consider as developer's background only the files modified by her in the last 6 months.

Descriptive Statistics

System	Commit Type	Median	Mean
Apache Ant	BUG	0.43	0.47
	NO BUG	0.38	0.41
JMeter	BUG	0.30	0.32
	NO BUG	0.24	0.27
Xerces-J	BUG	0.45	0.45
	NO BUG	0.34	0.37

BoxPlots

The figures 25, 26, 27 show the related boxplots.

Wilcoxon Test

System	p-value	Cliff's Delta
Ant	$< 2.2e-16$	0.22 (small)
JMeter	$< 2.2e-16$	0.25 (small)
Xerces-J	$6.843e-05$	0.28 (small)

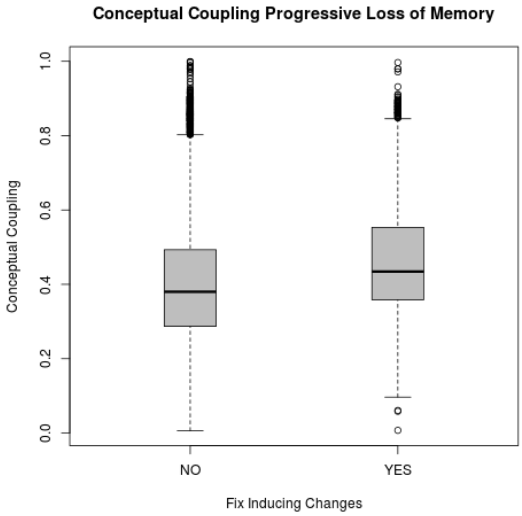

Figure 25: Conceptual Coupling PLoMB - Ant

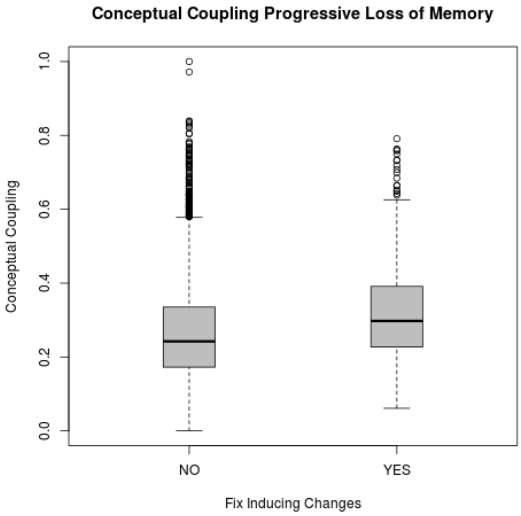

Figure 26: Conceptual Coupling PLoMB - JMeter

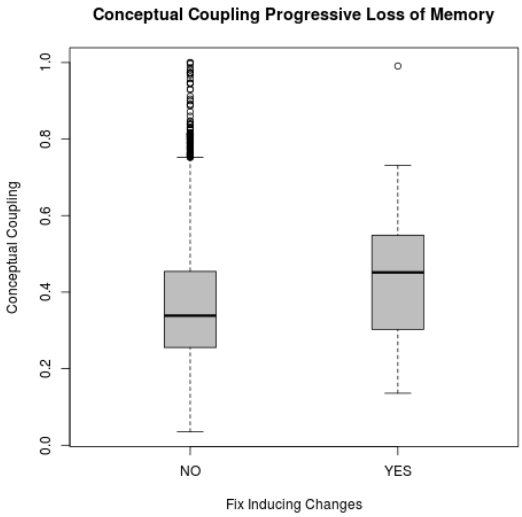

Figure 27: Conceptual Coupling PLoMB - Xerces-J

5.2 Knowledge

The higher the number of times a developer modified the files object of the commits in the past, the higher her knowledge.

In particular, the *knowledge* value is calculated as:

$$\frac{\#previousCommitOnInvolevedFiles}{\#previousCommitOnAllBackgroundFiles}$$

5.2.1 Complete Background

We consider as developer's background all files modified in the past by her.

Descriptive Statistics

System	Commit Type	Median	Mean
Apache Ant	BUG	0.012	0.042
	NO BUG	0.005	0.025
JMeter	BUG	0.014	0.042
	NO BUG	0.004	0.023
Xerces-J	BUG	0.015	0.094
	NO BUG	0.007	0.039

BoxPlots

The figures 28, 29, 30 show the related boxplots.

Wilcoxon Test

System	p-value	Cliff's Delta
Ant	$< 2.2e-16$	0.18 (small)
JMeter	$< 2.2e-16$	0.26 (small)
Xerces-J	0.004505	0.19 (small)

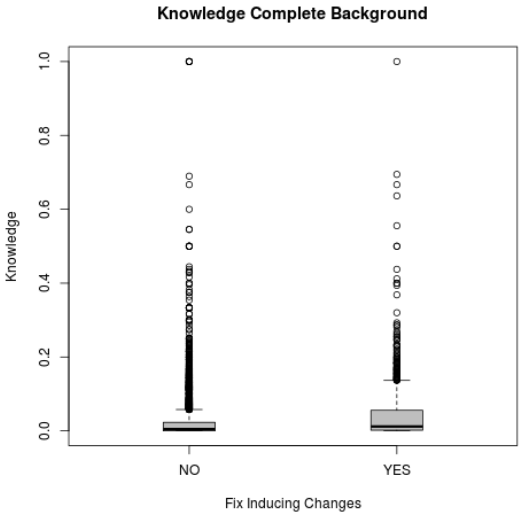

Figure 28: Knowledge CB - Ant

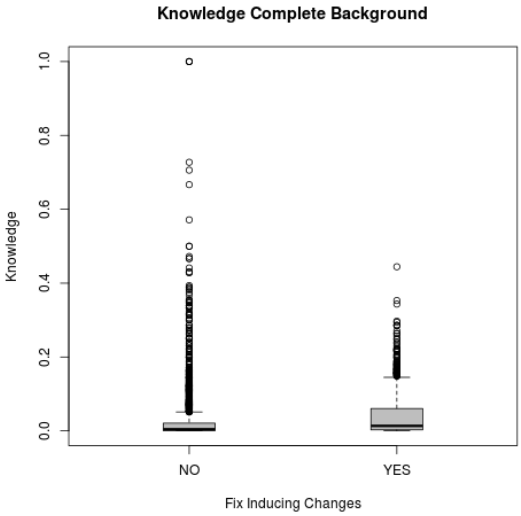

Figure 29: Knowledge CB - JMeter

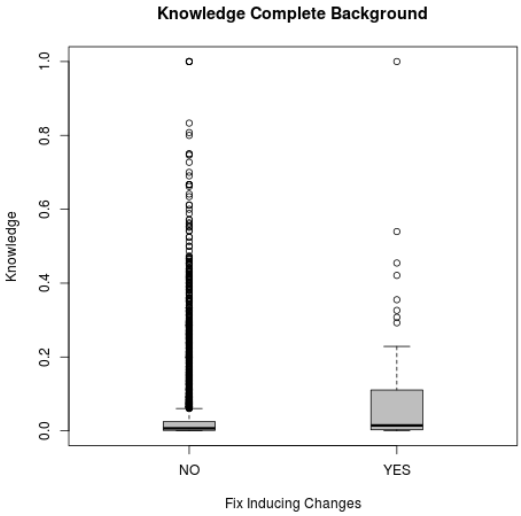

Figure 30: Knowledge CB - Xerces-J

5.2.2 Progressive Loss Of Memory Background

We consider as developer's background only the files modified by her in the last 6 months.

Descriptive Statistics

System	Commit Type	Median	Mean
Apache Ant	BUG	0.019	0.056
	NO BUG	0.006	0.041
JMeter	BUG	0.017	0.050
	NO BUG	0.005	0.027
Xerces-J	BUG	0.018	0.103
	NO BUG	0.010	0.044

BoxPlots

The figures 31, 32, 33 show the related boxplots.

Wilcoxon Test

System	p-value	Cliff's Delta
Ant	$5.96e-08$	0.09 (small)
JMeter	$< 2.2e-16$	0.26 (small)
Xerces-J	0.001195	0.22 (small)

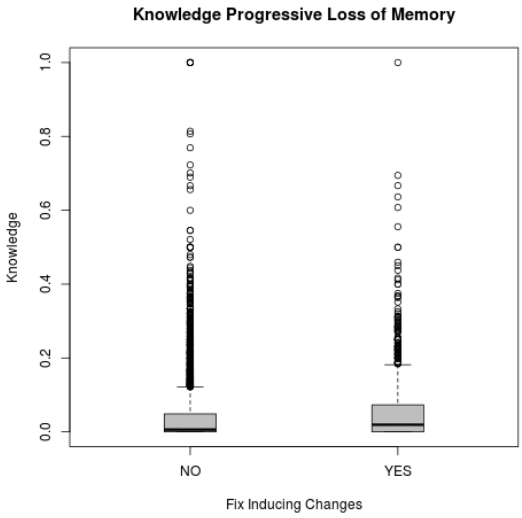

Figure 31: Knowledge PLoMB - Ant

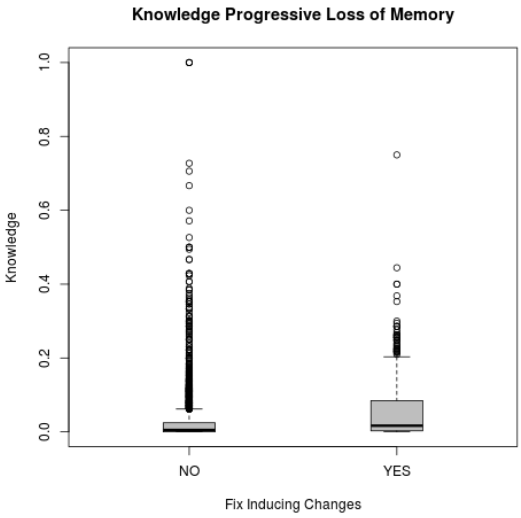

Figure 32: Knowledge PLoMB - JMeter

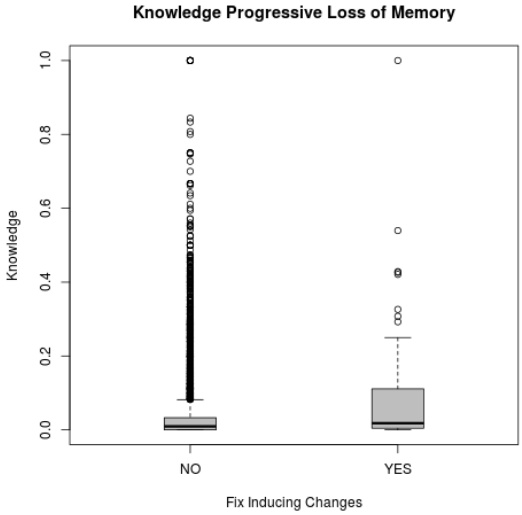

Figure 33: Knowledge PLoMB - Xerces-J

6 Developer's Interferences

6.1 Number of Interferences

Descriptive Statistics

System	Commit Type	Median	Mean
Apache Ant	BUG	1.00	5.34
	NO BUG	0.00	2.71
JMeter	BUG	0.00	2.00
	NO BUG	0.00	0.78
Xerces-J	BUG	0.00	1.15
	NO BUG	0.00	1.31

BoxPlots

The figures 34, 35, 36 show the related boxplots.

Wilcoxon Test

System	p-value	Cliff's Delta
Ant	$< 2.2e-16$	0.21 (small)
JMeter	$< 2.2e-16$	0.09 (small)
Xerces-J	0.03188	0.10 (small)

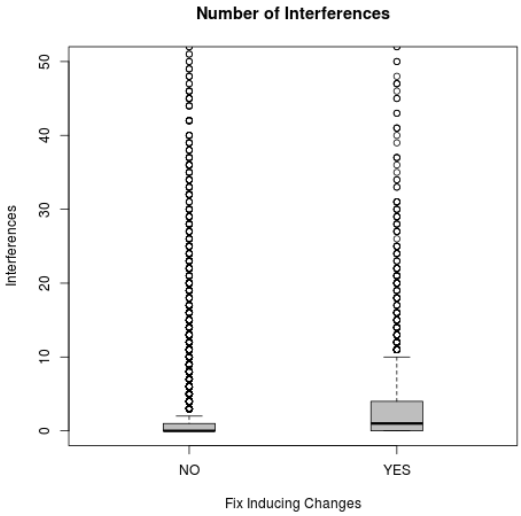

Figure 34: Number of Interferences - Ant

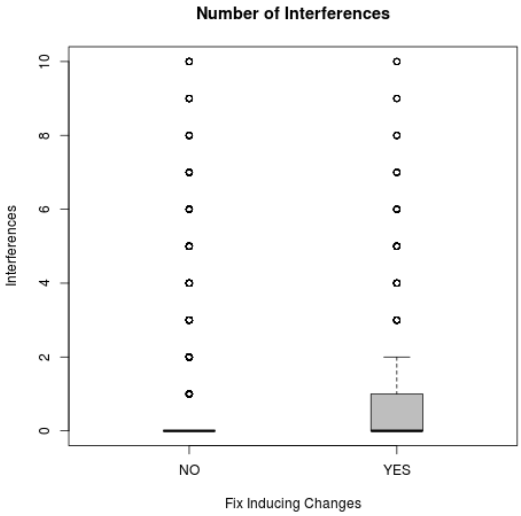

Figure 35: Number of Interferences - JMeter

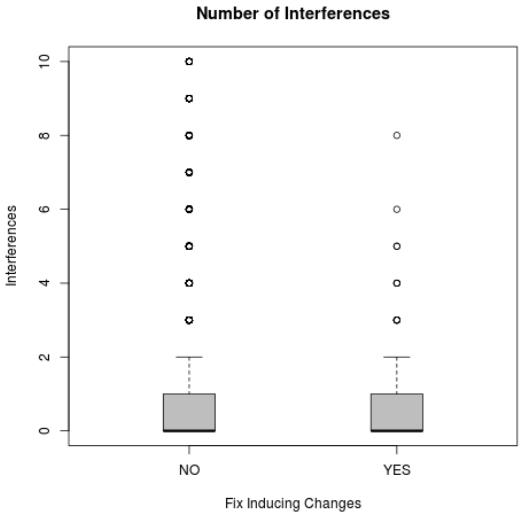

Figure 36: Number of Interferences - Xerces-J

6.2 Size of the Interferences in terms of LOCs

Descriptive Statistics

System	Commit Type	Median	Mean
Apache Ant	BUG	1.00	71.31
	NO BUG	0.00	37.99
JMeter	BUG	0.00	31.94
	NO BUG	0.00	10.51
Xerces-J	BUG	0.00	50.00
	NO BUG	0.00	38.88

BoxPlots

The figures 37, 38, 39 show the related boxplots.

Wilcoxon Test

System	p-value	Cliff's Delta
Ant	$< 2.2e-16$	0.21 (small)
JMeter	$< 2.2e-16$	0.09 (small)
Xerces-J	0.01991	0.11 (small)

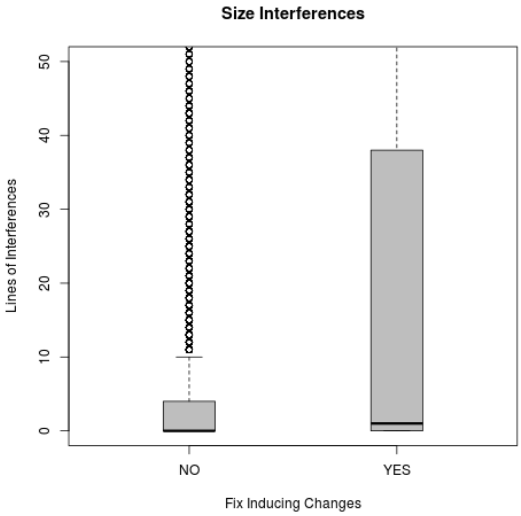

Figure 37: Size of Interferences - Ant

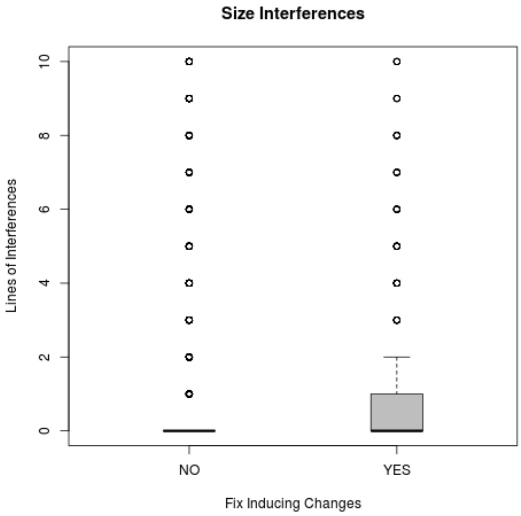

Figure 38: Size of Interferences - JMeter

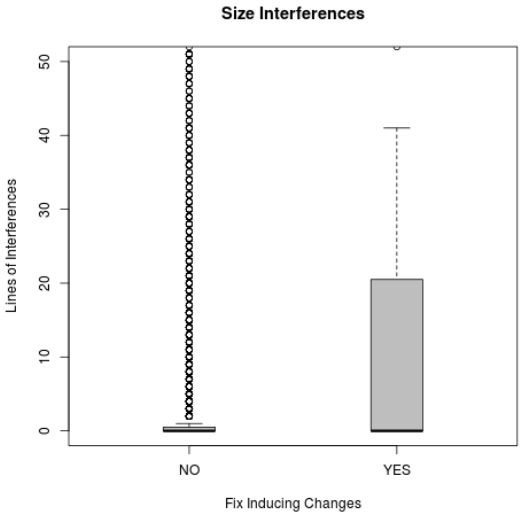

Figure 39: Size of Interferences - Xerces-J