

# OSS 開発者のプロジェクトへの貢献度と社交性の相関分析

プロジェクトマネジメントコース 矢吹研究室 1342081 辻岡 大知

## 1. 背景

ソフトウェア開発では、GitHub を用いることが多い。GitHub とはコンピュータ上で作成、編集されるファイルの変更履歴を管理するためのバージョン管理システムである。[1]

システムエンジニアはコミュニケーション能力が必要とされる仕事である。私は GitHub に公開されているプロジェクトを調べているうち、活発に活動しているユーザは Google+ や Twitter で交友関係が広い場合が多いことに気が付いた。そこで私は GitHub を用い、活発に活動するシステムエンジニアはコミュニティが広いのではないかという仮説を立て、その仮説を検証するため本研究を行った。

## 2. 目的

ソフトウェア開発で活発に活動しているユーザはコミュニティが広いのではないかという仮説を検証する。

## 3. 手法

活発に活動しているシステムエンジニアのユーザ情報を集めるため GHTorrent を使用した。GHTorrent とは GitHub のユーザ情報やプロジェクトの情報など様々なデータが格納されているダンプファイルのことである。

活発に活動しているシステムエンジニアのコミュニティを調べるため、Google+ と GitHub におけるユーザ情報の一つである contribution を使用した。contribution とはユーザがどの程度 GitHub 上で活動しているのかを定量的に知ることのできる値である。

集めたユーザのコミュニティ関係を調べるため Google+ を使用した。Google+ とは Google が運営する SNS のことである。

以下のような手順で研究を進めた。

1. GHTorrent から 2 万件のユーザ情報をランダムサンプリングで抽出した。
2. 1 で抽出したユーザのメールアドレスを、

GitHubAPI を用いて取得した。

3. 取得したメールアドレスを用い Google+ におけるフォロワー数を調べた。
4. ユーザの contribution 数と Google+ におけるフォロワー数の相関関係を調べた。

## 4. 結果

本研究の結果が図 1 の散布図である。回帰分析の結果、重相関 R の値が約 0.19 となったため、GitHub における contribution 数と Google+ におけるフォロワー数にほとんど相関はないということが分かった。

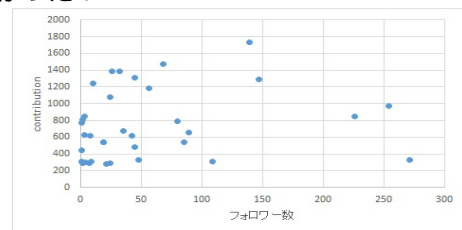


図 1 contribution 数とフォロワー数の相関

## 5. 考察

本研究を通して、活発に活動しているユーザは Google+ におけるコミュニティが必ずしも広いわけではないことが分かった。今回は Google+ を用いたが、現在流行している Facebook や Twitter など他の SNS を用いることにより、今回とは違った結果が得られるのではないかと考えた。

## 6. 結論

本研究を通して明らかとなった相関関係などはなかった。しかし研究を進めるための過程で GHTorrent 及び GoogleAPI を使用することができるようになった。今後の課題として研究を行う段階で最も流行している SNS などでデータを取得し、統計を取っていく必要がある。

## 参考文献

- [1] 大塚弘記. GitHub 実践入門による開発の変革. 技術評論社, 2014.