An Empirical Study of the Effect of File Editing Patterns on Software Quality

Feng Zhang, Foutse Khomh, Ying Zou

and Ahmed E. Hassan

Motivation

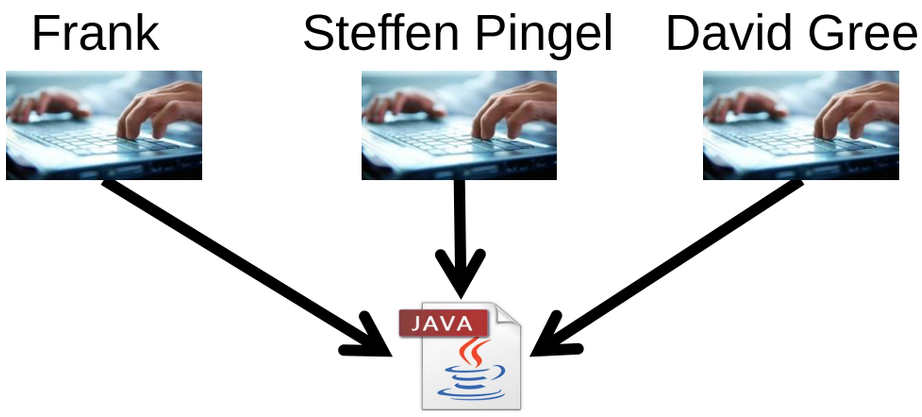
- Different developers may follow different file editing patterns.
- Different file editing patterns may happen together.

❖ Code change may introduce bugs

- ➢ Investigate risk of file editing patterns
- ➢ Investigate risk of patterns' interactions

Examples

Several developers edit the same file concurrently.



Concurrent Editing

Extended Editing

4.5 hours

DiscoveryViewer.java

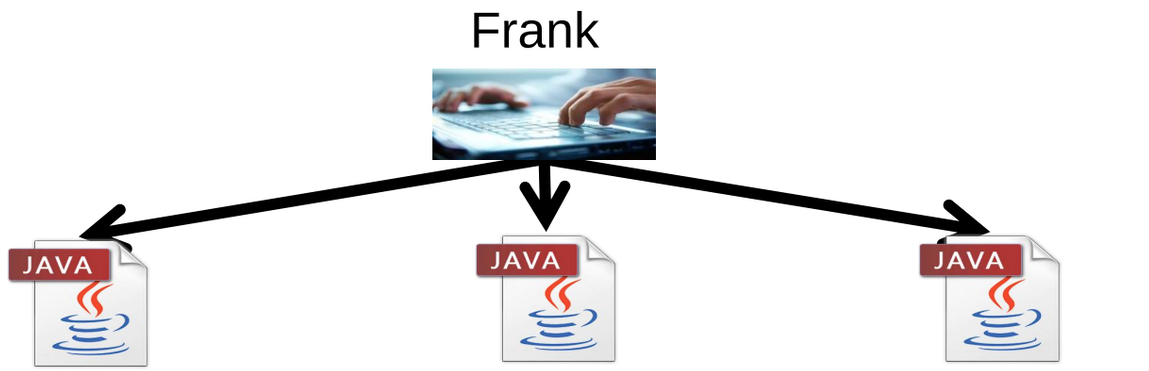
3.6 times

1.25 hours

Mylyn Project

Developers spend longer time editing a file. (threshold: third quantile)

Multiple files are edited in parallel by the same developer.



Parallel Editing

Interrupted Editing

338 hours

TaskCompareDialog.java

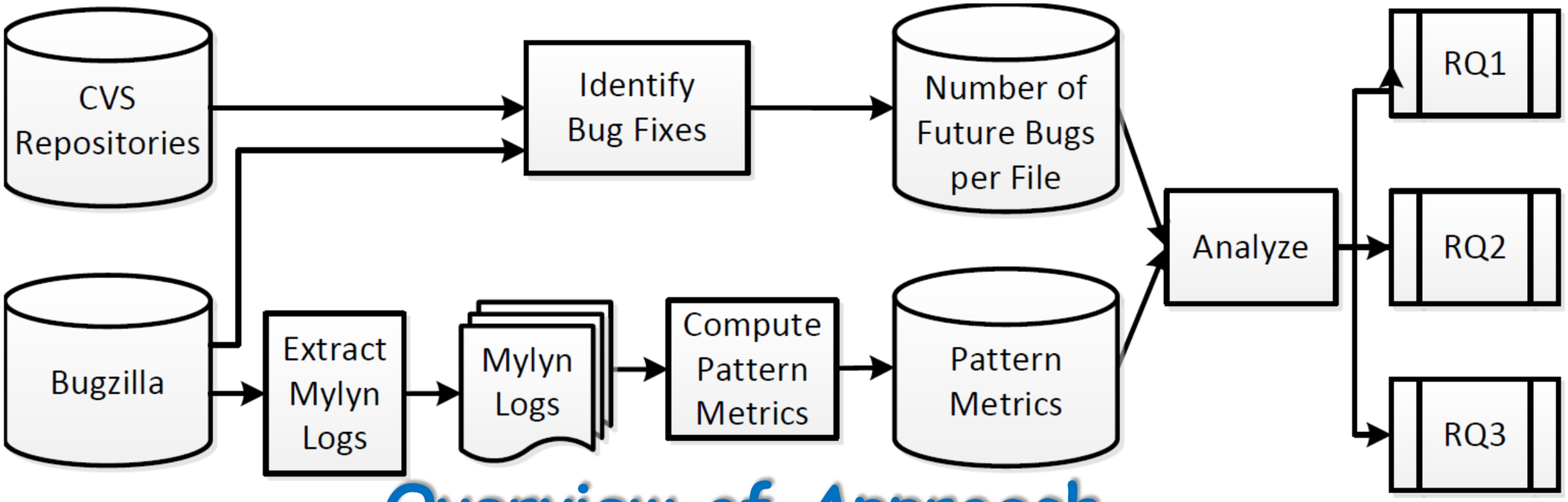
164 times

2 hours

Mylyn Project

There is a longer idle period during the editing of a file. (threshold: third quantile)

Case Study



Overview of Approach

Subject Systems

Mylyn	2,722 bugs	⟺	3,883 logs
Eclipse Platform	606 bugs	⟺	793 logs
PDE	524 bugs	⟺	638 logs

Description of Data

119 developers

2,140 files

5,070 CVS logs

98 bugs

Split Date: 2011-01-01

Research Questions

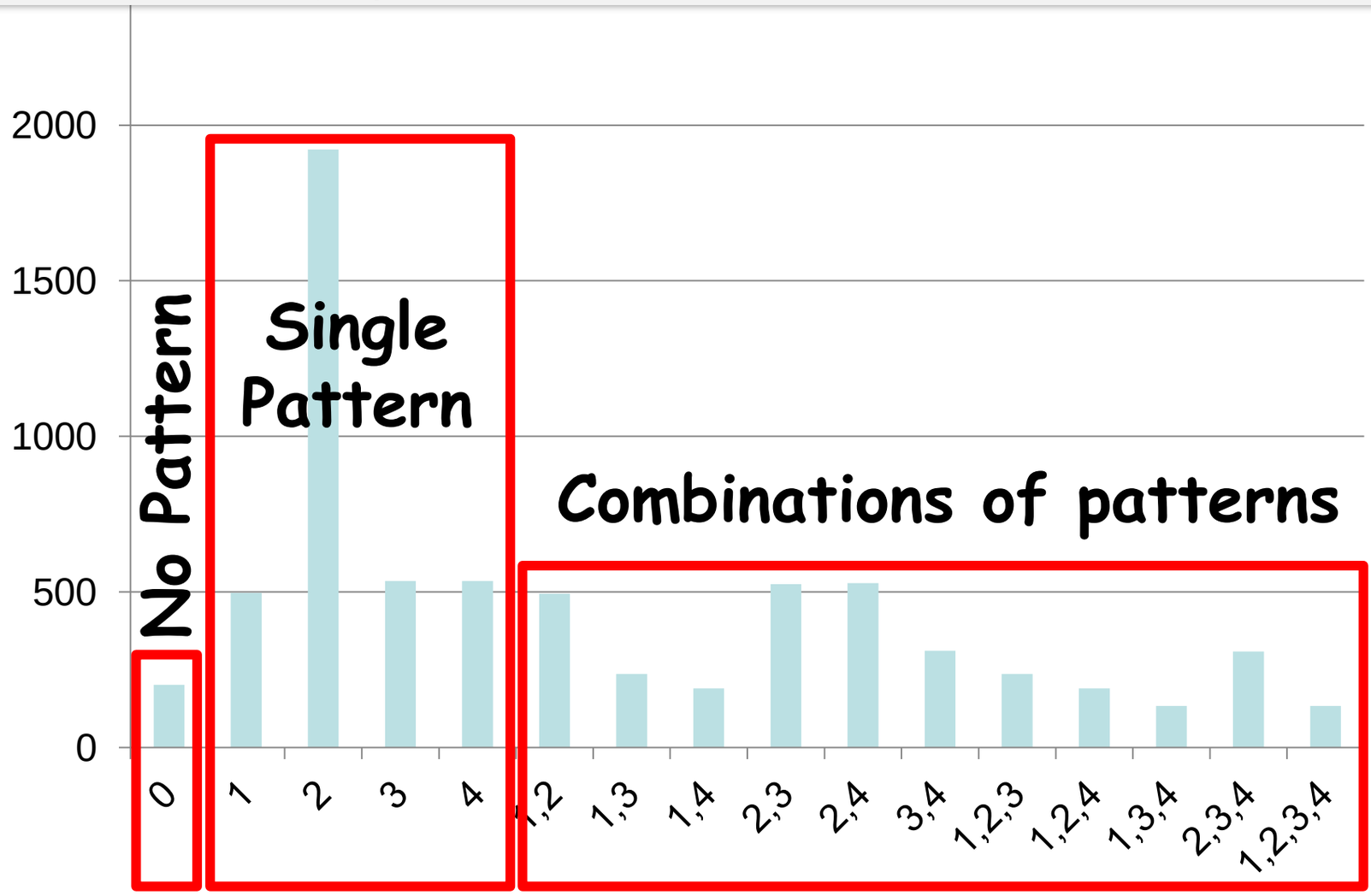
RQ1: Are there different file editing patterns?

RQ2: Do file editing patterns lead to more bugs?

RQ3: Do interactions among file editing patterns lead to more bugs?

RQ1: Are there different file editing patterns?

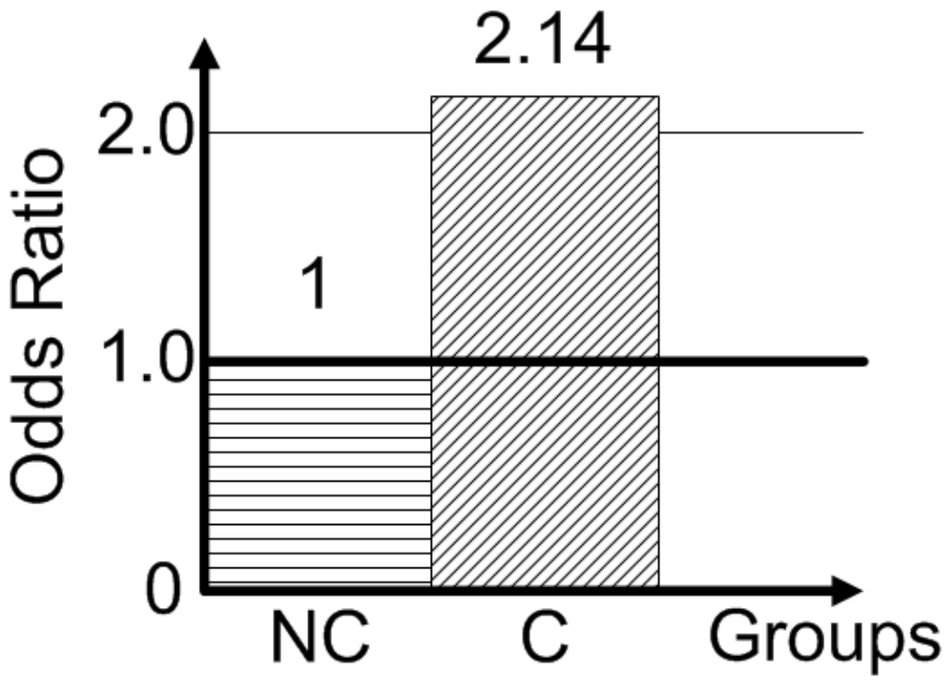
File editing patterns do exist



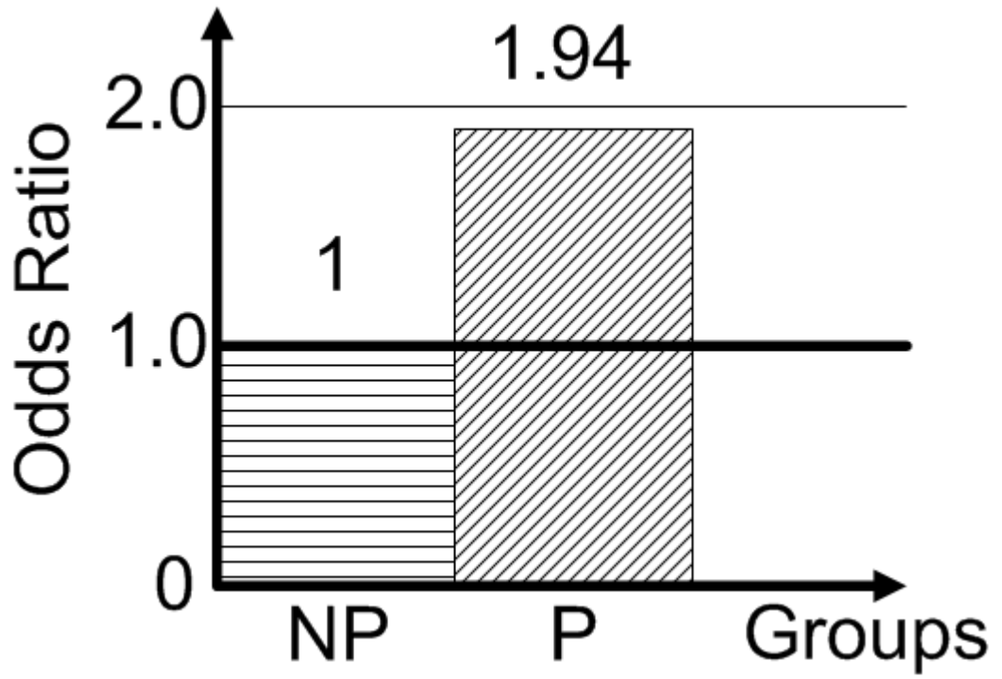
1. Concurrent
2. Parallel
3. Extended
4. Interrupted

Occurrences of file editing patterns and their interactions

RQ2: Risk of single editing pattern (Odds Ratio)

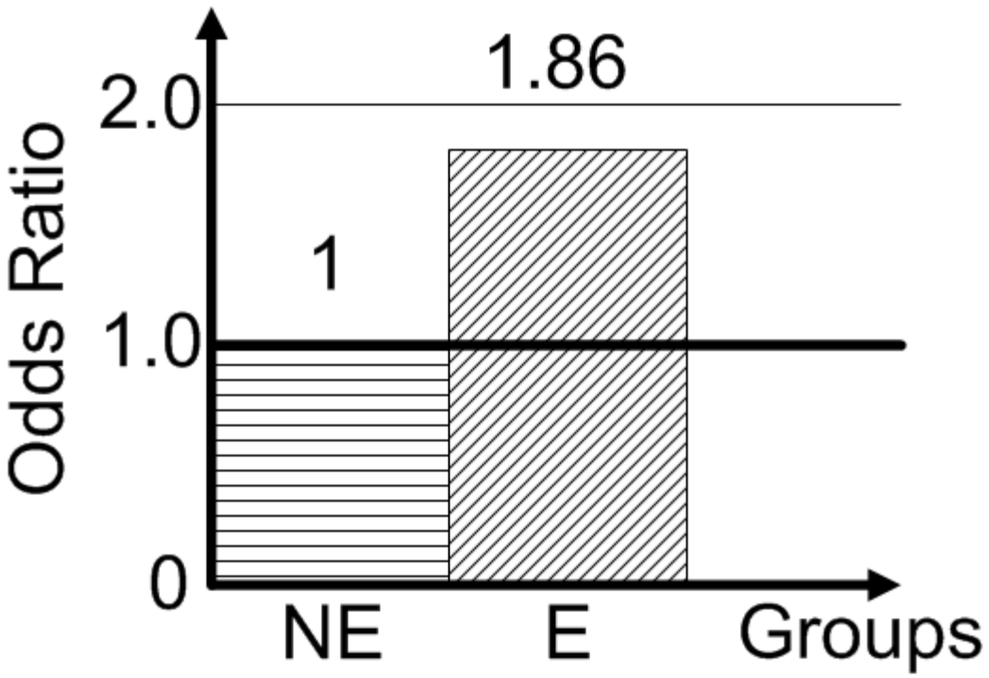


Concurrent Editing

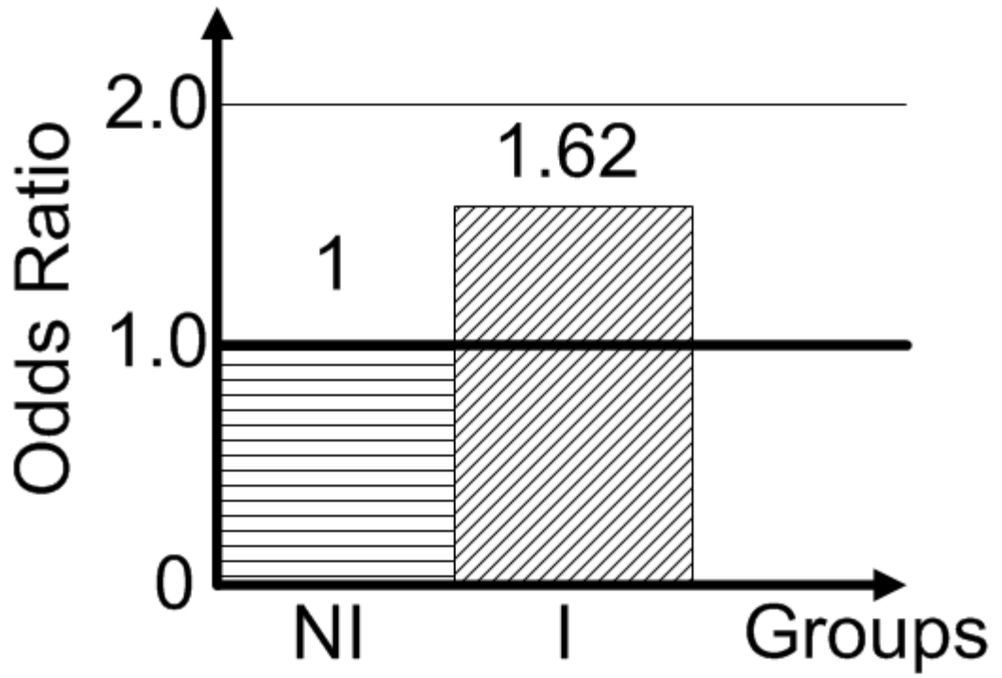


Parallel Editing

Extended Editing

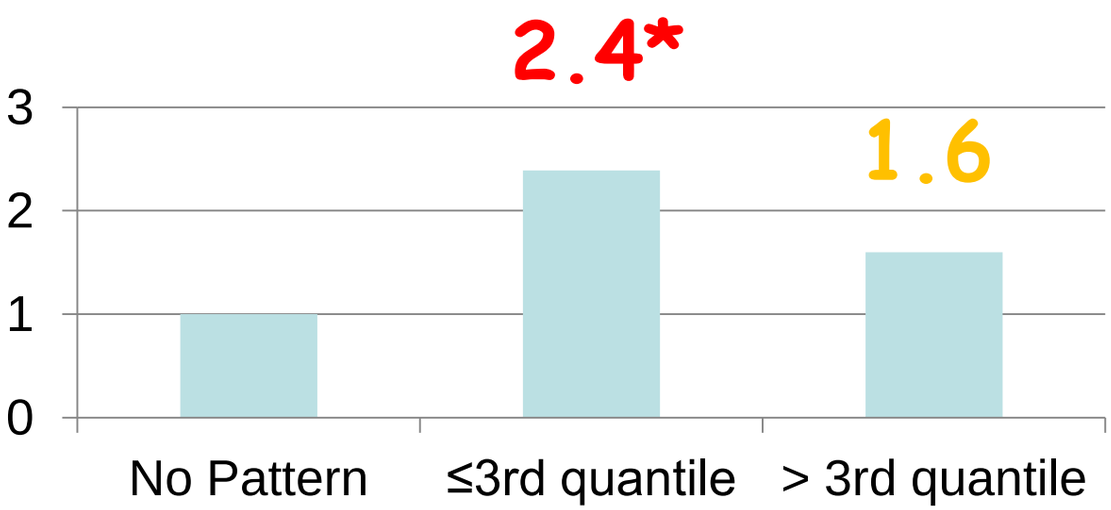


Interrupted Editing

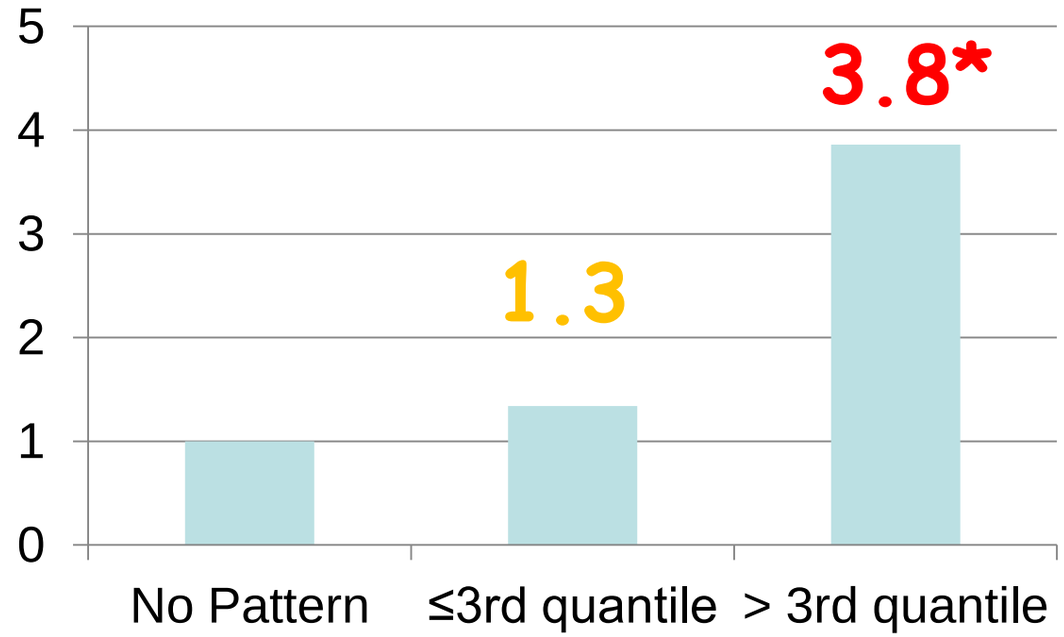


RQ2: Risk by level of involvement (Odds Ratio)

* Statistically significant

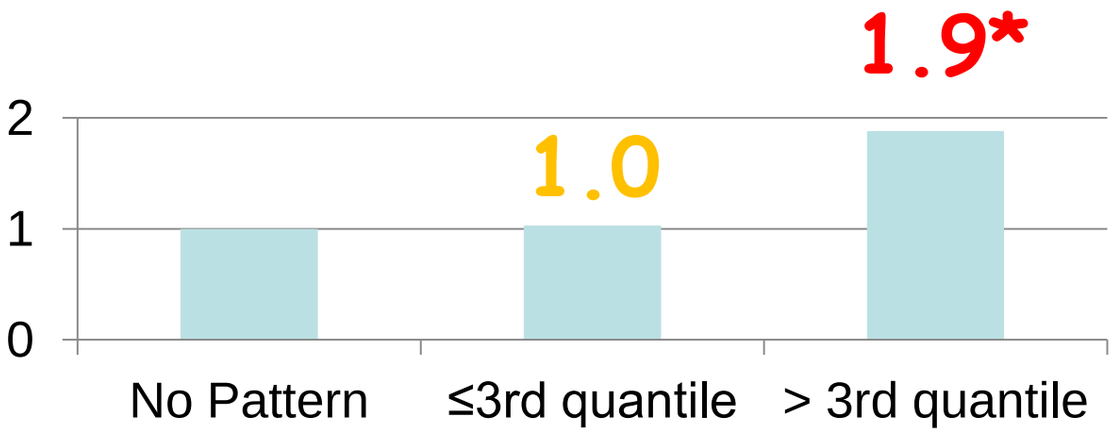


Concurrent Editing

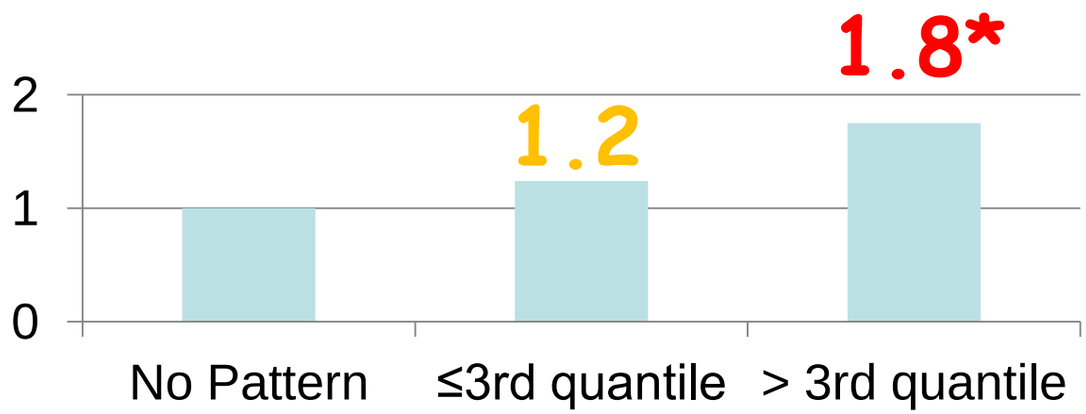


Parallel Editing

Extended Editing

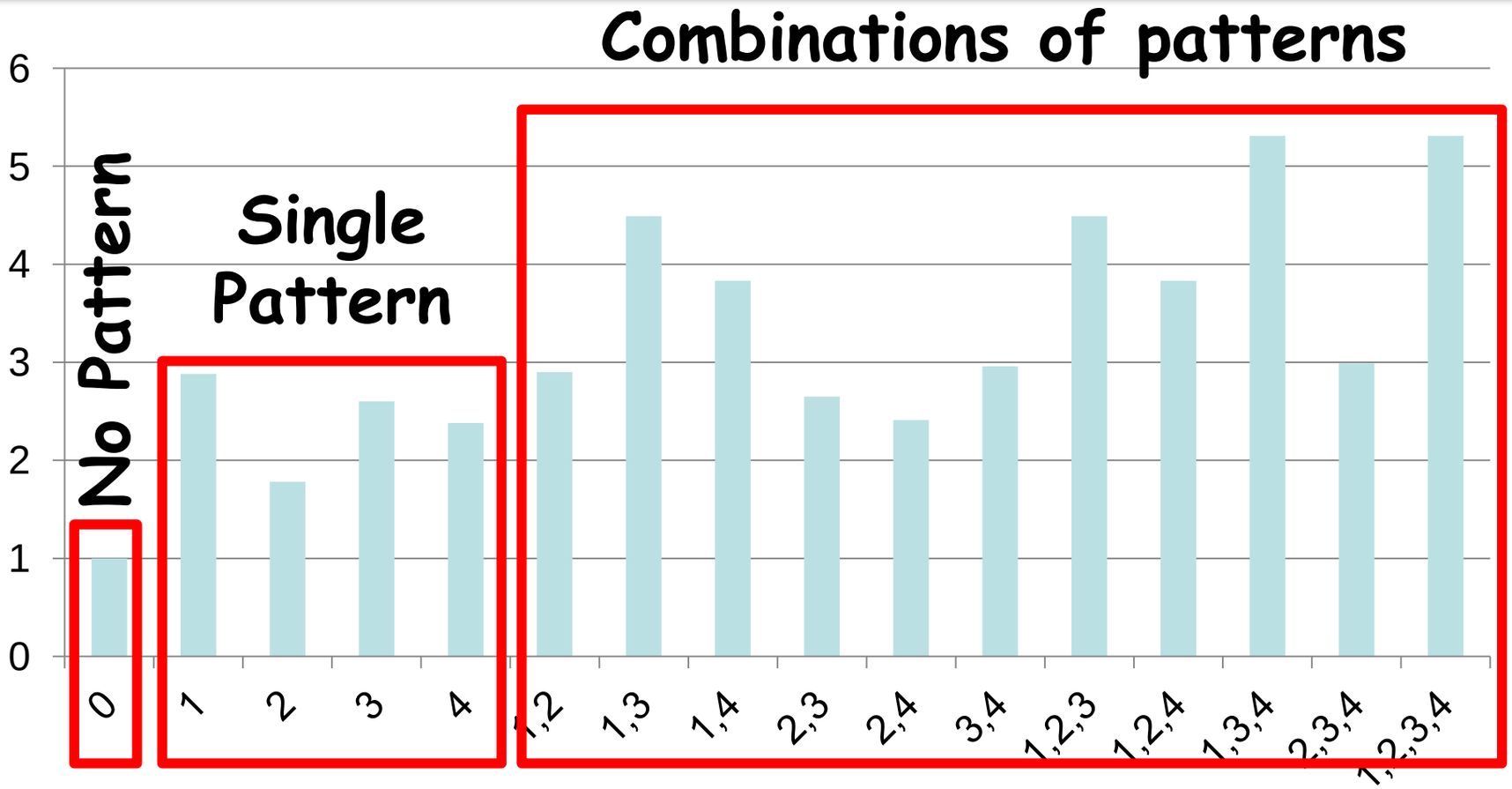


Interrupted Editing



RQ3: Do interactions among file editing patterns lead to more bugs?

Combinations of patterns are more risky than single pattern



1. Concurrent
2. Parallel
3. Extended
4. Interrupted

Odds ratio of 16 groups.

Conclusion

❑ Risk of Single File Editing Pattern

File Editing Pattern	Odds Ratio	Average Density of Bug
Concurrent Editing	2.14*	2.46*
Parallel Editing	1.94*	1.67*
Extended Editing	1.86*	2.28*
Interrupted Editing	1.62*	2.10*

❑ Risk of Combined File Editing Patterns

Average density of bug: up to 1.6* times than single pattern.

Feng Zhang, Foutse Khomh, Ying Zou, and Ahmed E. Hassan, **An Empirical Study of the Effect of File Editing Patterns on Software Quality**, Proceedings of the 19th Working Conference on Reverse Engineering **(WCRE)**, October 15-18, 2012, Kingston, Ontario, Canada. IEEE Computer Society Press.