ユーザー関係の分析による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ー数が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．自分以外のユーザーのツイートを読むためには，そのユーザーのページにアクセスするか，そのユーザーをフォローする必要がある．フォローしているユーザーのツイートは，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におけるユーザーのつながりの一つの表現である．仮にTwitter上でユーザー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ー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

　検索したいユーザーのTwitter上に持っているフォロー関係から，そのユーザー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

　ユーザー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

　今回の研究はMySQLのDB上で集めたデータの管理を行った．

まず調査を行うためにTwitterAPIのUser streamsの機能を使い，自分のフォローしているユーザーの行ったリツイートのみを抽出し保存してサンプル1を作成した．その中からリツイートしたユーザーが，そのリツイートされたツイートをつぶやいたユーザーをどれだけの割合でフォローしているのか調べた．

つぎにそのサンプル1の中からツイートごとに，リツイートを行った他のユーザー（最新100件のユーザーのみ）を集めてサンプル2を作成する．このサンプル2の中から最初にリツイートを行ったユーザーとどれだけフォローしているのかの割合を調べた．

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つの知見とする．1つ目はリツイートをしたユーザーが，そのツイートのつぶやきを行ったユーザーをフォローしている割合であり，結果を図1の左側に示す．2つ目はリツイートをしたユーザーがそれ以外にリツイートを行ったユーザーをフォローしている割合である．これらの結果を図1の右側に示す．

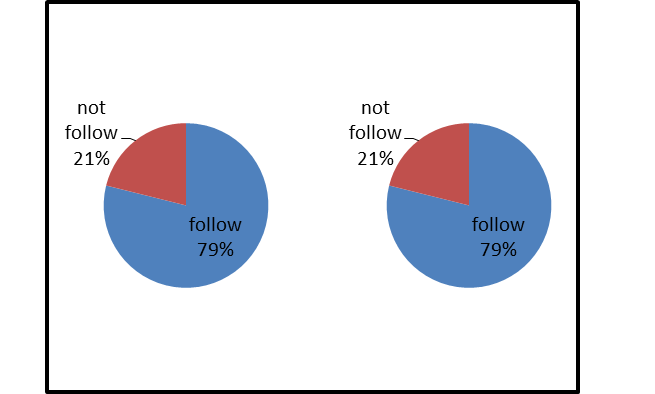


図1 研究結果

6．考察

　今回の結果からリツイートを行っている人は，全くフォローしていない人をリツイートしている人は20%もいることが分かった．これは多くのユーザー達がフォローしているユーザーのツイート以外のツイートを見ていることが分かったこのようにリツイートの行われる形を調べていくことで，ユーザー同士が本当はどのような形でツイートを確認しコミュニケーションを取っているのかが分かるための足がかりになると思われる．

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ユーザーの分類」『第75回全国大会講演論文集』pp.107—109 情報処理学会