Appendix - The Characteristics of False-Negatives in File-level Fault Prediction

Harold Valdivia-Garcia and Meiyappan Nagappan
Department of Software Engineering
Rochester Institute of Technology
Rochester, NY, USA
Email: hv1710@rit.edu, mei@se.rit.edu

Abstract

Over the years, a plethora of works has proposed more and more sophisticated machine learning techniques to improve fault prediction models. However, past studies using product metrics from closed-source projects, found a ceiling effect in the performance of fault prediction models. On the other hand, other studies have shown that process metrics are significantly better than product metrics for fault prediction. In our case study therefore we build models that include both product and process metrics taken together. We find that the ceiling effect found in prior studies exists even when we consider process metrics. We then qualitatively investigate the bug reports, source code files, and commit information for the bugs in the files that are false-negative in our fault prediction models trained using product and process metrics. Surprisingly, our qualitative analysis shows that bugs related to false-negative files and true-positive files are similar in terms of root causes, impact and affected components, and consequently such similarities might be exploited to enhance fault prediction models.

Identifying False-Negative Files

Table 1: Comparison between fault prediction models with and without PCA + SMOTE

		Original Model				Model + PCA + SMOTE			
Project	**Training $\rightarrow$ Testing release release**	**FN**	**TP**	**TN**	**FP**	**FN**	**TP**	**TN**	**FP**
Accumulo	1.4.0 → 1.5.0	70	16	792	29	66	20	744	77
	1.5.0 → 1.6.0	58	12	943	39	43	27	878	104
Bookkeeper	4.0.0 → 4.1.0	19	13	165	30	16	16	157	38
	4.1.0 → 4.2.0	16	9	260	26	13	12	254	32
Camel	2.9.0 → 2.10.0	136	11	2388	35	108	39	2281	142
	2.10.0 → 2.11.0	97	16	2661	78	68	45	2525	214
Cassandra	1.0.0 → 1.1.0	78	36	552	18	89	25	533	37
	1.1.0 → 1.2.0	107	49	725	41	107	49	706	60
CXF	2.2 → 2.3	48	43	2221	94	34	57	2031	284
	2.3 → 2.4	149	21	2420	23	134	36	2384	59
Derby	10.1.1.0 → 10.2.1.6	61	24	1395	78	47	38	1279	194
	10.2.1.6 → 10.3.1.4	118	12	1482	17	96	34	1441	58
Felix	scr-1.4.0 → scr-1.6.0	199	10	2363	5	160	49	2322	46
	scr-1.6.0 → scr-1.8.0	85	69	3120	138	60	94	2852	406
Hive	0.11.0 → 0.12.0	187	35	1998	26	152	70	1939	85
	0.12.0 → 0.13.0	156	54	2271	69	119	91	2161	179
OpenJPA	2.0.0 → 2.1.0	42	14	1405	28	33	23	1351	82
	2.1.0 → 2.2.0	29	6	1448	26	19	16	1392	82
Pig	0.9.0 → 0.10.0	47	9	969	16	40	16	918	67
	0.10.0 → 0.11	40	11	999	33	30	21	930	102
Wicket	1.3.0-final → 1.4.0	110	3	1223	2	99	14	1193	32
	1.4.0 → 1.5.0	86	21	1344	79	63	44	1218	205
All	_ _	1938 [5.3%]	494 [1.4%]	33144 [90.8%]	930 [2.5%]	1596 [4.4%]	836 [2.3%]	31489 [86.3%]	2585 [7.1%]

Quantitative analysis (RQ2-a).

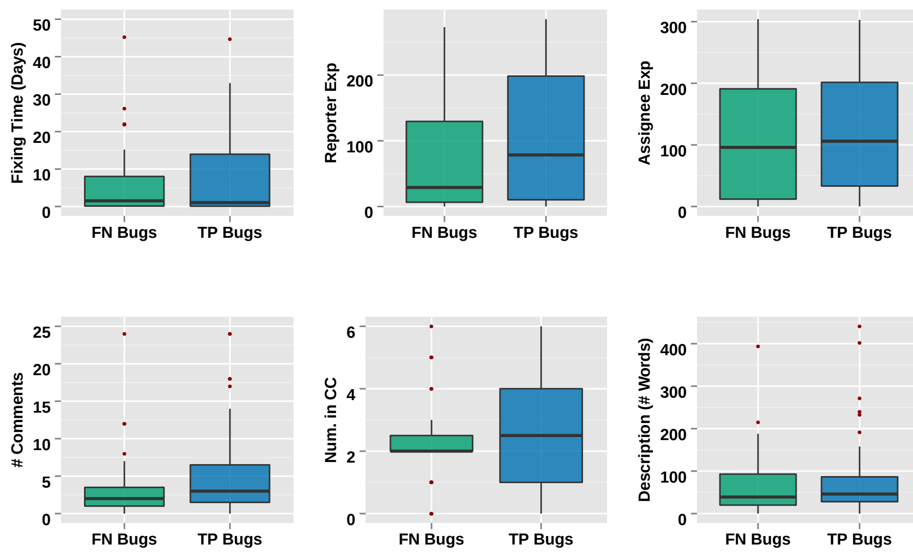

Figure 1: Boxplots of the metrics for FN-bugs and TN-bugs in Accumulo

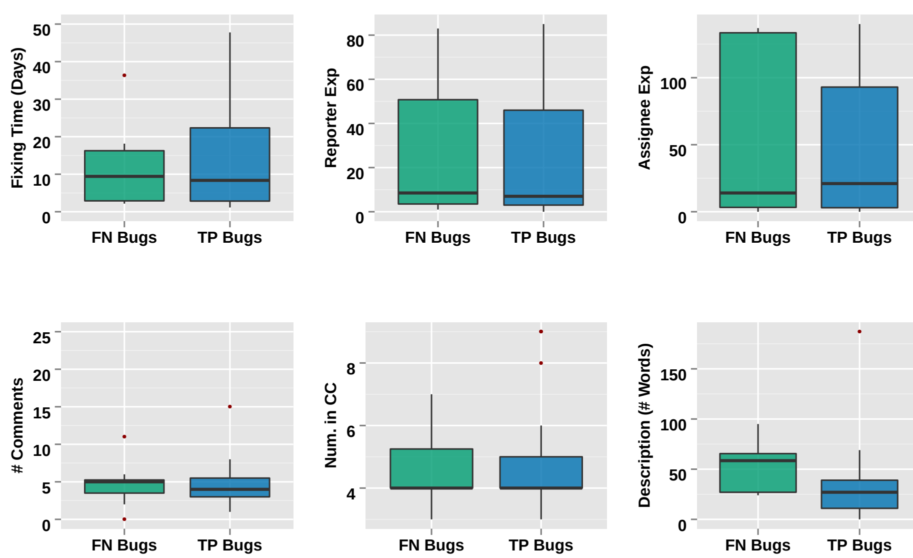

Figure 2: Boxplots of the metrics for FN-bugs and TN-bugs in Bookkeeper

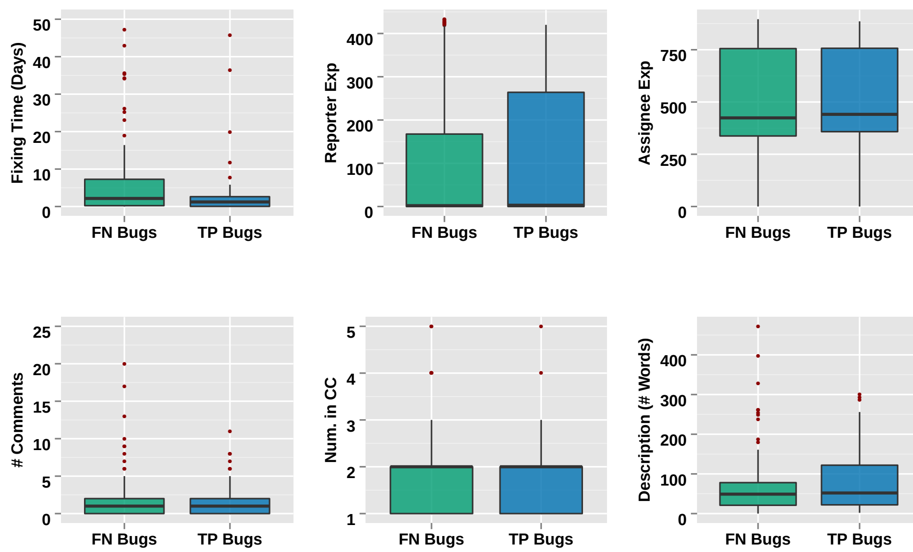

Figure 3: Boxplots of the metrics for FN-bugs and TN-bugs in Camel

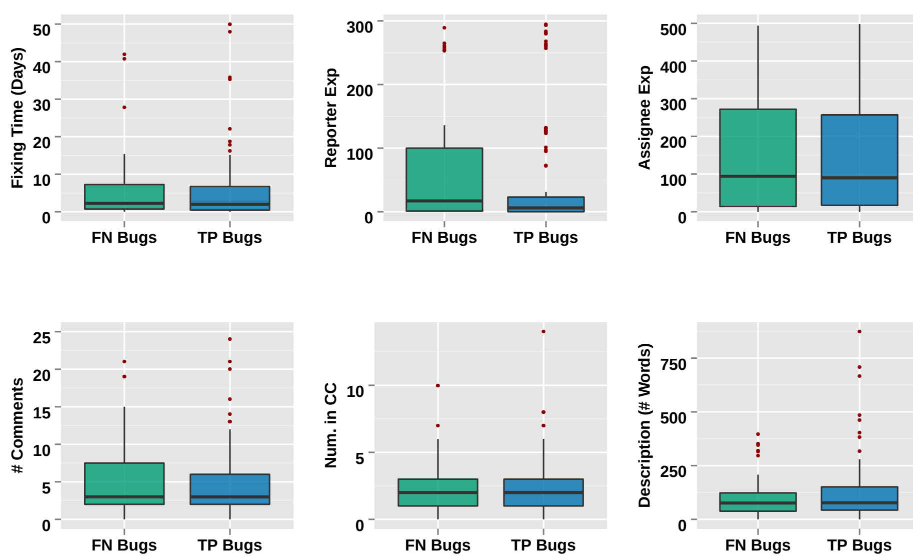

Figure 4: Boxplots of the metrics for FN-bugs and TN-bugs in Cassandra

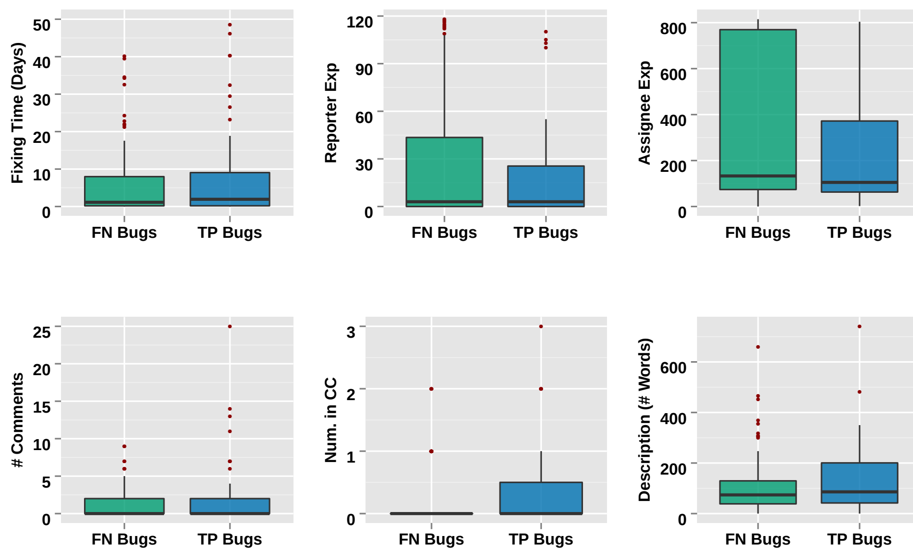

Figure 5: Boxplots of the metrics for FN-bugs and TN-bugs in CXF

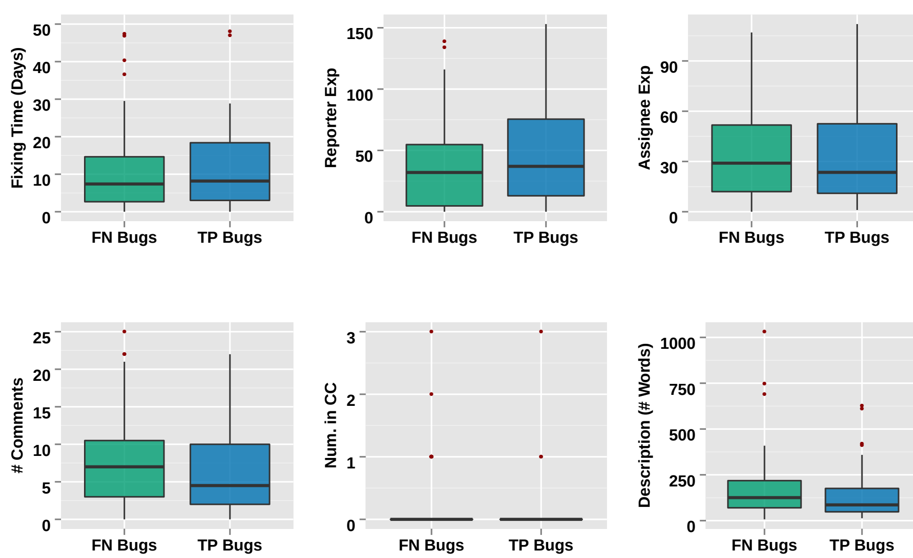

Figure 6: Boxplots of the metrics for FN-bugs and TN-bugs in Derby

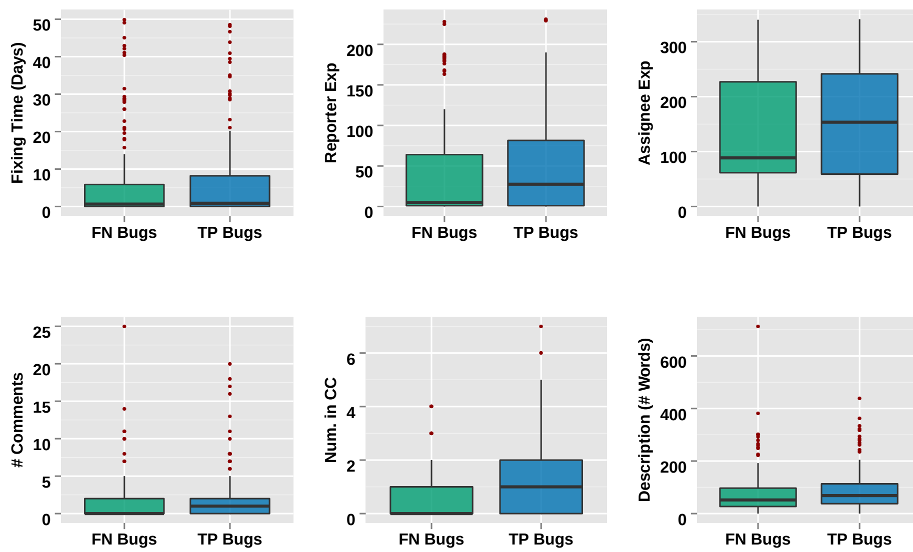

Figure 7: Boxplots of the metrics for FN-bugs and TN-bugs in Felix

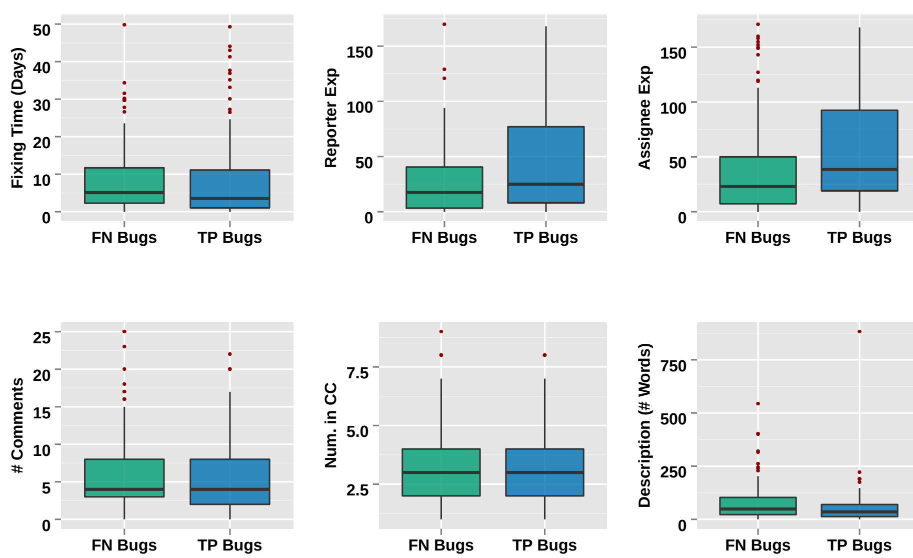

Figure 8: Boxplots of the metrics for FN-bugs and TN-bugs in Hive

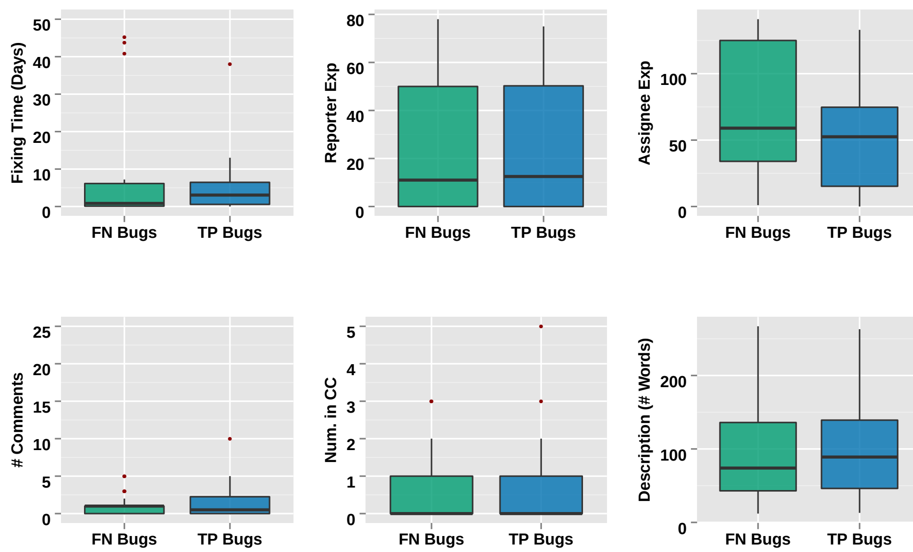

Figure 9: Boxplots of the metrics for FN-bugs and TN-bugs in OpenJPA

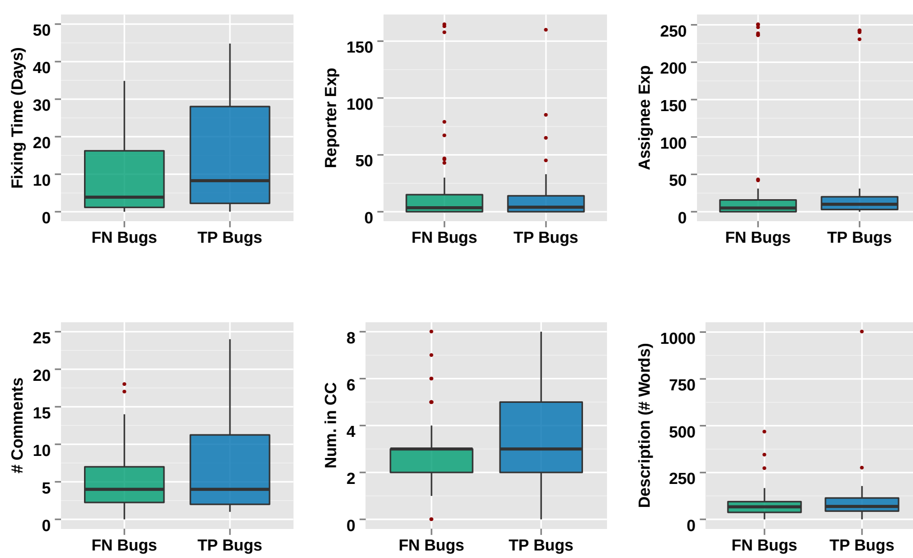

Figure 10: Boxplots of the metrics for FN-bugs and TN-bugs in Pig

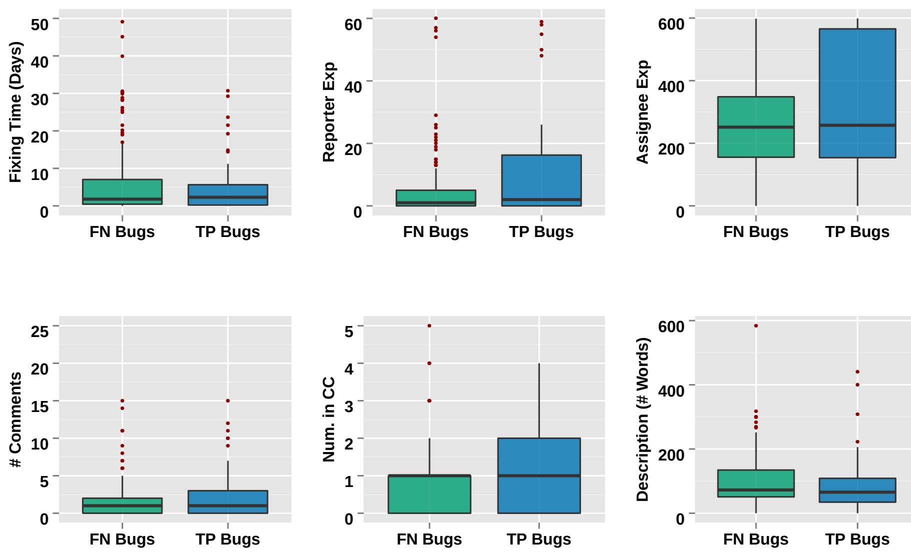

Figure 11: Boxplots of the metrics for FN-bugs and TN-bugs in Wicket

Qualitative analysis (RQ2-b). To support the categorization process of the bugs, we used the QDA Miner software from Provalis Research[1], which is a tool for qualitative analysis of text-based datasets. We provide all the data of our qualitative analysis (sources of information and the annotation of the categories) in our url address:

- https://github.com/harold-valdivia-garcia/fp-in-bug-pred/blob/master/fn-tp-bugs-qualitative-analysis.zip

1. http://provalisresearch.com/