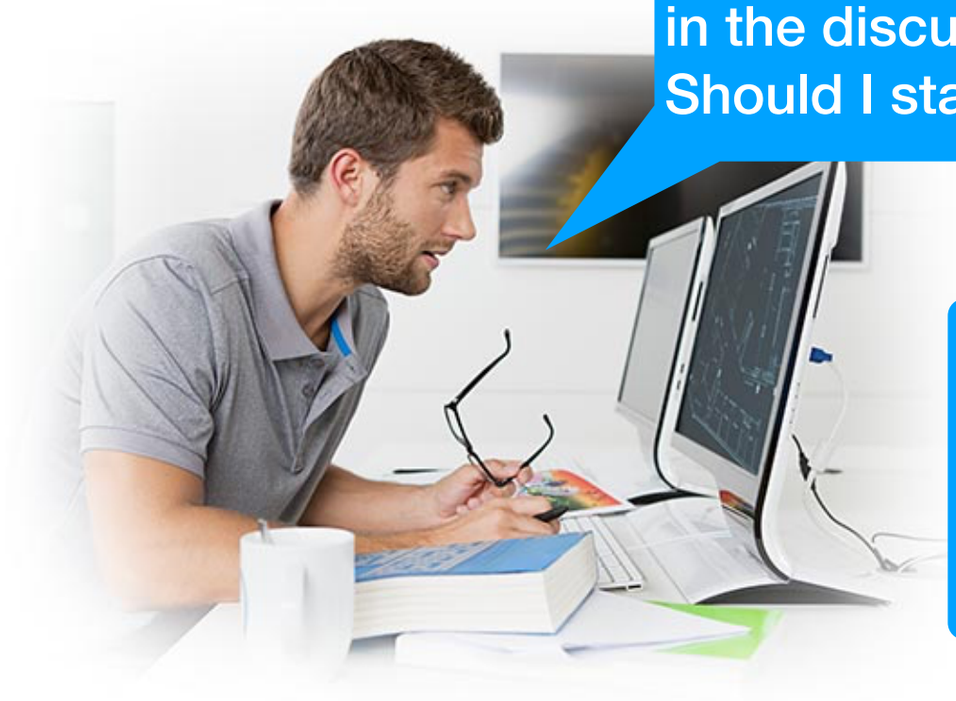
Recommending Related Code Reviews

DongGyun Han
28 Nov. 2017
https://dklab.net

Motivation



I've discussed about the similar issue like this long time ago and important context in the discussion, but I cannot locate it. Should I start the same discussion again?

My colleague asked me to review his change. Where should I start?

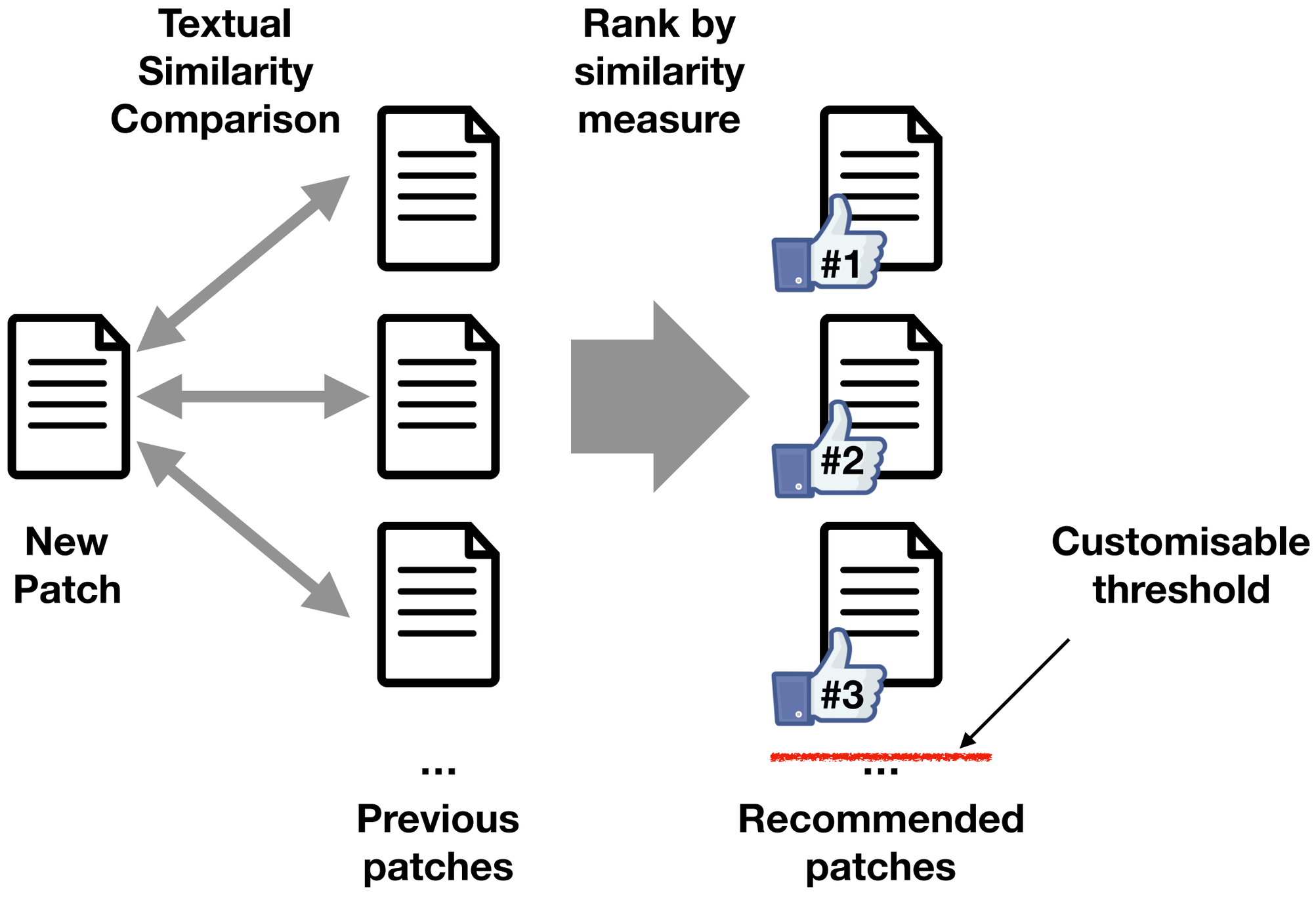
Motivation

Do you ever consult other/earlier reviews when doing a review?

82.65% (81 out of 98) open source developers

“(Referring other code review requests) is the most convenient way to understand what is going on in my team and the code base rather than go through the entire code base”

Related Code Review Recommendation



Requirements on similarity measurements

The similarity measurements must generally be applicable to compare patches (diffs) for any type of document, not just source code

The similarity measurements should be normalised

The similarity measurements must be computed efficiently

The similarity measurements should achieve high fidelity in their result

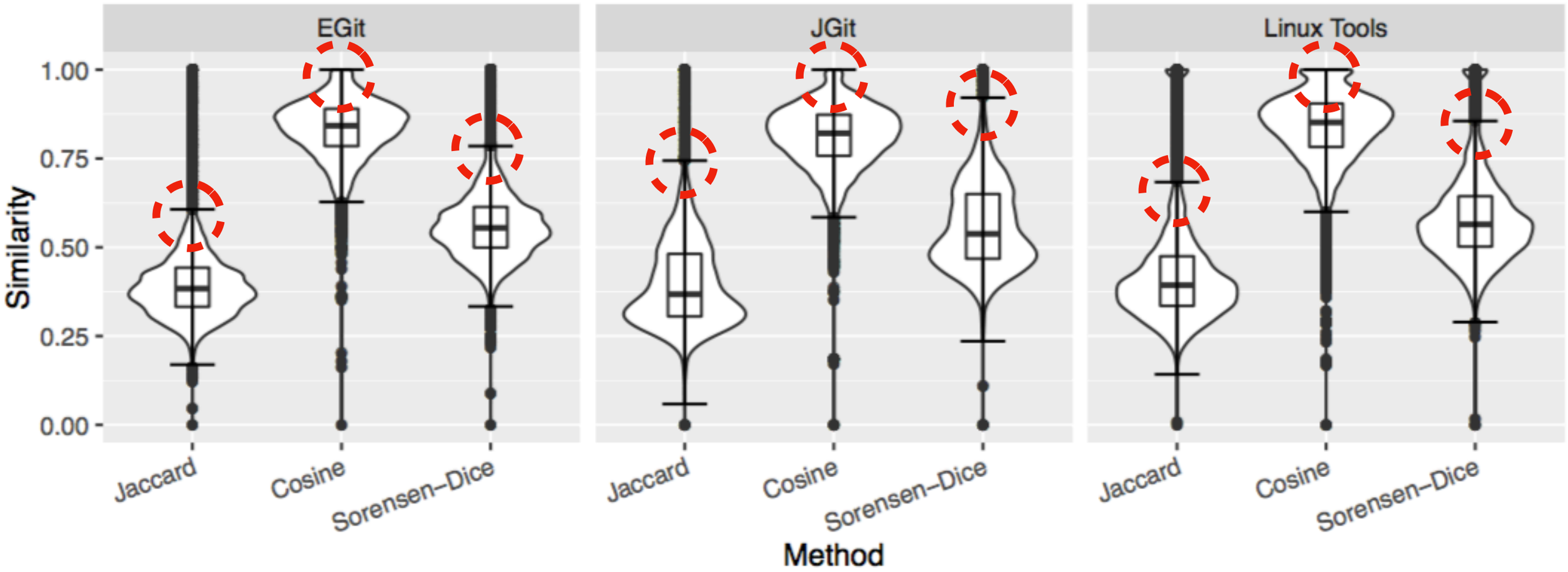
Jaccard

Sørensen-Dice

Cosine

Threshold (for evaluation)

Assumption: Similar patches are outliers of the similarity distribution



Research Questions

RQ1. How often are related patches recommended?

RQ2. How robust is the result compared to other similarity measurements?

RQ3. Do the recommended patches provide useful information during code review?

Research Questions

RQ1. How often are related patches recommended?

RQ2. How robust is the result compared to other similarity measurements?

RQ3. Do the recommended patches provide useful information during code review?

The number of recommendations

TABLE 3
Number of patches and code review requests having between 1 and 8 suggested code review request for each project

# of related review requests	**EGit**		**JGit**		**Linux Tools**		**Total**	
	Patches	Requests	Patches	Requests	Patches	Requests	Patches	Requests
1	83.9% (374)	91.0% (191)	98.2% (160)	96.3% (78)	94.7% (663)	91.8% (423)	91.4% (1197)	92.0% (692)
2	11.9% (53)	6.2% (13)	1.2% (2)	2.5% (2)	4.0% (28)	6.3% (29)	6.3% (83)	5.9% (44)
3	2.9% (13)	1.4% (3)	-	-	0.4% (3)	0.7% (3)	1.2% (16)	0.8% (6)
4	1.3% (6)	1.4% (3)	-	-	0.3% (2)	0.4% (2)	0.6% (8)	0.7% (5)
5	-	-	-	-	0.4% (3)	0.7% (3)	0.2% (3)	0.4% (3)
6	-	-	0.6% (1)	1.2% (1)	-	-	0.1% (1)	0.1% (1)
7	-	-	-	-	-	-	-	-
8	-	-	-	-	0.1% (1)	0.2% (1)	0.1% (1)	0.1% (1)
Total	**6.3%** (446)	**4.4%** (210)	**2.5%** (163)	**1.8%** (81)	**7.9%** (700)	**10.1%** (461)	**5.8%** (1309)	**5.5%** (752)

Identical patch resubmission

TABLE 4
Number of resubmitted identical patches

Project	EGit	JGit	Linux Tools
Patches	**0.7%** (49/7,050)	**0.4%** (24/6,457)	**2.7%** (252/9,232)
Reviews	**0.7%** (34/4,752)	**0.4%** (19/4,408)	**4.8%** (217/4,546)

TABLE 5
Reasons for resubmitting identical patches

Target branch	Category	EGit Reviews		EGit Patches		JGit Reviews		JGit Patches		Linux Tools Reviews		Linux Tools Patches	
Different branch	Cherry pick	76.47%	(26)	77.55%	(38)	78.95%	(15)	70.83%	(17)	88.48%	(192)	86.45%	(217)
	Merge	14.71%	(5)	16.33%	(8)	5.26%	(1)	12.50%	(3)	8.76%	(19)	11.16%	(28)
Same branch	Working on same base	5.88%	(2)	4.08%	(2)	5.26%	(1)	4.17%	(1)	2.30%	(5)	1.99%	(5)
	Mistake	2.94%	(1)	2.04%	(1)	-	-	-	-	0.46%	(1)	0.40%	(1)
	Take over	-	-	-	-	10.53%	(2)	12.50%	(3)	-	-	-	-

Research Questions

RQ1. How often are related patches recommended?

RQ2. How robust is the result compared to other similarity measurements?

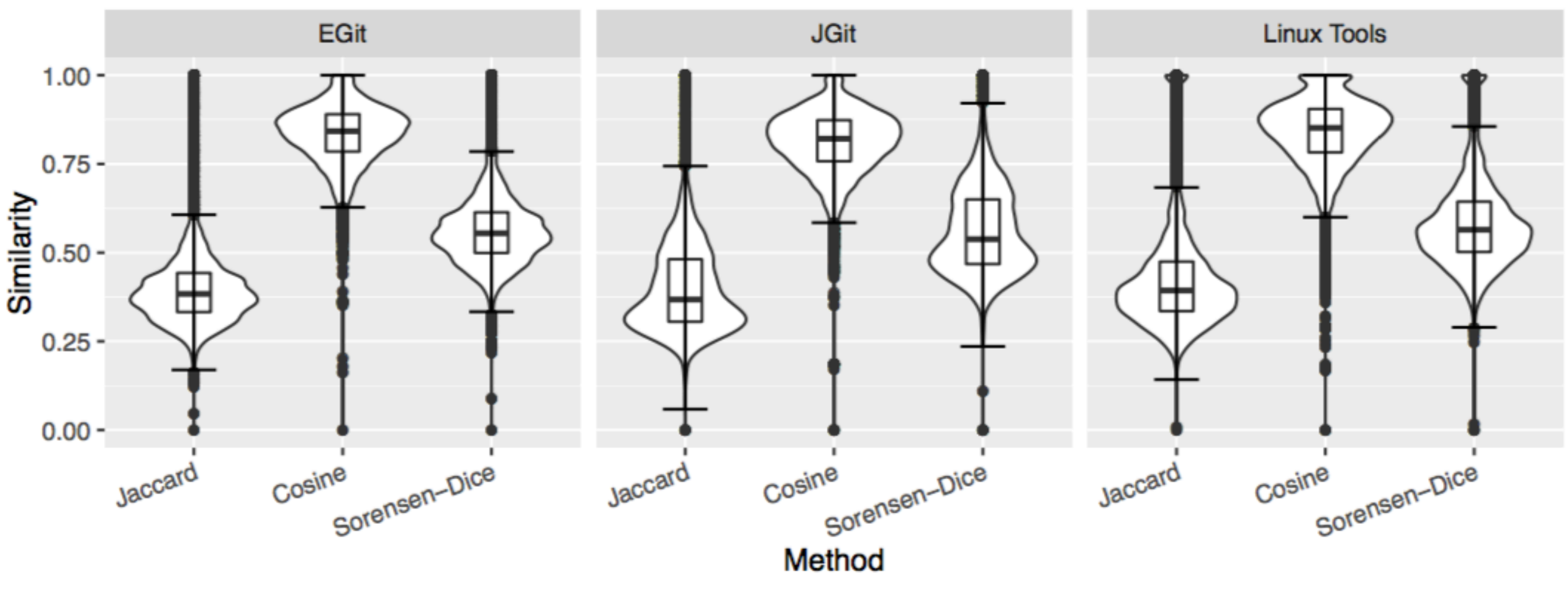
RQ3. Do the recommended patches provide useful information during code review? (objective evaluation)

Similarity Measurements Comparison

*** Sorensen-Dice recommends subset of Jaccard**
*** Cosine doesn't have recommendation**

TABLE 6
Statistics of patches with at least one recommended related patch

	EGit	JGit	Linux Tools
Jaccard	446	163	700
Sørensen-Dice	361	95	621
Cosine	0	0	0



Research Questions

RQ1. How often are related patches recommended?

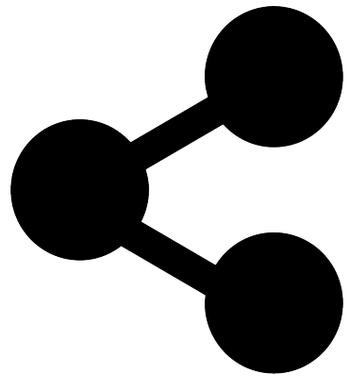
RQ2. How robust is the result compared to other similarity measurements?

RQ3. Do the recommended patches provide useful information during code review? (objective evaluation)

Manual Inspection

149 recommendations with 675 patches

Two inspector & cross check



Direct evidence / Relationship

Direct Evidence

Developers manually keep the related code reviews

Our technique find the manual code review pairs

TABLE 7
A statistical summary of the three evidence categories and the different types of evidence under each category

	EGit				JGit				Linux Tools			
	Review pairs		# Patches		Review pairs		# Patches		Review pairs		# Patches	
Total Sample Size	53		204		46		302		50		169	
Bug id	22.64%	(12)	29.9%	(61)	19.57%	(9)	17.22%	(52)	10%	(5)	3.55%	(6)
Topic	11.32%	(6)	7.84%	(16)	-	-	-	-	-	-	-	-
Related change	15.09%	(8)	17.65%	(36)	10.87%	(5)	9.60%	(29)	10%	(5)	9.47%	(16)
Recommended Change-Id	7.55%	(4)	8.82%	(18)	10.87%	(5)	14.24%	(43)	14%	(7)	5.33%	(9)
Same Change-Id	15.09%	(8)	18.63%	(38)	36.96%	(17)	33.11%	(100)	16%	(8)	23.08%	(39)
Recommended Change-Id in comments	11.32%	(6)	18.14%	(37)	6.52%	(3)	22.85%	(69)	8%	(4)	7.1%	(12)
Description in commit message	1.89%	(1)	0.98%	(2)	4.35%	(2)	3.31%	(10)	8%	(4)	6.51%	(11)
Todo in code	5.66%	(3)	2.45%	(5)	-	-	-	-	-	-	-	-
Total Evidence	67.92%	(36)	82.35%	(168)	76.09%	(35)	85.43%	(258)	58%	(29)	52.07%	(88)

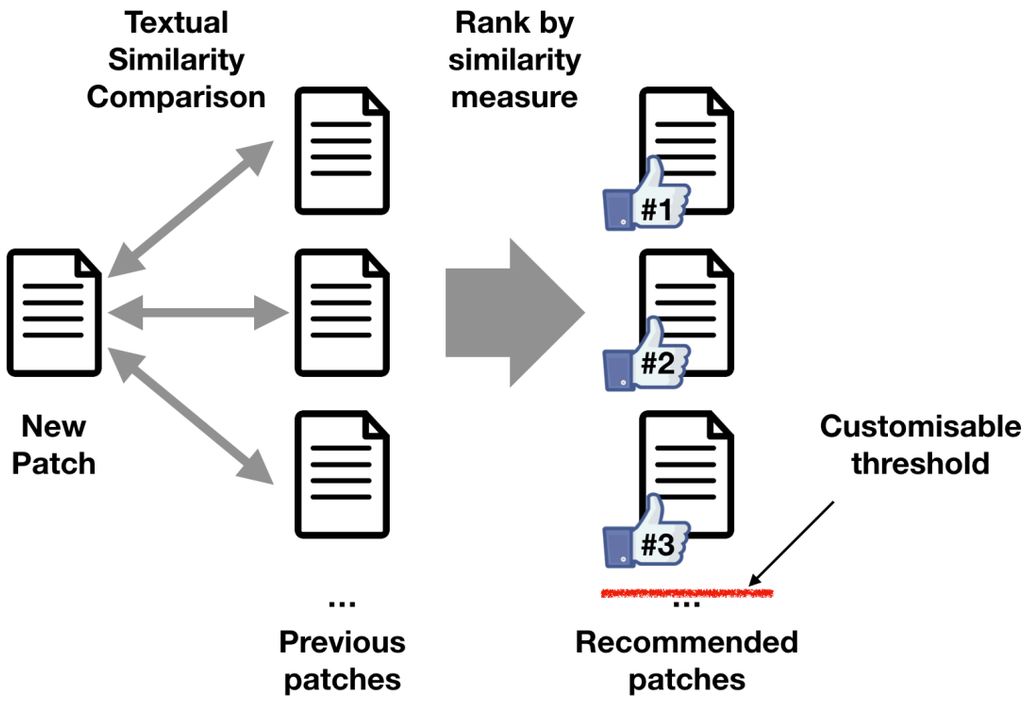
Manual Investigation

Investigate the content of patch manually

TABLE 8
A list of different relationship categories for the suggested patch sets

Category	EGit		JGit		Linux Tools		Total	
	Same files	Different files	Same files	Different files	Same files	Different files	Same files	Different files
Change of similar code	-	1.0% (2)	3.0% (9)	1.3% (4)	5.3% (9)	5.3% (9)	2.7% (18)	2.2% (15)
Related change	2.9% (6)	14.7% (30)	3.0% (9)	6.6% (20)	10.1% (17)	-	4.7% (32)	7.4% (50)
Revert change	2.9% (6)	-	5.3% (16)	-	4.1% (7)	-	4.3% (29)	-
Fix the same bug	3.9% (8)	-	11.3% (34)	-	-	-	6.2% (42)	-
Same change submitted twice (with minor change)								
- Same branch	26.0% (53)	-	24.5% (74)	-	26.0% (44)	-	25.3% (171)	-
- Different branch	31.9% (65)	-	33.1% (100)	-	25.4% (43)	-	30.8% (208)	-
Change in similar location	3.4% (7)	-	3.6% (11)	-	-	-	2.7% (18)	-
Refactoring	5.9% (12)	-	7.6% (23)	-	21.9% (37)	-	10.7% (72)	-
Fix newly introduced bug	1.0% (2)	-	0.3% (1)	-	-	-	0.4% (3)	
Update meta data	3.9% (8)	0.5% (1)	0.3% (1)	-	1.7% (3)	-	1.8% (12)	0.1% (1)
Not related	-	**2.0% (4)**	-	-	-	-	**0.6% (4)**	-

Related Code Review Recommendation

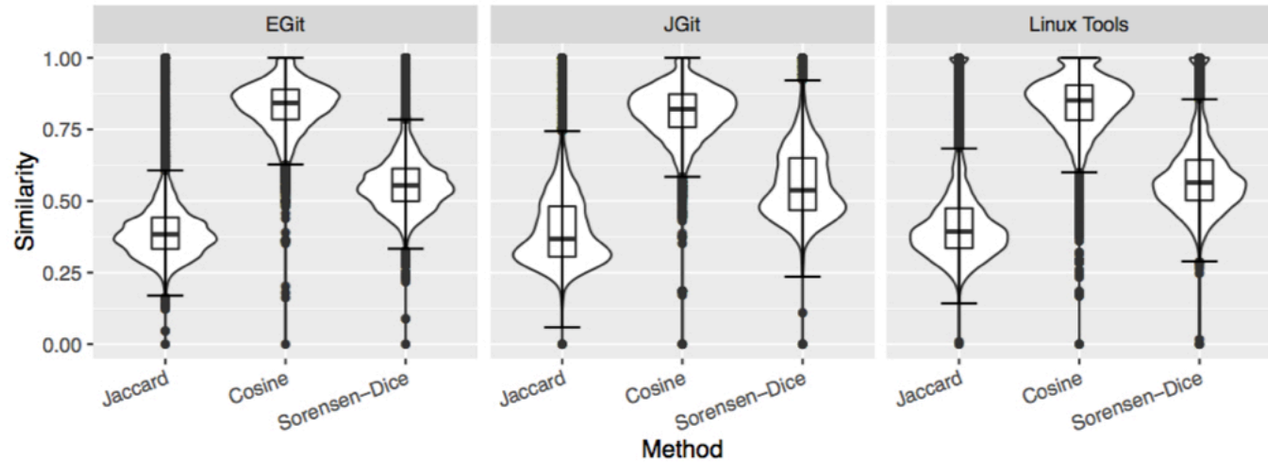


Similarity Measurements Comparison

* Sorensen-Dice recommends subset of Jaccard
* Cosine doesn't have recommendation

TABLE 6
Statistics of patches with at least one recommended related patch

	EGit	JGit	Linux Tools
Jaccard	446	163	700
Sørensen-Dice	361	95	621
Cosine	0	0	0



The number of resubmitted identical patches and reasons

TABLE 4
Number of resubmitted identical patches

Project	EGit	JGit	Linux Tools
Patches	**0.7%** (49/7,050)	**0.4%** (24/6,457)	**2.7%** (252/9,232)
Reviews	**0.7%** (34/4,752)	**0.4%** (19/4,408)	**4.8%** (217/4,546)

TABLE 5
Reasons for resubmitting identical patches

Target branch	Category	EGit Reviews		EGit Patches		JGit Reviews		JGit Patches		Linux Tools Reviews		Linux Tools Patches	
Different branch	Cherry pick	76.47%	(26)	77.55%	(38)	78.95%	(15)	70.83%	(17)	88.48%	(192)	86.45%	(217)
	Merge	14.71%	(5)	16.33%	(8)	5.26%	(1)	12.50%	(3)	8.76%	(19)	11.16%	(28)
Same branch	Working on same base	5.88%	(2)	4.08%	(2)	5.26%	(1)	4.17%	(1)	2.30%	(5)	1.99%	(5)
	Mistake	2.94%	(1)	2.04%	(1)	-	-	-	-	0.46%	(1)	0.40%	(1)
	Take over	-	-	-	-	10.53%	(2)	12.50%	(3)	-	-	-	-

Manual Inspection

149 recommendations with 675 patches

Two inspector & cross check

Direct evidence / Relationship