

進捗報告

知りたいこと

- ・ **修正にかかる時間を短縮できる要因**

- ・ 仮説

親密度が高いと修正までの時間が短い

親密度=関わっている回数が多い

→ 組み合わせ取得

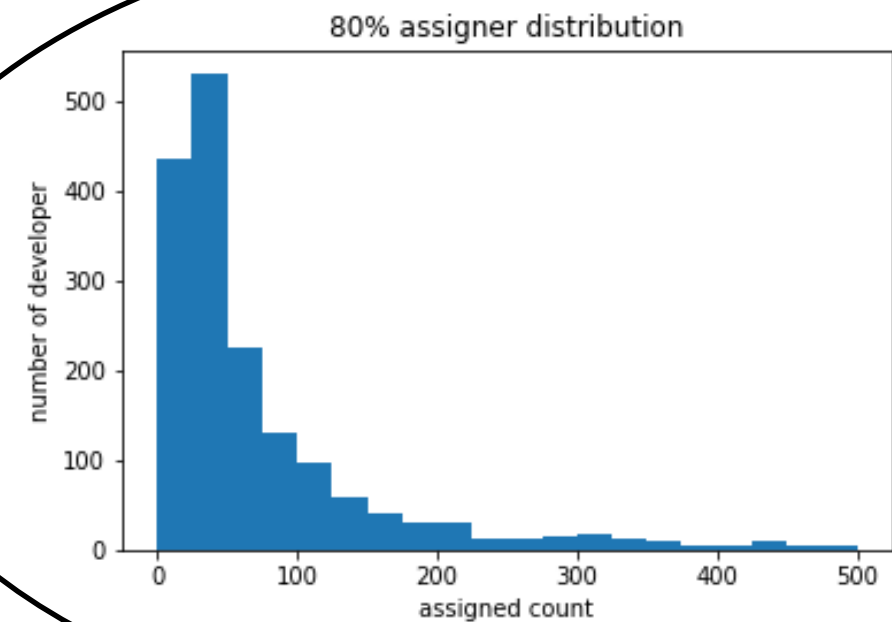
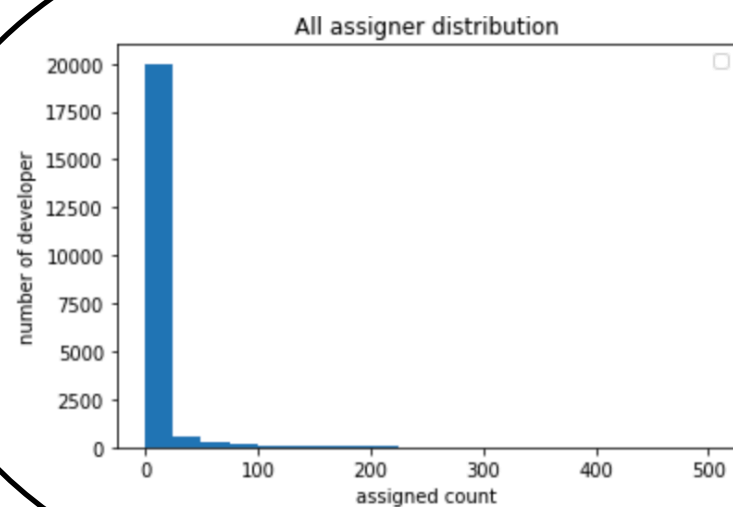
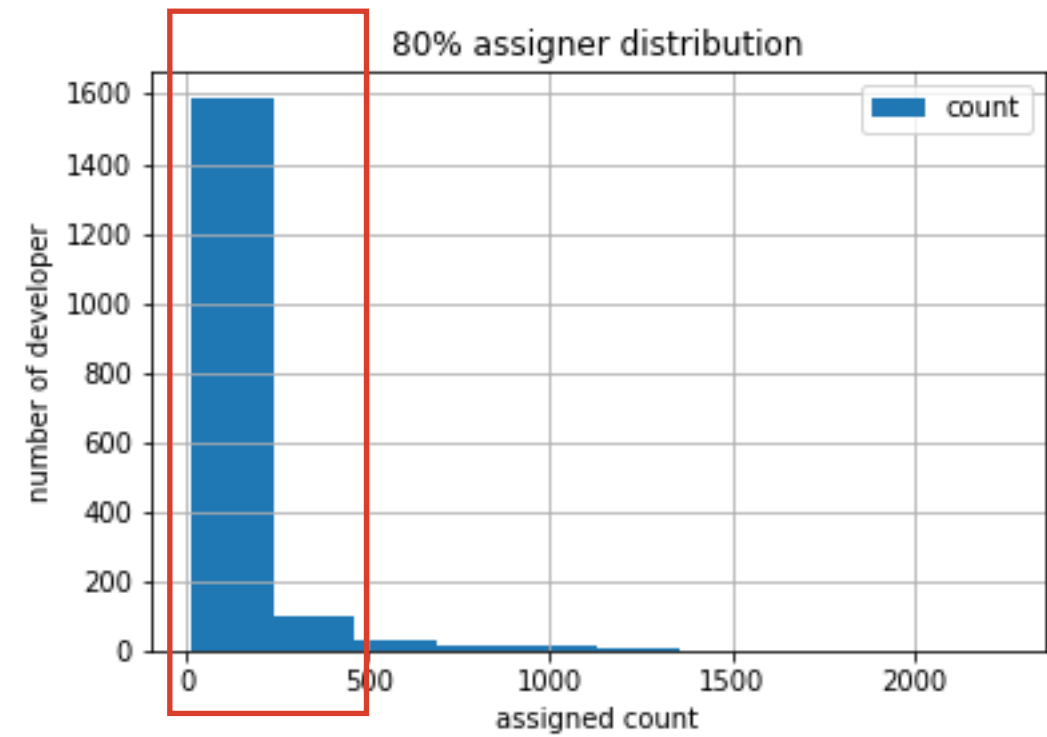
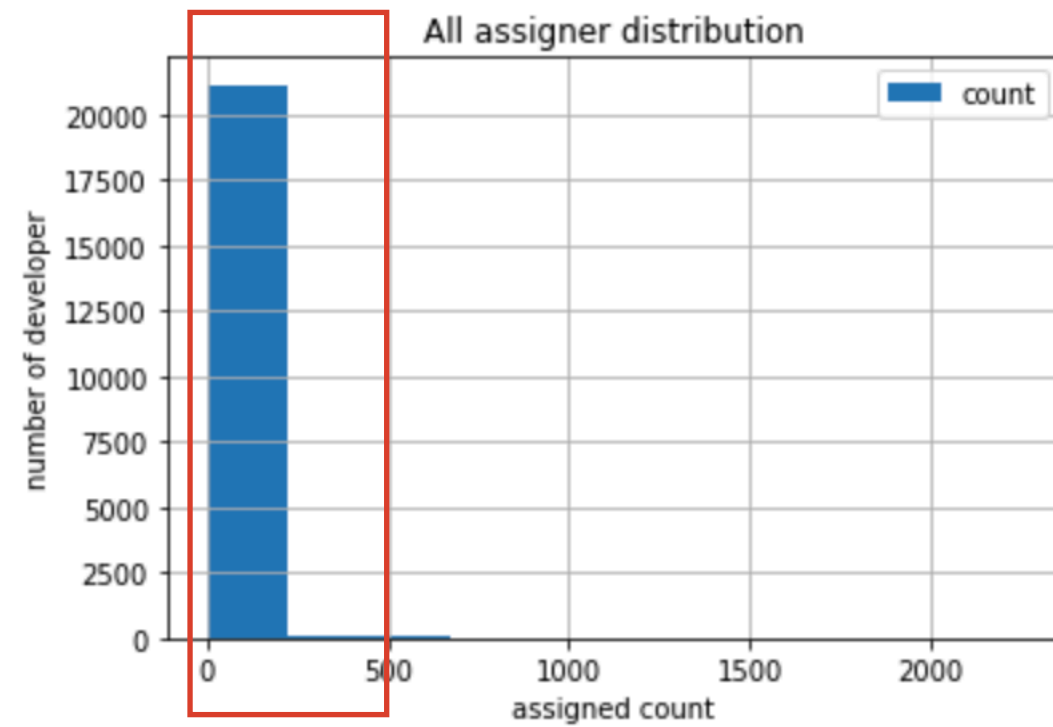
割り当て者の分布

- eclipseプロジェクトにおけるバグ管理において
割り当て者総数：21321人,
割り当て数：234854件
- 割り当て数の80%を担当している割り当て者：1760人

前処理

- 欠損値の削除
- 組織を削除
(nameにinbox,eclipse,platform,php,kdeが含まれている)
- unknown user の削除
- assigner と fixer が同じ人の場合を削除

割り当て者の分布



上位割り当て者とその内訳(一部)

	name	count
0	darin swanson	2299
1	dani megert	2232
2	markus keller	2128
3	daniel megert	1887
4	nick edgar	1693
5	snorthov	1633
6	david williams	1425
7	dirk baeumer	1423
8	tod creasey	1368
9	john arthorne	1291
10	stepper	1275
11	philippe mulet	1200
12	xiaodan wang	1159
13	jared burns	1132
14	liwen chen	1087
15	pwebster	1061
16	erich gamma	1033
17	caniszczyk	1026
18	olivier thomann	1016
19	martin aeschlimann	1013
20	curtis windatt public	970

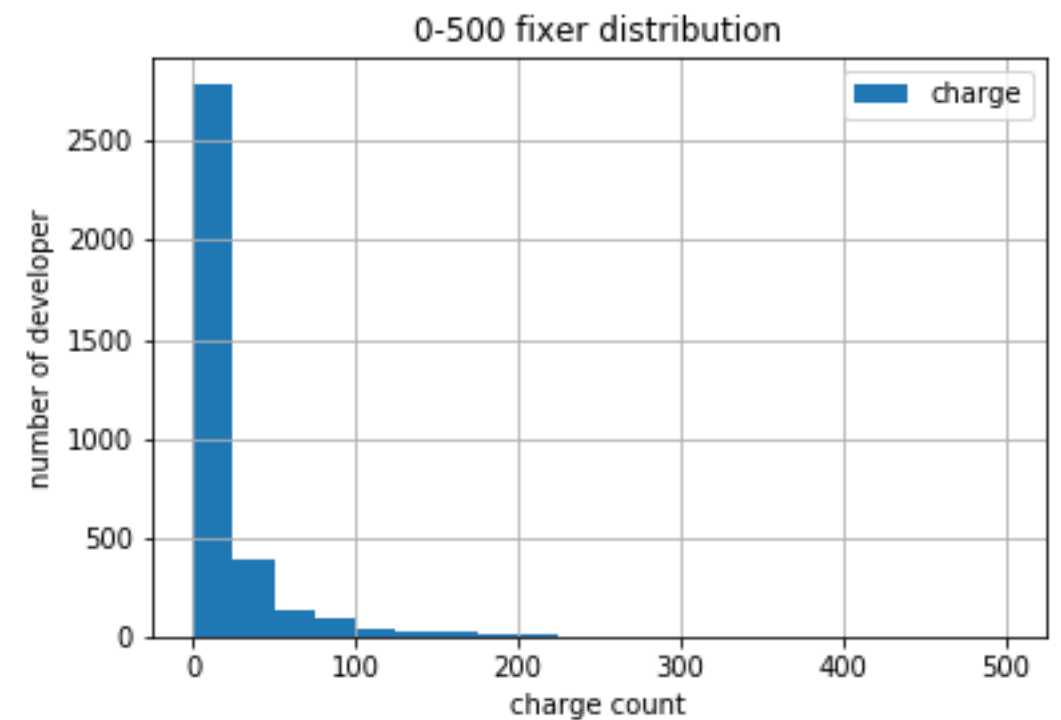
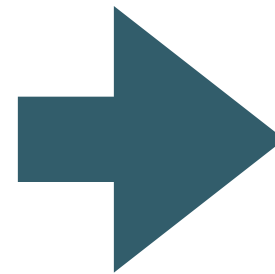
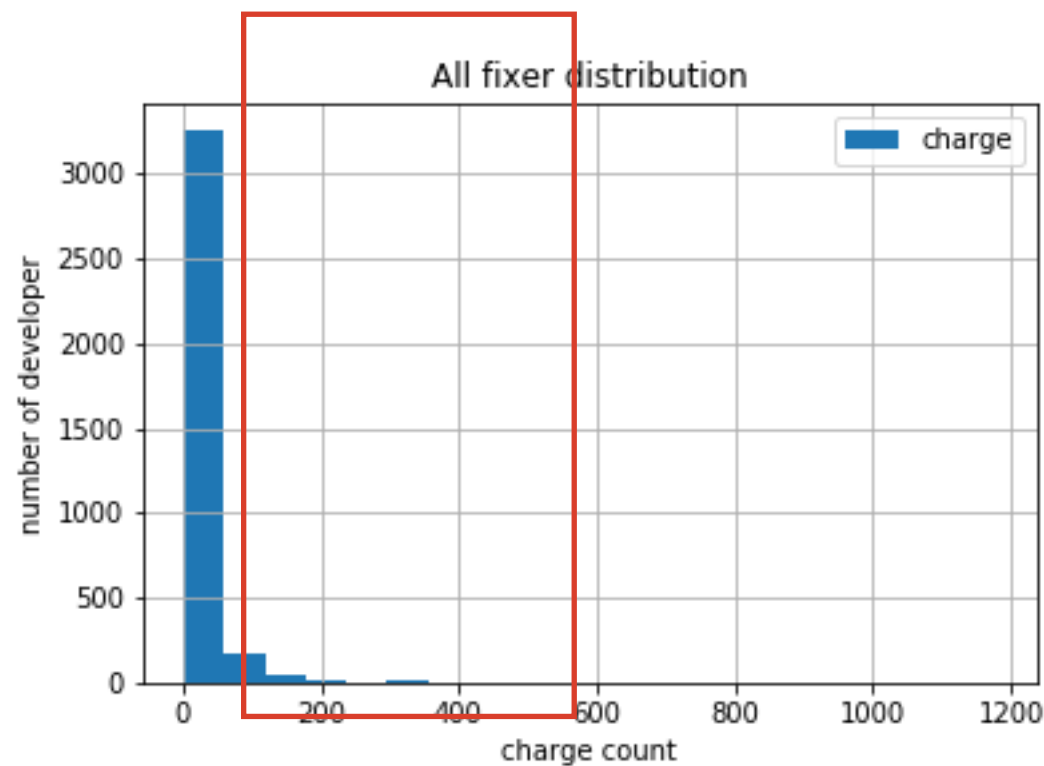
⋮

	assigner	fixer	charge	total
0	darin swanson	darin wright	771	2299
1	darin swanson	jared burns	529	2299
2	darin swanson	kevin barnes	318	2299
3	darin swanson	luc bourlier	246	2299
4	darin swanson	joe szurszewski	112	2299
5	darin swanson	michael reddie	51	2299
6	darin swanson	erich gamma	22	2299
7	darin swanson	tod creasey	21	2299
8	darin swanson	nick edgar	19	2299
9	darin swanson	john mason p shackelford	14	2299
10	dani megert	erich gamma	134	2232
11	dani megert	markus keller	129	2232
12	dani megert	tod creasey	101	2232
13	dani megert	martin aeschlimann	101	2232
14	dani megert	benno baumgartner	79	2232
15	dani megert	paul webster	63	2232
16	dani megert	tom hofmann	62	2232
17	dani megert	felipe heidrich	42	2232

⋮
⋮
⋮
⋮

修正者の分布

- 縦軸が修正者の数, 横軸が修正者が担当した数
- ほとんどが200以下 → 見たいデータではないので削除



担当数200未満を削除

- 3543件 → 42件
- assignerとfixerが同一人物である場合を手作業で削除
42件 → 17件

	assigner		fixer	charge	total
0	darin swanson	darin wright		771	2299
1	darin swanson	jared burns		529	2299
2	darin swanson	kevin barnes		318	2299
3	darin swanson	luc bourlier		246	2299
30	daniel megert	dani megert		1102	1887
40	nick edgar	tod creasey		222	1693
50	snorthov	steve northover		799	1633
51	snorthov	silenio quarti		213	1633
70	dirk baeumer	martin aeschlimann		344	1423
100	stepper	eike stepper		1179	1275
110	philippe mulet	philipe mulet		483	1200
111	philippe mulet	jerome lanneluc		225	1200
130	jared burns	darin wright		397	1132



	assigner		fixer	charge	total
0	darin swanson	darin wright		771	2299
1	darin swanson	jared burns		529	2299
2	darin swanson	kevin barnes		318	2299
3	darin swanson	luc bourlier		246	2299
40	nick edgar	tod creasey		222	1693
51	snorthov	silenio quarti		213	1633
70	dirk baeumer	martin aeschlimann		344	1423
111	philippe mulet	jerome lanneluc		225	1200
130	jared burns	darin wright		397	1132
131	jared burns	darin swanson		276	1132
210	michael reddie	darin wright		353	962
211	michael reddie	curtis windatt		203	962
220	martin oberhuber	david mcknight		250	902
280	mik kersten	robert elves		206	801
456	karen butzke	neil hauge		206	589
486	cocoaKevin	darin wright		255	575
834	give a damus	christian w damus		325	359

次にすること

- 算出したペアの修正時間を箱ひげ図にプロット
→ 短いかどうかを確認
- 短い場合, 詳細な要因を知るため会話データ等进行分析
長い場合は仮説が間違っていたということ。。

疑問

- 評価する際に比較する？
 - その際の比較対象は？修正時間が遅い組み合わせ？
- どのような観点で分析するのか
 - 会話内容(感情分析, 具体例の有無)
 - 参入時期
 - 母国語が同じ
 - 他のプロジェクトで関わったことがあるか

気づいたこと

- reporterのことを考えてなかったが3つの役割を異なる開発者で行う場合をみたいので考慮すべき？

$R=T \neq F$, $T \neq F=R$ の可能性があるので