

OSSにおける開発者のHigh Impact Bugに対する認識調査 (2015-10-27)

B4 松野祐介

背景

- 既存の研究では、ユーザや開発者に大きな影響を与えるバグを対象とした分析が行われている
- High Impact Bug
 - Surprise Bug , Dormant Bug , Blocking Bug , Security Bug , Performance Bug , Breakage Bug
- しかし、いずれも研究者によって定義されたもの
 - 現場で扱われているものとは異なる可能性がある
- 開発者が影響力が高いと考えるものをサポートする必要がある

目的

- 影響力の高いバグとはこういったバグであるかを開発者の意見を元に調査する
 - どのような役割の開発者が意見を述べているか
 - 何に対して影響力が高いものであるか
 - なぜ影響力が高いのか
- 調査することで、開発者の意見を加味したHIBの定義、その予防・対策などに寄与

アプローチ(1/2)

- GitHubに登録されている開発者に対してメールを送ってアンケートを行う
 - 約1000万人のユーザが登録
- その結果を集計し、調査する

アプローチ(2/2)

- より活動的な開発者にアンケートを行うために、github apiで開発者のcontributionを見たい
 - 開発者の属性にcontributionは存在しない
 - 活発なリポジトリで活動している開発者の意見が欲しい
- 全リポジトリのcontributorを見る
- リポジトリのIDは約3500万件存在

GitHub API

- URL: <https://api.github.com/> にアクセスすることでJSON形式で記述されているGitHub上のデータを閲覧できる

```
{
  "current_user_url": "https://api.github.com/user",
  "current_user_authorizations_html_url": "https://github.com/settings/connections/applications/{client_id}",
  "authorizations_url": "https://api.github.com/authorizations",
  "code_search_url": "https://api.github.com/search/code?q={query}{&page,per_page,sort,order}",
  "emails_url": "https://api.github.com/user/emails",
  "emojis_url": "https://api.github.com/emojis",
  "events_url": "https://api.github.com/events",
  "feeds_url": "https://api.github.com/feeds",
  "following_url": "https://api.github.com/user/following/{target}",
  "gists_url": "https://api.github.com/gists/{gist_id}",
  "hub_url": "https://api.github.com/hub",
  "issue_search_url": "https://api.github.com/search/issues?q={query}{&page,per_page,sort,order}",
  "issues_url": "https://api.github.com/issues",
  "keys_url": "https://api.github.com/user/keys",
  "notifications_url": "https://api.github.com/notifications",
  "organization_repositories_url": "https://api.github.com/orgs/{org}/repos?type,page,per_page,sort",
  "organization_url": "https://api.github.com/orgs/{org}",
  "public_gists_url": "https://api.github.com/gists/public",
  "rate_limit_url": "https://api.github.com/rate_limit",
  "repository_url": "https://api.github.com/repos/{owner}/{repo}",
  "repository_search_url": "https://api.github.com/search/repositories?q={query}{&page,per_page,sort,order}",
  "current_user_repositories_url": "https://api.github.com/user/repos?type,page,per_page,sort",
  "starred_url": "https://api.github.com/user/starred/{owner}/{repo}",
  "starred_gists_url": "https://api.github.com/gists/starred",
  "team_url": "https://api.github.com/teams",
  "user_url": "https://api.github.com/users/{user}",
  "user_organizations_url": "https://api.github.com/user/orgs",
  "user_repositories_url": "https://api.github.com/users/{user}/repos?type,page,per_page,sort",
  "user_search_url": "https://api.github.com/search/users?q={query}{&page,per_page,sort,order}"
}
```

<https://api.github.com/>

repositories(1/2)

- <https://api.github.com/repositories>にアクセスすることでリポジトリ一覧が表示できる
- 一度に表示される数は100が上限
 - https://api.github.com/repositories?since={repo_id} のように指定することでID以降から続きが表示される
- repo_idは約3500万件まで存在

repositories(2/2)

- https://api.github.com/repos/{full_name}/contributorsにアクセスすることでcontributorをcontributionの多い順に30名まで表示できる

```
[
  {
    "id": 1,
    "name": "grit",
    "full_name": "mojombo/grit",
    "owner": {
      "login": "mojombo",
      "id": 1,
      "avatar_url": "https://avatars.githubusercontent.com/u/1?v=3",
      "gravatar_id": "",
      "url": "https://api.github.com/users/mojombo",
      "html_url": "https://github.com/mojombo",
      "followers_url": "https://api.github.com/users/mojombo/followers",
      "following_url": "https://api.github.com/users/mojombo/following{/other_user}",
```

full_name=ユーザ名/リポジトリ名

～中略～

```
"blobs_url": "https://api.github.com/repos/mojombo/grit/git/blobs{/sha}",
"git_tags_url": "https://api.github.com/repos/mojombo/grit/git/tags{/sha}",
"git_refs_url": "https://api.github.com/repos/mojombo/grit/git/refs{/sha}",
"trees_url": "https://api.github.com/repos/mojombo/grit/git/trees{/sha}",
"statuses_url": "https://api.github.com/repos/mojombo/grit/statuses{/sha}",
"languages_url": "https://api.github.com/repos/mojombo/grit/languages",
"stargazers_url": "https://api.github.com/repos/mojombo/grit/stargazers",
"contributors_url": "https://api.github.com/repos/mojombo/grit/contributors",
"subscribers_url": "https://api.github.com/repos/mojombo/grit/subscribers",
"subscription_url": "https://api.github.com/repos/mojombo/grit/subscription",
"commits_url": "https://api.github.com/repos/mojombo/grit/commits{/sha}",
"git_commits_url": "https://api.github.com/repos/mojombo/grit/git/commits{/sha}",
"comments_url": "https://api.github.com/repos/mojombo/grit/comments{/number}",
"issue_comment_url": "https://api.github.com/repos/mojombo/grit/issues/comments{/numl",
"contents_url": "https://api.github.com/repos/mojombo/grit/contents{/path}"
```

ここにアクセス

<https://api.github.com/repositories> (一部)

contribution

- GitHub APIで表示されるcontributionは該当リポジトリに対してコミットした回数
- ソースコードを一回編集すると1カウント
- ユーザAのリポジトリ α をユーザBがフォークし、ユーザBのプルリクエストがユーザAによってマージされた場合
 - リポジトリ α に対してユーザAのcontributionが1増加
 - ユーザBも編集した分だけ増加
- その他のissue,コメントなどはカウントされない

repo_contribution(1/2)

- リポジトリ3500万件までのcontributorsを参照し、contributionを合算したランキングを作成

repo1

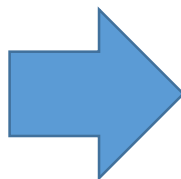
contributor	contribution
user1	3000
user2	2500
...	
user30	7

repo2

contributor	contribution
userA	1000
userB	500
...	
userK	1

...

上位30名まで
表示

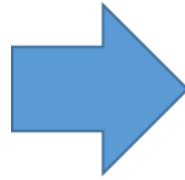


repository	contribution
repo1	30000
repo3	28500
repo8	27000
...	
...	
...	
...	
repo4	100
...	

repo_contribution(2/2)

- フォークされたリポジトリを除外するために同名のリポジトリをcontributionの一番多いものを残してランキングから削除
- ランキングにしたリポジトリの中から contribution1000以上のrepo1211913件の contributorを参照 ユーザのcontributionおよびメールアドレスなどを取得

repository	contribution
repo1	30000
repo8	27000
...	
...	
repo4	1000



contributor	contribution
user1	3000
user14	2900
...	
...	

repo_contributor(1/2)

```
[
  {
    "login": "mojombo",
    "id": 1,
    "avatar_url": "https://avatars.githubusercontent.com/u/1?v=3",
    "gravatar_id": "",
    "url": "https://api.github.com/users/mojombo",
    "html_url": "https://github.com/mojombo",
    "followers_url": "https://api.github.com/users/mojombo/followers",
    "following_url": "https://api.github.com/users/mojombo/following{/other_user}",
    "gists_url": "https://api.github.com/users/mojombo/gists{/gist_id}",
    "starred_url": "https://api.github.com/users/mojombo/starred{/owner}/{/repo}",
    "subscriptions_url": "https://api.github.com/users/mojombo/subscriptions",
    "organizations_url": "https://api.github.com/users/mojombo/orgs",
    "repos_url": "https://api.github.com/users/mojombo/repos",
    "events_url": "https://api.github.com/users/mojombo/events{/privacy}",
    "received_events_url": "https://api.github.com/users/mojombo/received_events",
    "type": "User",
    "site_admin": false,
    "contributions": 177
  },
  {
    "login": "schacon",
    "id": 70,
    "avatar_url": "https://avatars.githubusercontent.com/u/70?v=3",
    "gravatar_id": "",
    "url": "https://api.github.com/users/schacon",
    "html_url": "https://github.com/schacon",
    "followers_url": "https://api.github.com/users/schacon/followers",
    "following_url": "https://api.github.com/users/schacon/following{/other_user}",
    "gists_url": "https://api.github.com/users/schacon/gists{/gist_id}",
    "starred_url": "https://api.github.com/users/schacon/starred{/owner}/{/repo}"
  }
]
```

ユーザに関する情報を閲覧できる

リポジトリに対するcontribution

<https://api.github.com/repos/mojombo/grit/contributors>

repo_contributor(2/2)

- このページでメールアドレスなどを取得

```
{
  "login": "mojombo",
  "id": 1,
  "avatar_url": "https://avatars.githubusercontent.com/u/1?v=3",
  "gravatar_id": "",
  "url": "https://api.github.com/users/mojombo",
  "html_url": "https://github.com/mojombo",
  "followers_url": "https://api.github.com/users/mojombo/followers",
  "following_url": "https://api.github.com/users/mojombo/following{/other_user}",
  "gists_url": "https://api.github.com/users/mojombo/gists{/gist_id}",
  "starred_url": "https://api.github.com/users/mojombo/starred{/owner}/{/repo}",
  "subscriptions_url": "https://api.github.com/users/mojombo/subscriptions",
  "organizations_url": "https://api.github.com/users/mojombo/orgs",
  "repos_url": "https://api.github.com/users/mojombo/repos",
  "events_url": "https://api.github.com/users/mojombo/events{/privacy}",
  "received_events_url": "https://api.github.com/users/mojombo/received_events",
  "type": "User",
  "site_admin": false,
  "name": "Tom Preston-Werner",
  "company": "",
  "blog": "http://tom.preston-werner.com",
  "location": "San Francisco",
  "email": "tom@mojombo.com",
  "hireable": false,
  "bio": null,
  "public_repos": 60,
  "public_gists": 65,
  "followers": 18242,
  "following": 11,
  "created_at": "2007-10-20T05:24:19Z",
  "updated_at": "2015-07-23T21:47:46Z"
}
```

contributors

- 同じユーザが複数取得されるのでcontributionを合算してランキングを作成

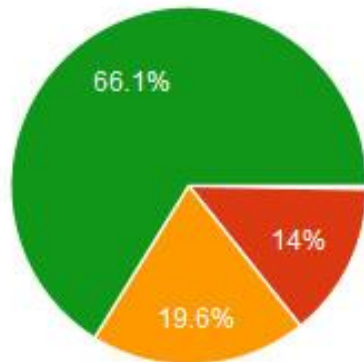
ID	user	contribution	email
2000	userA	30000	aaa@gmail.com
1000	userB	29000	bbb@gmail.com
...			
...			
3000	userC	20000	ccc@gmail.com
...			

- contribution100以上のユーザ22228人中18298人にアンケートメールを送信した

アンケート結果(1/1)

- アンケートメールを対象者18298人に送信した結果、321件の回答が得られた

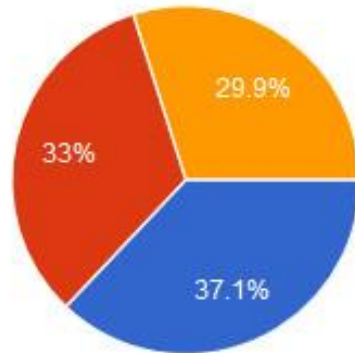
Experience with FLOSS development



	less than one year	1	0.3%
■	one to three years	45	14%
■	three to five years	63	19.6%
■	more than five years	213	66.1%

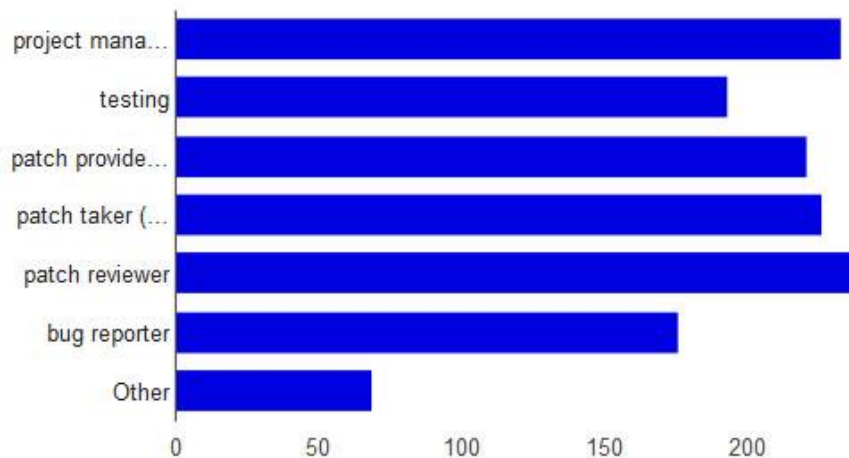
アンケート結果(2/2)

Motivation to participate in FLOSS development



for work	119	37.1%
for hobby	106	33%
Other	96	29.9%

Your role in your main project

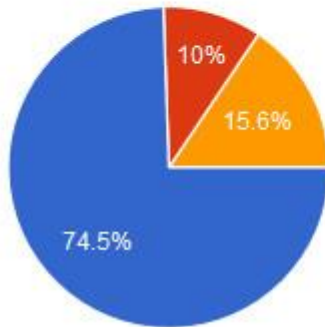


project management	233	72.4%
testing	193	59.9%
patch provider (pull-request sender)	221	68.6%
patch taker (merging pull-requests)	226	70.2%
patch reviewer	239	74.2%
bug reporter	176	54.7%
Other	69	21.4%

アンケート結果(3/3)

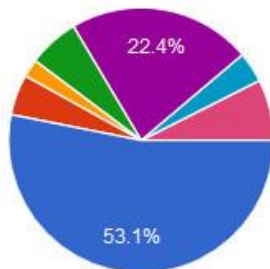
Step 2. High impact bugs

2-1. Which bug bothers you in bug fix?



a bug decreasing user satisfaction	239	74.5%
a bug requiring schedule changes	32	10%
Other	50	15.6%

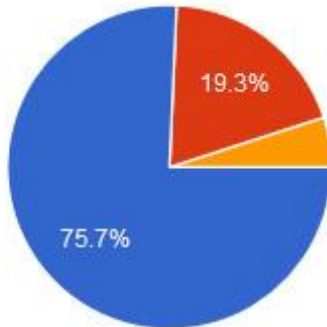
2-2-a. What kind of bug would be much more important to be fixed? (We call the following bugs "high impact bugs")



a bug threatening systems' security	171	53.1%
a bug blocking other bug fixes	16	5%
a bug found in unexpected timing and location	7	2.2%
a bug deteriorating system's performance	20	6.2%
a bug introduced in a newer release and breaking functions of older releases	72	22.4%
a bug introduced in older releases and found in a newer releases	12	3.7%
Other	24	7.5%

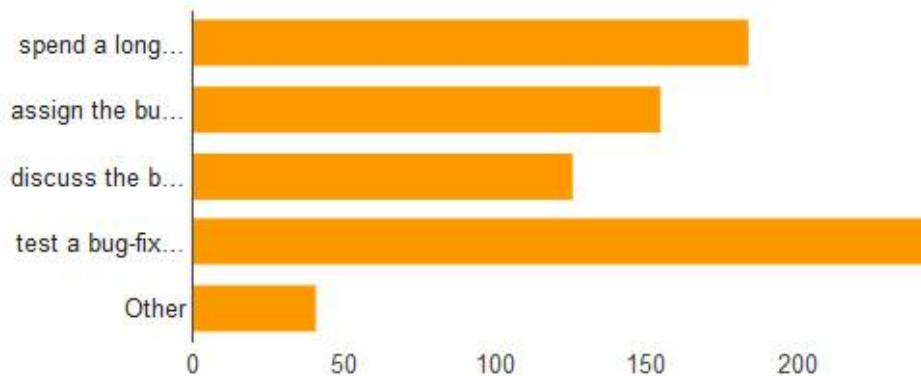
アンケート結果(4/4)

2-2-b. Who is mainly influenced by the above bug?



users	243	75.7%
developers (due to priority changes and etc.)	62	19.3%
Other	16	5%

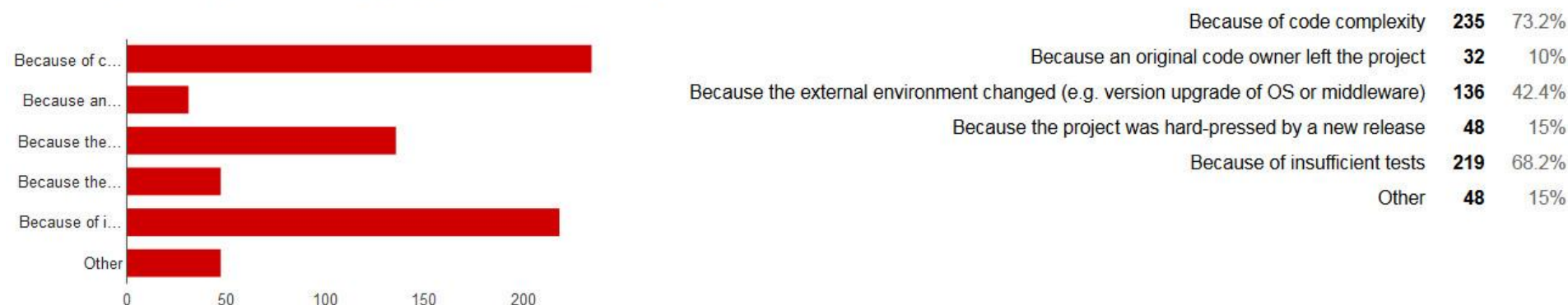
2-2-c. How do you (or project) fix the bug? [Multiple selection allowed]



spend a long time to reproduce the bug	184	57.3%
assign the bug to an experienced developer	155	48.3%
discuss the bug at length	126	39.3%
test a bug-fix-patch carefully	244	76%
Other	41	12.8%

アンケート結果(5/5)

2-2-d. Why is the bug introduced? [Multiple selection allowed]



今後

- どのカテゴリのリポジトリで活動している人がどのバグを影響力が高いと答えているかの分布を出す
- バグを読む観点を考える
- 卒論目次を完成させる