

Twitter におけるリツイート経路の重ね合わせによるユーザ発見支援

Visualization of overlapping pathways of Retweet in Twitter

太田 侑介*
Yusuke Ota

寺田 実*
Minoru Terada

丸山 一貴†
Kazutaka Maruyama

1 背景

近年, Twitter^{*1}が普及してきた。Twitter は「ツイート」という 140 文字以内の短いメッセージのやり取りで、友達や家族、同僚、自分と繋がっているユーザなどと情報交換を行えるサービスである。ユーザは他のユーザを「フォロー」することで、そのユーザのツイートを自分のタイムライン上で閲覧することが可能となる。グラフ理論で言えば、ユーザをノードとすると、他のユーザをフォローすることは、ノード間に単方向のエッジを繋ぐこと相当する。このため、Twitter においてユーザのフォロー、フォロワーの関係からソーシャルグラフが浮かび上がる。

2 目的

Twitter を利用するにあたって、ユーザは同一の興味を持つ他のユーザを探す必要がある。自分のフォロー、フォロワーなどから探したり、興味あるワードで検索したり出来るが、探すのは容易ではない。そこで本研究では、リツイート(3章にて後述)がどのように伝播したかを可視化することで、同一の興味を持つユーザの発見を支援した。特に、複数のリツイート情報を重ね合わせると、より興味が似ているユーザを判断でき、よりよい結果を導くと思われる。

3 リツイート

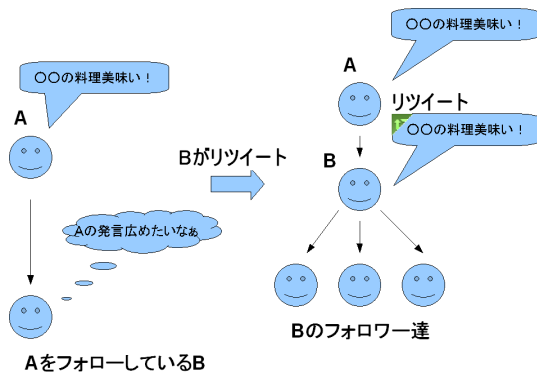


図1 リツイート例

本研究では Twitter のリツイートという機能に着目した。あるユーザが他人のツイートを自分のフォローしているユーザにも提示する機能がリツイートである。リツイートは図1のように、他人のツイートを自分のツイートとして再投稿することで、自分のフォローしているユーザにもそ

の内容を広めることができる。また、それを見たユーザがさらにリツイートをしていくほど、多数のユーザに情報が拡散され、元のユーザから関係が遠いユーザにも情報が共有される。

Boyd らの研究 [1] によると、ユーザはツイート内容を他の人に広げるためやツイート内容に対して同意をするためといったことがリツイートする理由であると言及している。そのため、リツイートを行ったユーザは、そのツイートに対して少なからず関心を持っていると考えられる。つまり、あるツイートをリツイートしているユーザ達は、その内容に対して同一の興味を持った集団であると言える。また、リツイートを行っているユーザは、自分の興味あることを他のユーザに伝えようと Twitter を積極的に利用していることから、すでに Twitter を利用しなくなったユーザと区別することができる。

4 関連研究

4.1 viratter^{*2}

viratter はリツイートの経路をツリー形式で表示するサービスである。ツリー形式で観察することで、リツイートの伝播状況の理解を支援している。リツイート伝播経路から同一興味ユーザを発見するのは、本研究と同じである。だが、一つのリツイート情報だけでは、本当にそのユーザが自分と同じ興味を持っているかを判断するのは難しい。

4.2 Mentionmap^{*3}

Mentionmap はユーザ間のフォロー関係をソーシャルグラフとして可視化したサービスである。指定したユーザとリプライのやりとりをよく行なう親密な相手を判断して表示される。リプライ関係からは、ユーザ間の親密度を理解することができる。だが、特定のユーザ間が親密であるからといって、自分と同一の興味を持っていると判断するのは難しい。リツイートをしているユーザはその内容に関して興味を持っていることが多いので、この問題を解決することができる。

5 提案システム

5.1 複数のリツイートの重ね合わせ

3章に示したとおり、リツイートを行っているユーザ達はそのリツイートに対して興味を持っていると考えられる。そのため、複数のリツイートを多く共有しているほど、ユーザ同士がより興味が似ていると考えられ、ユーザ同士フォロー対象者となる。このことは、単一のリツイート情報から

* 電気通信大学 The University of Electro-Communications

† 東京大学 情報基盤センター

*1 <http://twitter.com/>

*2 <http://viratter.jp/>

*3 <http://apps.asterisq.com/mentionmap/>

は得られない。本研究では、リツイート情報を複数重ね合わせた結果を可視化し、フォロー対象者発見を促した。

5.2 オーバーラップグラフ (図4)

5.2.1 概要

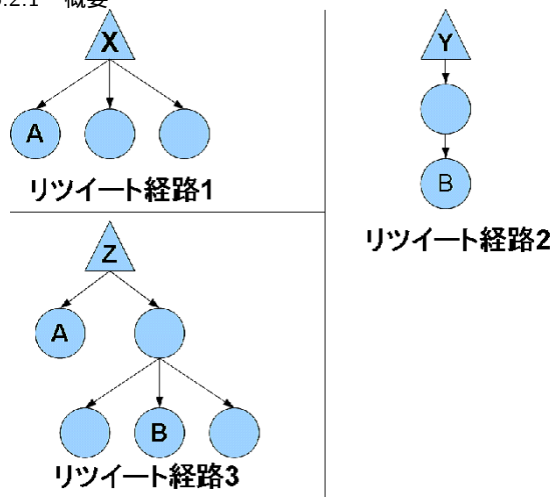


図2 3つのリツイート伝播経路情報

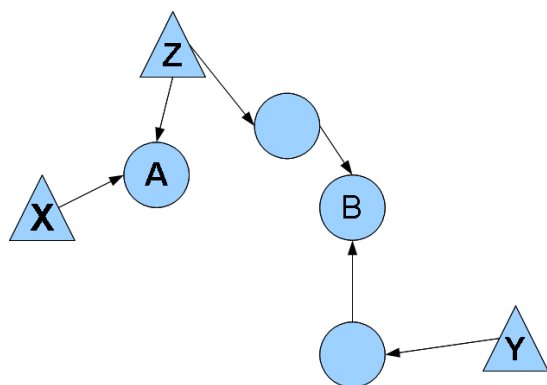


図3 オーバーラップグラフ例

オーバーラップグラフは、収集した複数のリツイート経路情報について、二回以上登場しているユーザを集め、そのユーザからリツイート元までの経路を重ね合わせたものとなっている。構成例として、3つのリツイート経路の図2とオーバーラップグラフの図3を示す。オーバーラップグラフの構成は、まず複数のリツイート伝播経路情報に二回以上登場するユーザをピックアップする。次に、伝播経路情報を順に見ていき、該当するユーザを発見したら、そのユーザからリツイート元までの経路を辿り、その経路中に出てきたユーザをオーバーラップグラフに追加していく。図2はX, Y, Zがそれぞれリツイート元のユーザであるリツイート経路情報である。ここでA, Bのユーザがそれぞれ二回登場しているため、オーバーラップグラフに表示されている。また、リツイート経路2, 3において、Bには一度しか登場していない他のユーザを介してリツイートが伝播している。これらのユーザもオーバーラップグラフに表示しているが、これはBまでの伝播経路を示すために表示している。

5.3 オーバーラップグラフ例

実際にTwitterからリツイート情報を約30件取得して作成したオーバーラップグラフが図4である。オーバーラップグラフ中における要素についての説明は以下の通り

である。

- 三角のノード：リツイート元
- ノードの濃さ：ユーザの登場回数
- エッジの向き：リツイートの伝播方向
- エッジの太さ：経路を通った回数
- 赤色のエッジ：現在注目しているリツイート経路

5.3.1 リツイート元

ノードの形が三角であるものは一度でもリツイート元となったユーザである。リツイート元となるということは、他のユーザにツイート内容がおもしろいと思われるユーザである。

5.3.2 ノードの濃さ

ノードの色の濃さは、ユーザの登場回数に比例している。また、自分のリツイートしたツイートからリツイート経路情報を取得しているため、最も色の濃いユーザ(図4ではX)が自分となる。

5.3.3 エッジの太さ

エッジの太さはその経路を通った回数に比例している。すなわち、太いエッジが繋がっているユーザは事実上、登場回数が多く、エッジが太いユーザ間ではリツイートの伝播が活発に行われていることを示している。一方のユーザについて興味があるならば、そのユーザと関係が深いもう一方のユーザにも興味に向く可能性がある。

5.3.4 赤色のエッジ

赤色で表現しているエッジは、現在注目して見ているリツイート経路である。注目するリツイートに関しては、ユーザが自由に変更することができる。このように、色を変更して見ることで、伝播経路がより分かりやすくなる。

5.3.5 潜在的ユーザ

オーバーラップグラフでは、二回以上登場しているユーザをピックアップしている。このとき、登場しているユーザの中には自分とフォロー関係が遠いユーザが存在している。このユーザは、自分の知らないところで、複数の同じツイートに対して自分と同様にリツイートを行っているユーザである。図4ではAのようなユーザが該当する。個々のリツイート経路を可視化してもそのユーザはただのユーザと扱われていた。だが、リツイートの情報を重ね合わせたことがこのユーザの発見に繋がっている。よって、オーバーラップグラフではこのような潜在的に自分と同じ興味を持っているユーザをグラフ表示することで直感的に発見することに優れていると言える。

6 実装

6.1 処理の概要

本システムの流れとして、大まかに情報取得部とグラフ作成部に分かれている。情報取得部では、TwitterAPIのJAVAラッパであるTwitter4j^{*4}を使用し、必要な情報を取得している。グラフ作成部では、得られた情報よりJUNGライブラリ^{*5}を使用してグラフを作成している。

^{*4} <http://twitter4j.org/ja/index.html>

^{*5} <http://jung.sourceforge.net/>

表 1 メソッド及び API の情報取得対応表

メソッド	呼び出される API	取得情報
getRetweetByMe	statuses/retweeted_by_me	ログインユーザによってリツイートされたツイート
getRetweets	statuses/retweets/:id	与えられたツイートの 100 件までのリツイート
getFriendsIDs	friends/ids	指定したユーザがフォローしているユーザの UserID

デルを採用している。バネモデルはエッジが繋がる二つのノード間をバネであるとして伸縮力を働かせ、さらに繋がっていない他のノードと斥力を働かせノードの位置を計算するレイアウトである。バネモデルを使うことで、ノードの重なりを除去し、見やすいレイアウトとすることができる。JUNG ライブラリには、SpringLayout として用意してある。

7 現状の仕様

本システムを構成するに当たって、TwitterAPI 制限などにより、やむなく仕様を変更した箇所がある。

7.1 TwitterAPI の利用制限

TwitterAPI を利用する場合、ユーザは API 利用回数を制限されている。TwitterAPI は、ユーザのフォローリストやリツイートしたツイートのリストなど、情報を取得するたびに利用回数がカウントされる。利用回数の制限は、API 利用による負荷を軽減するものであり、通常は 150 回/h、OAuth 認証経由の場合は 350 回/h となっている。

本システムでは逐次 Twitter より情報を取得する予定であった。だが、リツイートを行なったユーザのフォローリストを取得する必要があった。逐次情報を取得する場合、TwitterAPI 制限により、すぐ利用不可能になるため、各々のユーザのフォロワーリストとそれぞれのリツイートがいつ誰にリツイートされたかというデータをローカルに保存することにした。

7.2 リツイート経路情報の消失

リツイート経路推定時において、特定のリツイートに対して誰がリツイートを行なったかという情報は最大 100 件しか取得できなかった。このため、リツイートが 100 人以上にされたものに対しては伝播経路情報が一部失われる。経路情報が不明となってしまったユーザに関しては直接リツイート元とエッジを繋ぐことにした。直接リツイート元と繋がっているエッジの色は青くして区別をつけた。

8 評価実験

8.1 評価方法

本大学の学生 11 名に評価実験を行った。実験は、オーバーラップグラフを操作し可視化結果を見せ、アンケートに答えさせた。

8.2 結果

フォローしたいと思うユーザの選択に関するアンケート結果を表 1 に示す。ただし、回答は複数回答可能である。

8.3 考察

表 2 では、*をつけたユーザのように複数のリツイート情報でないと発見できないようなユーザがフォロー対象として多く選択された。これらのユーザは、登場回数の多いユー

表 2 オーバーラップグラフの結果

フォローしたいと思うユーザ	選択回数
*登場回数の多いユーザ	7
*太いエッジが繋がっているユーザ	8
リツイート元のユーザ	3
経路途中のユーザ	0
いない	0
その他のユーザ	1

ザである。登場回数の多いユーザは、自分と同様のリツイートを何度も行ったユーザであり、すなわち同じ内容に興味を持っていると考えられる。

また、被験者が表 2 で選択したようなユーザのプロフィールを確認するしぐさが何度か見受けられた。さらに、実験においてリツイートの情報を赤色のエッジで示したりするだけではなく、例えば、あるエッジにはどのリツイート情報が含まれているかをまとめて表示して欲しいといったような意見も上がった。これは、やはりただオーバーラップグラフを見るだけでは情報が少ないからだと思われる。だが、少なくともユーザを探す目安にはなっていたと考えられる。

9 結論

本研究では、Twitter の複数のリツイート経路の可視化によってユーザのフォロー支援を促した。評価実験の結果より、被験者はフォローしたいと思うユーザとして、主に登場回数の多いユーザ、すなわち自分と同じリツイートを多くしているユーザを選ぶ傾向があった。そのため、自分と興味似ているユーザを発見しやすくなった。特に、自分との距離が遠いユーザに関しては、経路を可視化することで直感的な発見に繋がった。そのため、リツイート経路を重ね合わせることが有用であると言及できる。Twitter ユーザは、このように普段見つけにくいユーザを発見することがより出来るようになり、それは Twitter をよりよく活用する手助けになるだろう。

参考文献

- [1] danah boyd, et al. "Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter". Proceedings of the 43rd Hawaii International Conference on System Sciences, 2010.
- [2] Zi Yang, et al. "Understanding Retweeting Behaviors in Social Networks". CIKM '10, October 25.30, 2010,
- [3] 風間一洋, 今田美幸, 柏木啓一郎. "Twitter の情報伝播ネットワークの分析". 第 24 回人工知能学会全国大会, 2010.