

## Twitter における会話ネットワークの特徴分析

### Feature Analysis of Conversation Networks on Twitter

竹田 凌輔<sup>†</sup> 武田 利浩<sup>†</sup>  
Ryosuke Takeda Toshihiro Taketa

#### 1. はじめに

SNS のなかでも、Twitter はとりわけアクティブユーザーが多く [1]、様々なコミュニケーションが行われている。そのため、Twitter における発言(ツイート)を分析した研究がしばしば行われている。例えば、高橋らの研究 [2] では、炎上の検出のために各ツイートの分析を行っている。また、村上らの研究 [3] では、ツイートの感情分析を行い、その可視化を行っている。これらはいずれもツイート単位での分析であり、一連の会話(会話ネットワーク)ごとに分析している研究はほとんどない。Twitter においては、1 つのツイートで完結せず、リプライ(返信)機能を使って会話が行われることも多い。会話の特徴を分析することで、ツイートの分類や検出に役立てることができるのではないかと思われる。

本論文では、Twitter における会話ネットワークを分析し、会話ごとの特徴、傾向の発見を目指す。

#### 2. 会話ネットワークの定義と分析

本章では、会話ネットワークの定義とその分析について説明する。図 1 に Twitter における実際の会話ネットワークの例を示す。図示のように始点となるツイートを根とする木構造になる。

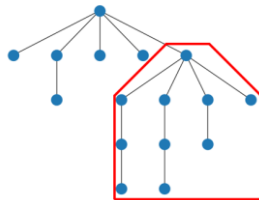


図 1 会話ネットワークと分岐の例

ここで、赤枠で囲った部分では、リプライに対してさらに複数のリプライが付いている。これを話題の分岐とみなす。根からも複数のリプライが伸びているが、これらの大半は根のツイートに対するコメントで話題自体は変わっていないので分析の対象としない。

本論文では、リプライしているユーザーとリプライされたユーザーのフォロー関係や会話の分岐に注目し、以下の 3 つの分析を行う。

##### 分析 1:

会話の話題に着目し、話題ごとにフォロー率に偏りがあるのかを確認する

##### 分析 2:

リプライまでの時間をフォロー/非フォローで分けて計測し、違いがあるか調べる

##### 分析 3:

会話の分岐に着目し、どのようなときに分岐がおこるのか調べる。

#### 3. 分析データの収集

本章では、各ツイートのデータを収集する方法について述べる。データの収集には、Twitter API v2 [4] を用いた。これを用いることで、ツイートやユーザーの各種情報を取得できる。特に v2 を選んだ理由としては、会話ごとに ID が割り振られる仕様により、会話の収集が非常に容易であることが挙げられる。以前のバージョンである v1.1 では ID による紐づけがなく、会話単位での収集は困難であった。

データ収集プログラムの作成には Python 3.8.3 [5] を用いた。作成したプログラムを用いて、リプライが付いているツイートの収集を行った。今回は 99 件の会話を収集した。

ツイートの収集後、会話ごとに割り振られる ID をもとに、リプライを収集した。また、ツイートごとの作成者情報や作成日時などの情報も表 1 のとおり取得した。

表 1 収集データ一覧

要素名	内容
id	ツイートの ID
author_id	ツイートしたユーザーの ID
text	ツイート本文
created_at	ツイート作成日時
retweets_count	リツイート数
likes_count	いいね数
followers_count	フォロワー数
replies_count	リプライ数

Twitter API では、直接フォロー関係を取得することはできないため、リプライされたユーザーのフォロワーのユーザー ID を取得し、リプライしているユーザーの ID と照らし合わせることでフォロー関係の判定を行った。

リプライまでの時間の計測は、リプライされたユーザーとリプライしたユーザーのツイート作成日時を取得し、その差分を取ることで行った。

会話ネットワークを木構造のグラフにし、画像として出力することで、分岐が起きている会話を分析していく。会話ネットワークの木構造の作成には NetworkX 2.6.2 [6] を、可視化には Matplotlib 3.4.2 [7] を使用した。

#### 4. 分析結果と考察

本章では、3 つの分析結果とその考察を述べる。

##### 4.1 分析 1 の結果とその考察

本分析では、収集したツイートのデータから、リプライしているユーザーのフォロー率と話題の傾向を調査した。

図 2 は、収集したツイートのフォロー率をヒストグラムで表したものである。

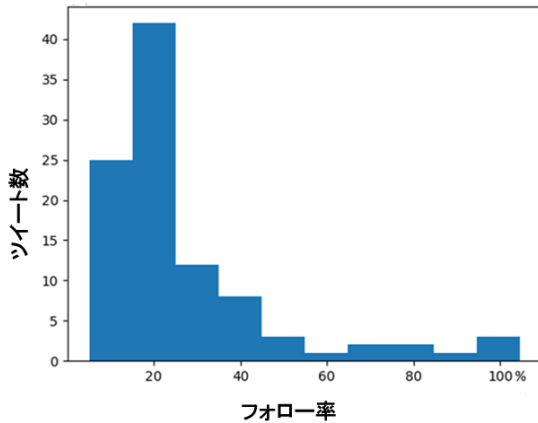


図 2 収集ツイートのフォロー率

全体の 65%以上が、フォロー率 20%未満となっていることが分かる。ここで、20%未満とそれ以上で違いがあるか見たところ、ほとんどの場合、ツイート内容に違いがあることが分かった。

20%未満のフォロー率の話題の 1 つに、意見表明や問題提起をしているツイートがあった。ただし、フォローしていないからといって否定意見が多いというわけではなく、むしろ賛同意見のほうが多かった。

もう 1 つの話題として、一般的に見て面白いと感じるツイートがあった。これらは数千～数万 RT されている場合もあり、フォローしていない人の目にも多く留まるからだと考えられる。

逆に 20%以上のフォロー率の話題として、イラスト、動画などの創作作品のツイートや、アニメやゲームなど、特定のコミュニティ内でのツイートがあった。創作もコミュニティ内でのツイートも、そのコンテンツが好きな人同士での会話になるので、フォロー率が高くなると考えられる。

#### 4.2 分析 2 の結果とその考察

本分析では、リプライされたユーザーとリプライしたユーザー間に、フォロー関係がある・なしに分けて、リプライまでにかかる時間に傾向があるか調査した。

図 3 は、フォローしているユーザー（フォロー）とそうでないユーザー（非フォロー）とでリプライまでの時間の割合を表したものである。

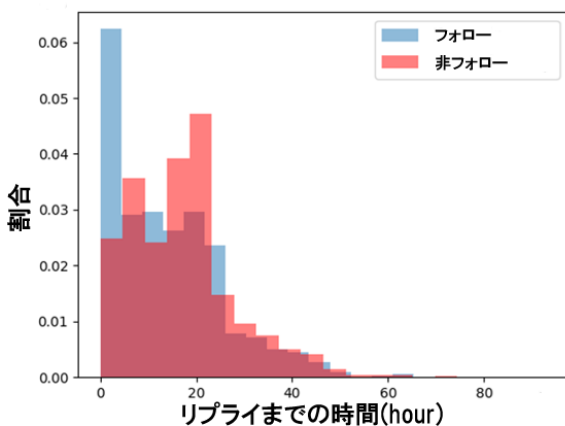


図 3 フォロー/非フォローユーザーのリプライまでの時間

フォローしているユーザーがツイートから 4 時間以内にリプライする割合は、フォローしていないユーザーの 2 倍以上であることが読み取れる。

また、12～20 時間以内にフォローしていないユーザーのリプライ率が、フォローしているユーザーのそれを大きく上回っている。

該当ユーザーをフォローしているならタイムラインにすぐ表示されるので、リプライも早くなり、フォローしていなければタイムラインに表示されるのは遅くなるので、図のような結果になったと考えられる。

#### 4.3 分析 3 の結果とその考察

本分析では、会話ネットワークを木構造で表し、特に話題が分岐しているものを選択して話題に傾向があるか調査した。

内容に差はあれども、大半はお互いの意見を言い合うときに会話が分岐していた。ごく一部、大喜利のようにリプライし合っていて分岐している会話もあった。

よって、会話の分岐に着目することで、意見を言い合っている会話の検出に役立てられると考えられる。

#### 5. まとめと今後の課題

本論文では、3 つの分析を行った。その結果、リプライされたユーザーとリプライしたユーザー間のフォロー関係に注目することで、おおよそ話題を絞ることができると分かった。また、リプライまでの時間が短ければ、フォローしている可能性が高い傾向にあることも分かった。そして、会話の分岐が起こる場合は、意見の言い合いが起きていることも分かった。これらの特徴によって、精度は高くないものの、ツイートの分類の一つの指標として用いることができると言える。

今後の課題は、サンプル数が少ないため、もっと増やすべきであることと、リツイート数やいいね数など、他のパラメータとの関係も分析することである。

#### 参考文献

- [1] 人気ソーシャルメディアのユーザー数まとめ <https://www.comnico.jp/we-love-social/sns-users>
- [2] 高橋直樹, 檜垣泰彦, “Twitter における感情分析を用いた炎上の検出と分析”, IEICE Technical Report LOIS2016-86(2017-03)
- [3] 村上奈緒, 尼岡利崇, “Twitter 上で任意の検索語句に対するネガポジ度を判定し可視化するアプリケーションの開発と研究”, EC2014(2014-09)
- [4] Twitter API Documentation | Docs | Twitter Developer Platform <https://developer.twitter.com/en/docs/twitter-api>
- [5] About Python™ | Python.org <https://www.python.org/about/>
- [6] NetworkX — NetworkX documentation <https://networkx.org/>
- [7] Matplotlib — Visualization with Python <https://matplotlib.org/>