

# リツイートユーザー一覧間の関係分析

プロジェクトマネジメントコースソフトウェア開発管理グループ 矢吹研究室 1442014 岩橋 瑠伊

## 1. 序論

Twitter は 2006 年 7 月 15 日に開設された「ツイート」と称される 140 文字以内の短文の投稿を共有するウェブ上の情報サービスである。2015 年 12 月時点で、1 カ月間 Twitter にログインしたアクティブユーザー数は 3500 万人、世界全体では 3 億 2000 万人で、約 1 割が日本国内からのアクセスである。[1]。

Twitter は「ミニブログ」や「マイクロブログ」ともいわれるが、投稿毎にタイトルを決める必要などないし、何より 140 文字という制限があるので、ひとつのツイートで書けることもただか知れている。ツイート毎に固有リンクが与えられるが、大半のツイートはタイムラインからそのうち流され、忘れ去られる。数日前のものでさえ、過去のツイートを掘り返す人は少ない。しかし、そういうサービスだからこそユーザーは気兼ねなくカジュアルに投稿できる [2]。

Twitter では、ツイートをすると自分のフォロワーのしているタイムラインに自分のツイートが表示される。逆に、自分がフォローしている人がツイートをすると自分のタイムラインにそのツイートが表示される。その他の Twitter の機能にリツイートというものがあり、簡単に説明すると、元のツイート者のユーザー名のまま、自分のフォロワーのしているタイムラインに転送する仕組みである。この機能を使うことで自分が興味深い、拡散したいと思ったツイートを自分のフォロワーに伝えることが可能になる。リツイートされて自分のタイムラインに表示されるツイートを見ていると、ユーザーのクラスタによってリツイートするツイートの内容がそのクラスタに関連するものに絞られているように感じられた。このことから、同じツイートをリツイートしているユーザー間にクラスタの一致を見い出せるのではないかと考えた。これが正しければ、同じツイートをリツイートしたユーザー同士趣味嗜好が一致するということになる。つまり、Twitter の機能の一部であるおすすめユーザーの表示機能の精度に繋がると考えられ、Twitter 社の機能向上の手助けを出来るのではないかと考えた。

## 2. 目的

比較的レスポンスの高いツイートをリツイートしているユーザーとユーザーの間でクラスタは一致しているのかを調べる。どのような種類のツイートにはどのようなクラスタのユーザーが集まるのか分析する。最終的には、Twitter のおすすめユーザーの機能の精度向上に貢献する。

## 3. 手法

TwitterAPI を利用し、300 リツイート以上されたツイートをランダムで 100 個集める。次に、それらのツイートをリツイートしたユーザーのプロフィールをランダムで 100 個集めてそのツイートの内容にユーザーのプロフィールが一致するかを調べる(ここでのプロフィールとはアイコンとプロフィール本文とする)[3]。プロフィールのみでユーザーのクラスタが判断できないときはそのユーザーのツイートを読みに行き、特徴のある単語からクラスタ分類を図る。プロフィールかツイートでクラスター一致と判断した場合は一致と判定する。非公開アカウントに関してはツイートの内容が見られない為、プロフィールから判断不能な場合対象から除外し、別のユーザーで補完を行う。プロフィールとツイート内容を見てもクラスタ判断不能、クラスタ違いとなった場合不一致と判定する。一致数と不一致数を洗い出し、その比率はどうなっているかで関係分析を行う。

## 参考文献

- [1] THE HUFFINGTON POST. Twitter が国内ユーザー数を初公表「増加率は世界一」.  
[http://www.huffingtonpost.jp/2016/02/18/twitter-japan\\_n\\_9260630.html](http://www.huffingtonpost.jp/2016/02/18/twitter-japan_n_9260630.html)  
(2016.06.30 閲覧)。
- [2] 近藤淳也. ネットコミュニティの設計と力. 角川学芸出版, 第 05 版, 2015.
- [3] 井上乃佑. Twitter におけるユーザープロフィールと拡散力の関係分析. 2014.