ユーザ関係の分析によるTwitterコミュニティ抽出

PMコース　矢吹研究室　1142123　渡邊雄大

1. 研究背景

コミュニケーションツールとして，Social Networking Service（SNS）を使用している人はとても多くいる．その中でもTwitterは，SNSを代表する１つである. なぜならTwitterはアクティブユーザ数が2億3千万人もいることだけでなく，ツイートと呼ばれるマイクロブログが一日平均で5億件も送信されているからだ（2014年9月現在）．そのためTwitterは，調査する価値のあるSNSの１つであると考えられる．

　Twitterはツイートと呼ばれる短い文字列を投稿するためのサービスである．自分以外のユーザのツイートを読むためには，そのユーザのページにアクセスするか，そのユーザをフォローする必要がある．フォローしているユーザのツイートは，ひとまとめにされ，タイムラインを形成する．誰が誰をフォローしているかという情報（フォロー関係）は，Twitterにおけるユーザのつながりの一つの表現である．仮にTwitter上でユーザがコミュニティを形成していたとすれば，フォロー関係にもそれが反映されていると思われる．そのコミュニティを抽出することができれば，フォローすべきユーザの発見が容易になるなど，Twitterのユーザビリティが大きく向上することが期待される．

2.　研究目的

　検索したいユーザのTwitter上に持っているフォロー関係から，そのユーザの持つ実際のコミュニティをTwitterの機能であるリツイートを解析することによって見つけ出す．さらに, この研究を行うことで，プロジェクトを円滑に行う為の人的資源マネジメントとして活用することを目指す．

3.　類似研究

　ユーザのフォロワーとのつながりについての研究として[1]が挙げられる．この研究では，フォロー関係からクラスタリング分析を行い，その結果からクラスタを出力できるシステムを作成している．この研究はシステムが結果として出力したクラスタの内容と，手作業で作成した理想のクラスタを比べることで研究の評価を行っている．

4. 研究方法

　この研究を説明する際に使用する用語を以下の表1にまとめた．

表1 使用する用語

|  |  |
| --- | --- |
| リツイートA | 観測されたリツイート |
| ツイートA | リツイートAのオリジナルのツイート |
| ユーザA | ツイートAのユーザ |
| ユーザB | リツイートAのユーザ |
| 指標1 | ユーザBがユーザAをフォローしていない確率 |
| 指標2 | ユーザBがツイートAをリツイートした人たちをフォローしていない確率 |

まずこの研究では指標1と指標2を導き出す．そのためにTwitterAPIのUser streamsの機能を使い，自分のフォローしているユーザの行ったリツイートのみを抽出し保存する．この保存されたリツイートの中からリツイートAを使用し研究を行ったとする．この場合はユーザBのフォローしている相手とツイートAをリツイートしたユーザB以外の相手を取得する．取得したリツイートの全体と指標1，指標2の関係は以下の図1のようになる．

指標1

指標2

リツイート全体

図1 リツイート全体，指標1，指標2の関係

5．研究結果

　この研究では成果物を研究の結果から得られた2つの指標とする．指標1は60％であった．指標2は40%であった．

6．考察

　指標1と指標2の結果から確認できるように，多くのユーザ達がフォローして無いユーザのツイートも確認し，リツイートを行っていることが分かった．これはフォロー関係によって繋がっていると思われていた従来のTwitterユーザ同士のネットワークの形ではない別の形のネットワークがある可能性が示された．

参考文献

[1] 山下拓也・佐藤晴彦・小山聡・栗原正仁(2013)「フォロー関係に基づくTwitterユーザの分類」『第75回全国大会講演論文集』pp.107—109 情報処理学会