オープンソースソフトウェア開発における役割分担の実態調査

プロジェクトマネジメントコース　矢吹研究室　1042067　関口 元基

1. 研究背景

オープンソースのソフトウェア開発のためには，バージョン管理システムやWiki，バグ追跡システム，メーリングシステムがよく利用される[1]．これらのツールをまとめて提供するのがホスティングサイトであり，その代表的なものがGitHubである．

現在，オープンソースソフトウェア（以下OSS）開発は数多く世界中で行われており，その多くはGitHubで行われている．GitHubのAPIを使用することによりプロジェクトの活動の様子やプロジェクトメンバの活動ログを収集することができる．

GitHubのAPIには以下のようなものがある．

()内の英記号は主成分分析でグラフの値が読みにくくなってしまうので英記号として置き換える．

* DeleteEvent (c) : デリートを行ったイベント．プロジェクトで行われていたイベントを削除したという活動のログ．
* IssueCommentEvent (d) : Issuesにコメントを行ったイベント．プロジェクトメンバに限らず，第三者もIssuesにコメントしたという活動のログ．
* IssuesEvent (e) : Issuesを行ったイベント．プロジェクトメンバに限らず，第三者もIssuesを発行したという活動のログ．
* PullRequestReviewCommentEvent (g) : プルリクエストにコメントを行ったイベント．管理者がプルリクエストにコメントをした活動のログ．
* PushEvent (h) : プッシュを行ったイベント．変更履歴をアップロードした活動のログ．
* ForkEvent (j) : フォークを行ったイベント．自分のアカウント内に既存のリポジトリの複製をつくったという活動のログ．
* WatchEvent (i) : スター（お気に入り）をしたイベント．自分が気になっているリポジトリにスターを付けたという活動のログ．

このようなGituHubのAPIを使い，活動ログを取得・解析し，結果を解釈することによって，今まで明らかになっていなかったOSSプロジェクトにおける各メンバの役割の分担状況を明らかにすることが期待できる．

1. 研究の目的

GitHub上で公開されているOSSを調べることで，プロジェクトメンバの役割分担の実態を明らかにする．

1. 研究方法

以下の手法で研究する．

1. APIを使用し，GitHub上で行われているプロジェクトメンバ全員の活動ログを収集する．
2. 各イベントが何回行われているかの活動ログを，プロジェクトメンバごとの一覧表にまとめる．
3. 統計解析ツールRを使用し，②で得られた一覧表のデータを主成分分析し，図1のようなグラフで表し，結果を解釈する．

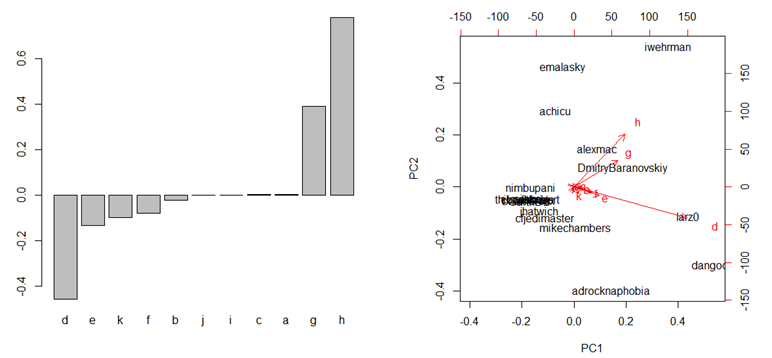
****

図 1 左がAdobe Systemsの第一主成分，右が主成分のスコア

1. 結果と考察

各プロジェクトの主成分には次の傾向が見られた．PushEventとWatchEventの絶対値が大きく，正負が逆であることが分かった．つまり，Pushする行為とリポジトリにスターを付ける行為は別々のメンバが行っていることが多い．

GitHubのプロジェクトメンバの活動ログを調べることで，プロジェクトのメインプログラマや外部のサポートを行っているメンバなどの役割分担の実態を解明することができた．また，主成分分析を行うことで，OSSプロジェクトの役割を把握することができるとともに，第三者がより深くOSSでのプロジェクトを知ることができ，今後のOSSの発展やプログラマの技術の向上に役立てることができる．

1. 参考文献

[1] Fogel Karl. オープンソースソフトウェアの育て方. ツールが一通り揃ったホスティングサイト.http://producingoss.com/ja/web-site.html,2013.12.24