

OSS 開発における不具合修正プロセスの改善 に向けたモジュールオーナー候補者推薦

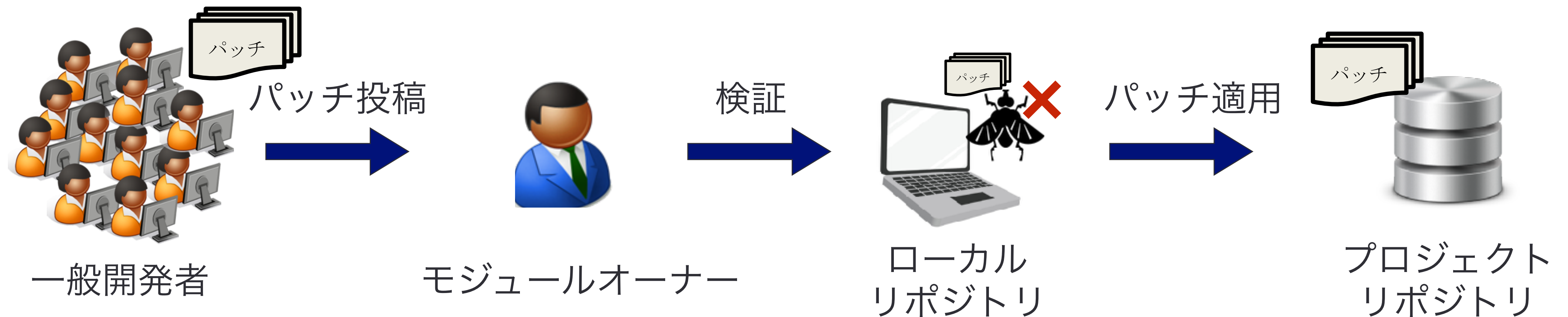
Module Owner Candidate Recommendation
toward Improving Bug-Fixing Process in OSS Development

和歌山大学*
日本学術振興会特別研究員 (DC)†

© 柏 祐太郎*†, 山谷 陽亮*, 大平 雅雄*

背景

- 大規模OSSプロジェクトには日々、
数百件を超える不具合が報告されている[1]
 - プロジェクトに参加する開発者によって、パッチを作成・投稿
 - 作成されたパッチは**モジュールオーナー**によって検証・適用

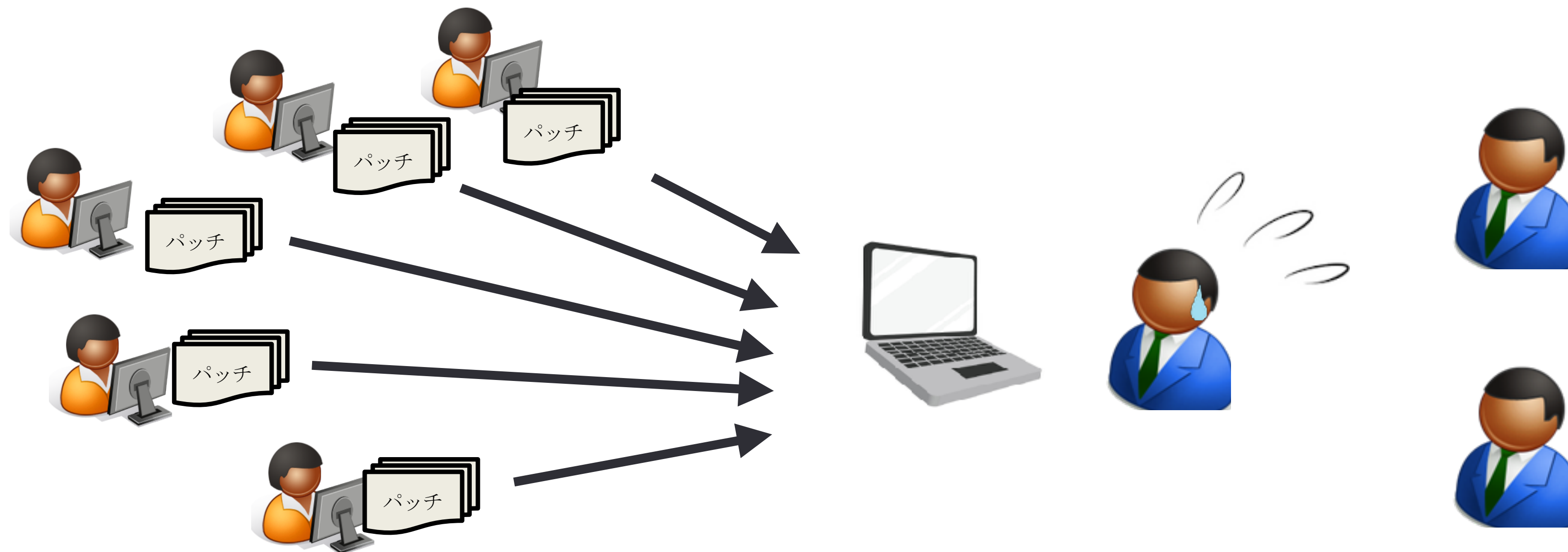


[1] John Anvik, Lyndon Hiew, and Gail C. Murphy. Who should x this bug? , In Proc. of ICSE'06, pp. 361--370, 2006.

モジュールオーナーへの過度な負担

3

- コミット権限を持たない開発者に比べ少人数で構成
 - 一般開発者100人に対してモジュールオーナーは1人程度 (Firefox)
 - パッチの検証は日々実施されるが、**検証待ちパッチが多数存在**^[2]



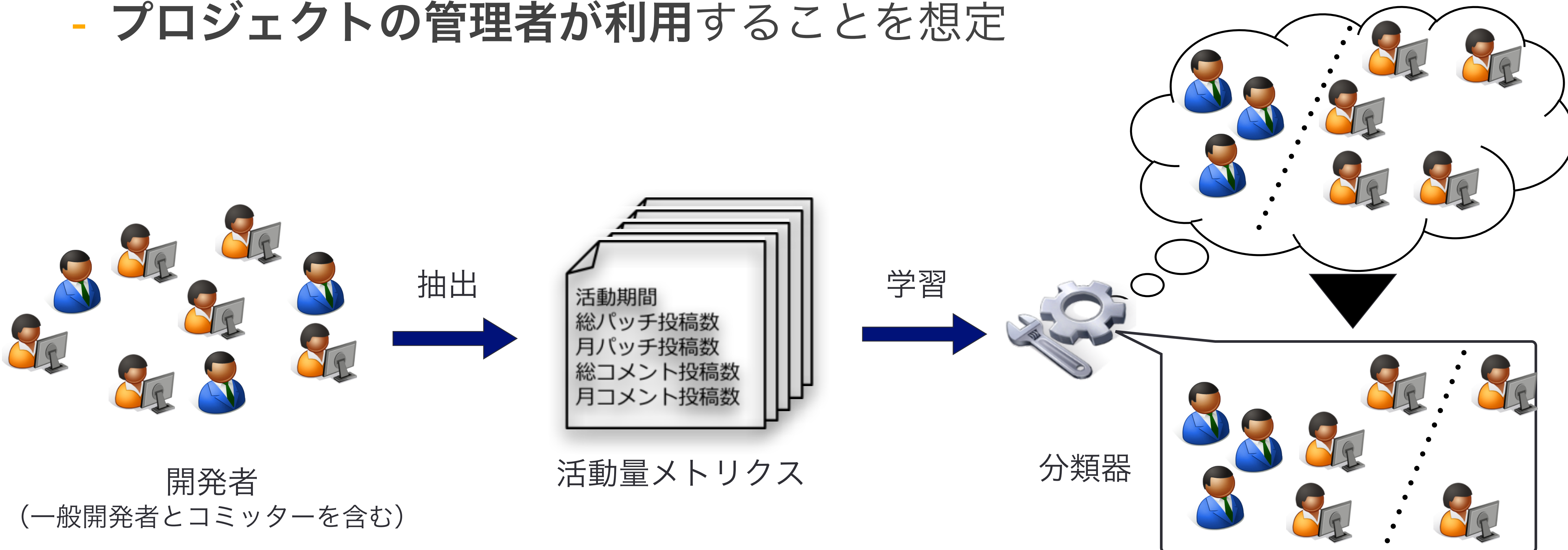
モジュールオーナーへの昇格

- 開発者のこれまでのプロジェクトでの活動を基に,
既存のモジュールオーナーが慎重に判断
 - ほとんどのプロジェクトでは,
モジュールオーナーへの昇格に具体的な指標はない
- 多数の開発者がプロジェクトに参加するも,
ほとんどがモジュールオーナーにならずに去ってしまう
 - **開発者は平均1年でプロジェクトを去ってしまう**
 - 一方, **プロジェクトとしては1年以上の貢献が必要な場合が多い**

どれだけ頑張ればコミット権限が得られるかを示すことで,
将来の活躍が期待できる開発者のプロジェクト離れを防ぐことを目指す

先行研究[3]

- 活動量メトリクスを用いたコミッター候補者推薦
 - コミッターとして活躍が期待できる開発者の早期発見を目指す
 - プロジェクトの管理者が利用することを想定

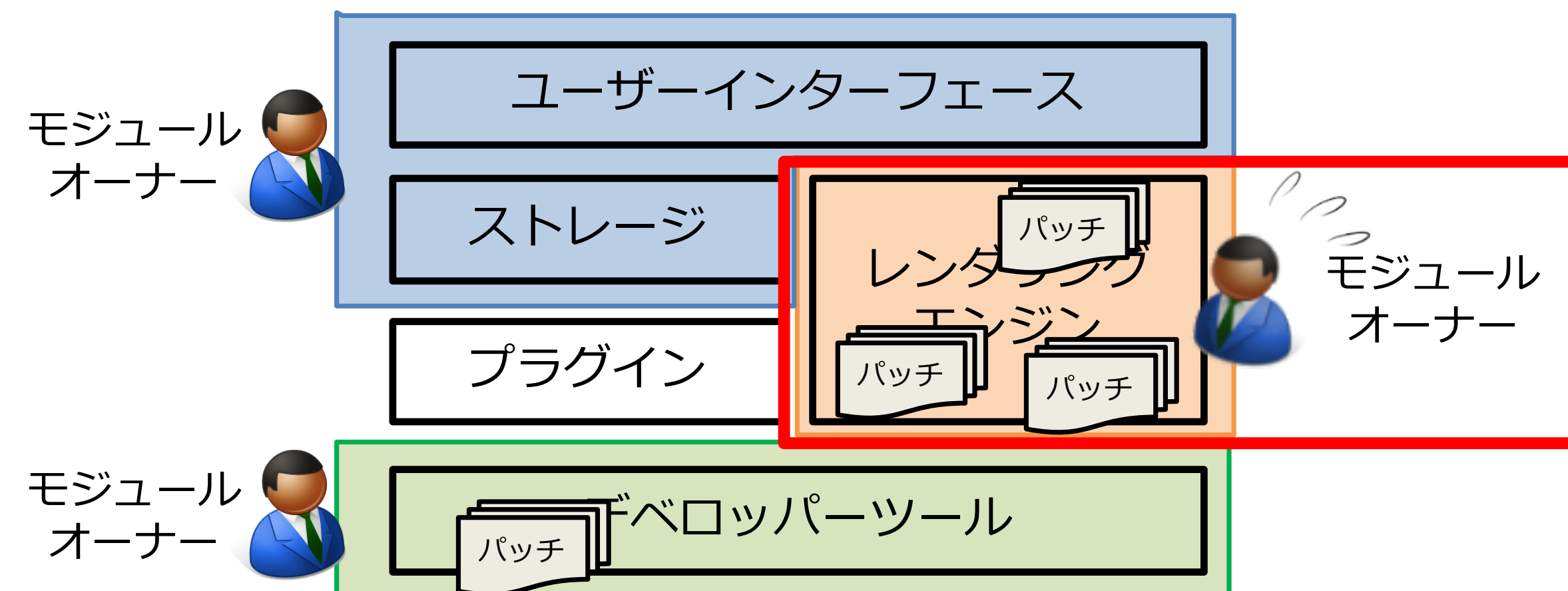


[3] Akinori Ihara, Yasutaka Kamei, Masao Ohira, Ahmed E. Hassan, Naoyasu Ubayashi, and Kenichi Matsumoto. Early identification of future committers in open source software projects.

In Proceedings of the 14th International Conference on Quality Software (QSIC'14), pp. 47--56, 2014.

本研究の狙い

- コミット権限取得を目指す開発者の支援
 - 自分がコミット権限を得られるかを予測
 - コミット権限の取得に必要な活動量を提示
- モジュールごとに必要とされる専門性や活動量を考慮
 - 既存手法では、どのモジュールの担当かは考慮されていない
 - 慢性的に人員不足のモジュールや新モジュールは昇格しやすいと想定



Research Questions

RQ1. MOの昇格人数は、モジュールごとで違いはあるのか？

RQ2. MOの昇格に必要とされる活動量は、
モジュールで違いはあるか？

RQ3. モジュールオーナーを予測できるか

- 対象プロジェクト

プロジェクト			
	Mozilla Firefox	Eclipse Platform	Webkit
期間	1999/07～2013/01	2001/10～2012/10	2000/12～2014/04
ドメイン	WEBブラウザ	総合開発環境 (IDE)	レンダリングエンジン
モジュール数	5	3	7
モジュールオーナー数	530	195	363
開発者数	7,6931	1,3297	8,017

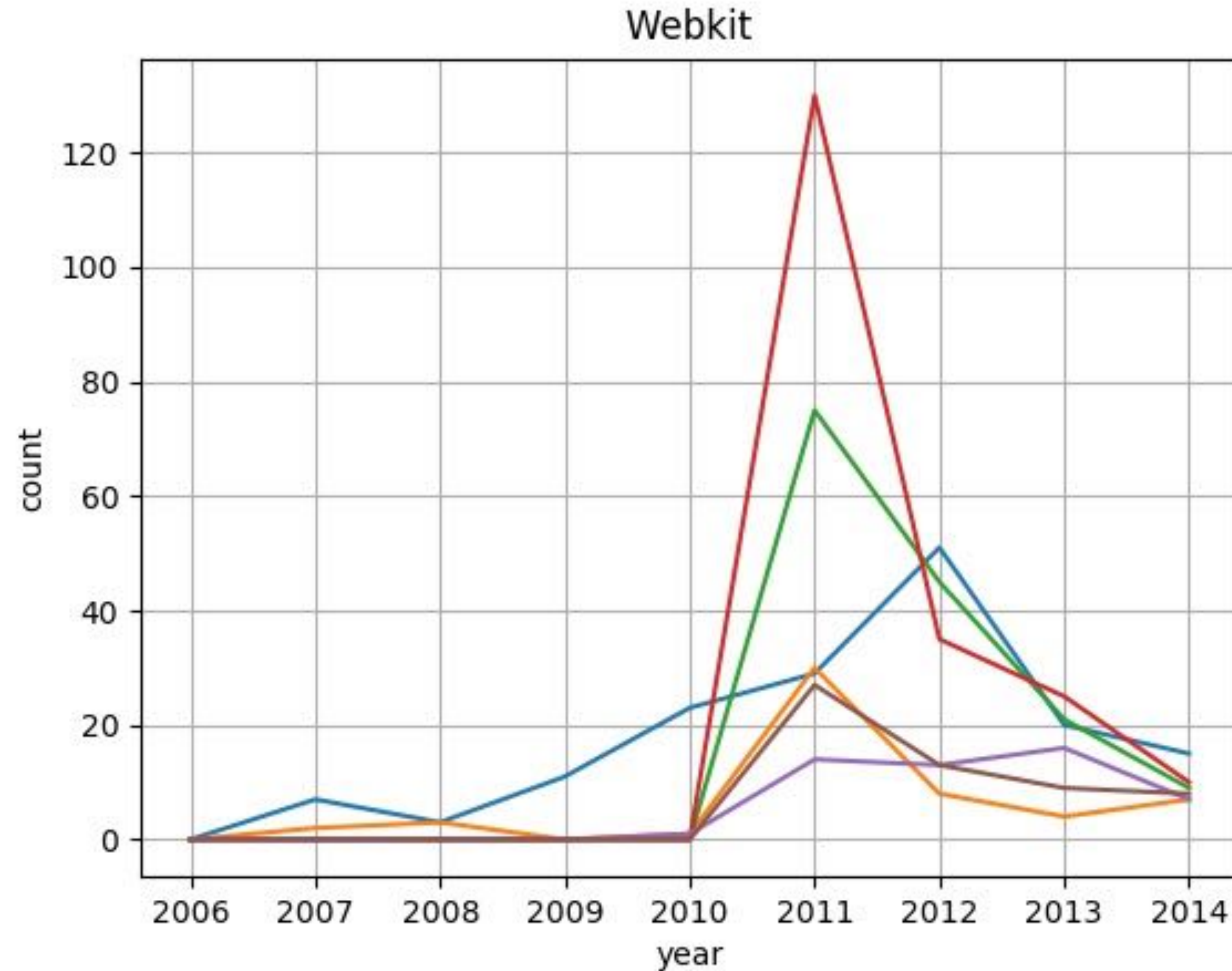
※モジュール数は、不具合票に記載されるモジュール名をベースにした。
ただし、モジュールの粒度が異なるため、プロジェクトのwikiページなどを参考に適時統合

RQ 1:モジュールごとの昇格人数の違い

- 各モジュールごとの昇格人数を集計

モジュール名	総昇格数	平均	最小値	中央値	最大値
Build	159	17.67	0	15.00	51
JavaScriptCore	55	6.11	0	3.00	30
Layout	150	16.67	0	0.00	75
WebCore	200	22.22	0	0.00	130
WebKit2	51	5.67	0	1.00	16
Inspector	57	6.33	0	0.00	27

RQ 1:モジュールごとの昇格人数の違い



- 先行研究で用いられている活動量メトリクスを
モジュールごとに計測
 - モジュールオーナーとなる開発者は、昇格するまでのデータを利用

メトリクス名	内容
活動期間	初めて活動した日から最後に活動した日までの期間の日数
総パッチ投稿数	各開発者が投稿したパッチの総数
月平均パッチ投稿数	各開発者が1か月にパッチを投稿した回数の中央値
総コメント投稿数	各開発者が投稿したコメントの総数
月平均コメント投稿数	各開発者が1か月にコメントを投稿した回数の中央値

※計測したメトリクスは分布に偏りが存在するために、対数変換を実施

RQ2:モジュールごとの活動量の違い

12

- モジュールオーナーとモジュールオーナーでない開発者の活動量を比較
 - 有意差があるメトリクスを太字で表示
 - » * はp値が0.05以下
 - » ** はp値が0.01以下

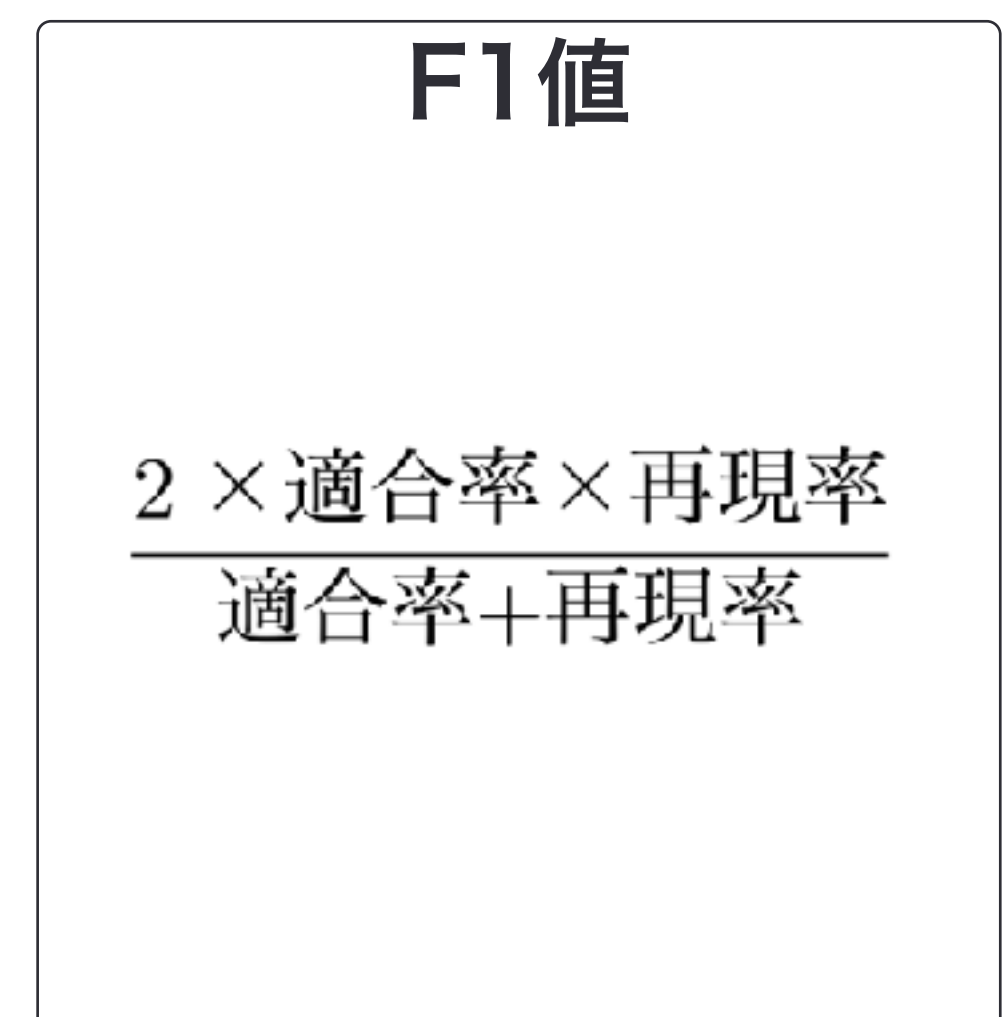
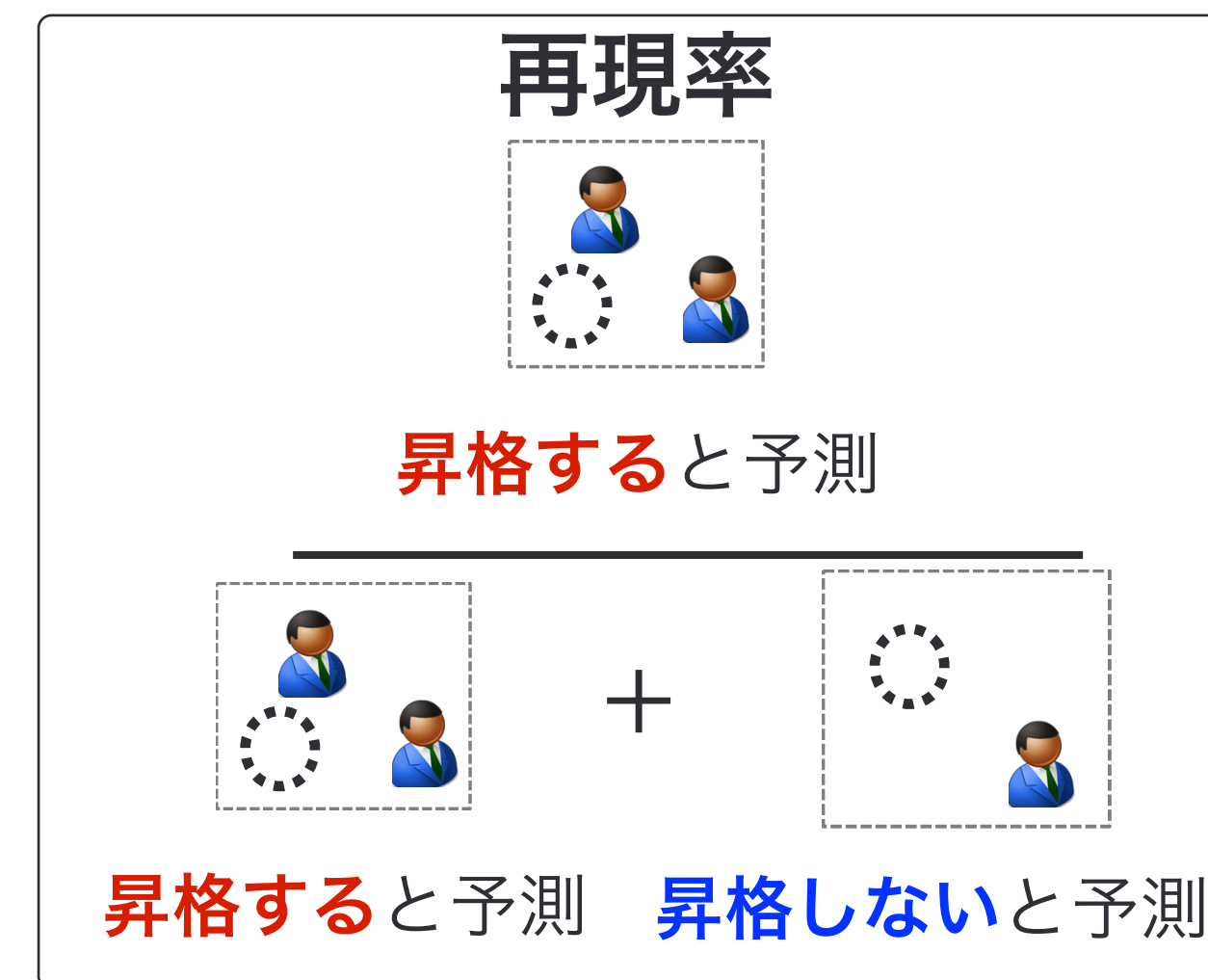
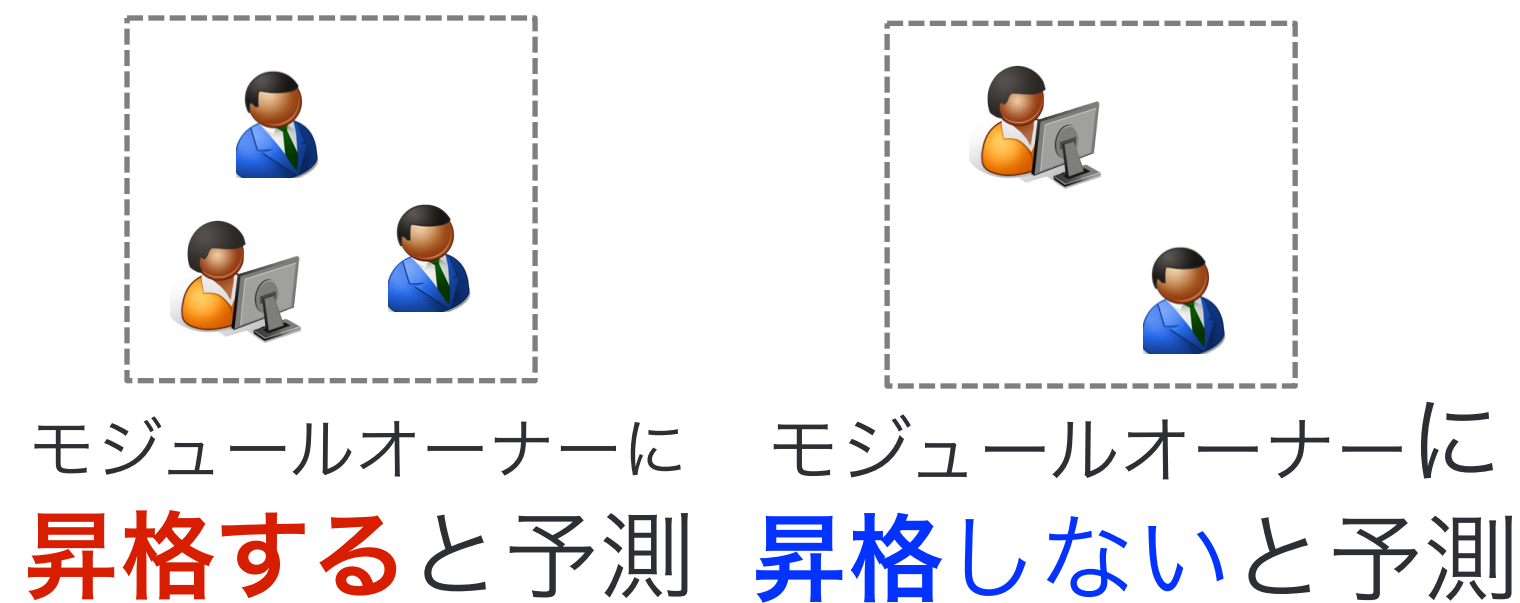
モジュール名	活動期間	総コミット	月コミット	総コメント	月コメント
Build	5.64*	2.20	0.00	3.04	0.00
JavaScriptCore	6.78	1.95	0.00	3.22**	0.00**
Layout	6.33	2.94**	0.00**	4.14**	0.69**
WebCore	6.34*	2.74**	0.00**	3.88**	0.00**
WebKit2	5.66	2.30	0.00	3.30	0.41
Inspector	6.54	1.79	0.00	3.56	0.00

目的

- モジュールごとに計測した活動量メトリクスを用いて、モジュールオーナー候補者を予測することができるかを確かめること

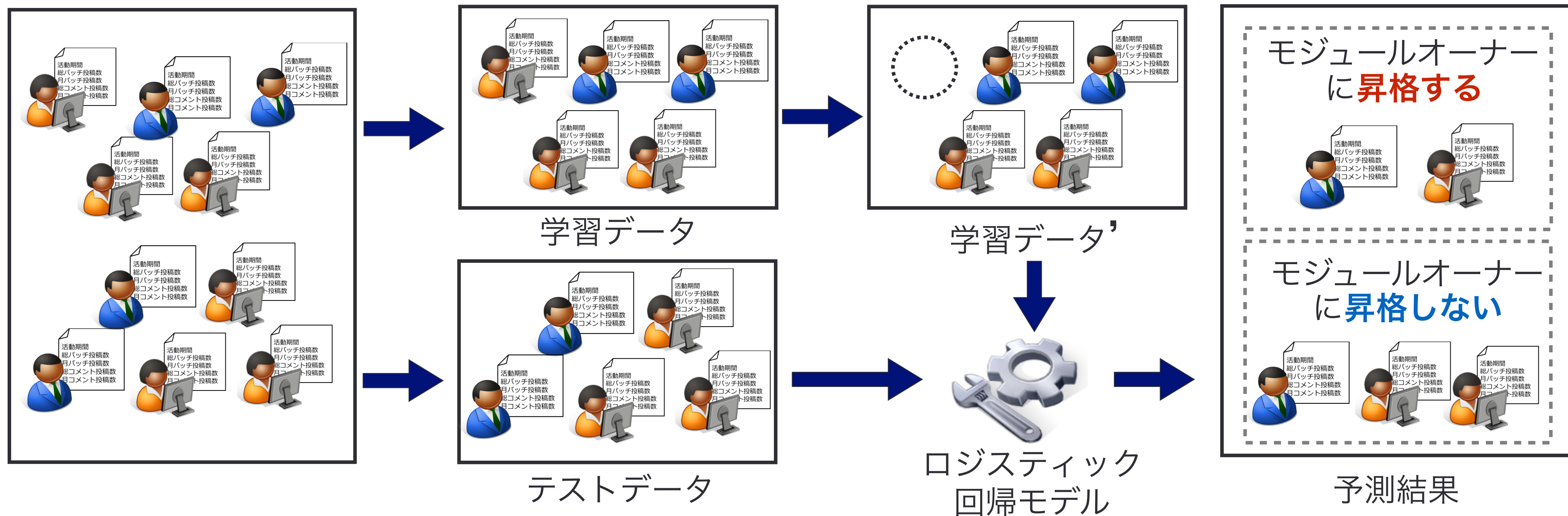
評価指標

- 適合率と再現率, F1値で比較



実験手順

- 以下の処理を1000回繰り返して精度を測定



① 学習データとテストデータにデータをランダムに2分割

② 過学習を避けるために一般開発者をモジュールオーナーの数と同数に調整

③ テストデータを入力し予測精度を算出

RQ3:モジュールオーナーの予測

15

- Firefoxプロジェクトのモジュールごとの予測精度

モジュール名	適合率	再現率	F1値
Developer Tools	0.80	0.77	0.80
App	0.57	0.56	0.55
Preferences	0.55	0.61	0.59
Bar	0.50	0.75	0.59
Bookmark	0.48	0.56	0.51
General	0.46	0.59	0.51
平均	0.56	0.55	0.59

モジュール別に求めた活動量メトリクスを用いて
2人に1人の精度でモジュールオーナーを予測可能

考察：予測に失敗した開発者

16

- 予測精度の向上のため、
正しく予測できなかった開発者を目視で確認

	モジュールオーナー に昇格した	モジュールオーナー に昇格しなかった
モジュールオーナー に昇格すると予測	TRUE POSITIVE	FALSE POSITIVE
モジュールオーナー に昇格しないと予測	FALSE NEGATIVE	TRUE NEGATIVE

- モジュール内での活動量が比較的**低い傾向**
- 他のモジュールでも活動
- ⇒他モジュールとの調整役を担っている可能性
 - » 依存関係のモジュールのコミット回数などを考慮

- モジュールオーナーと同様に、モジュール内での活動量が**多い傾向**
- ⇒プロジェクト内での**評価が低い可能性**
 - » パッチに対する評価を考慮（感情分析などを実施）

まとめと今後の課題

- モジュールごとに計測した活動量メトリクスを用いてモジュールオーナー候補者を推薦するモデルを構築
 - 3種類のOSSプロジェクトを対象に実験を行い、適合率が0.5前後で予測できることがわかった

今後の課題

- 活動量だけでなく、プロジェクトでの評価も考慮
- 依存関係のあるモジュールにおける活動を考慮