

影響力の範囲を考慮した Twitter における影響力推定手法

An Influence Estimation Method in Twitter Considering to Range of Influence

岡本 拓馬†
Takuma Okamoto

松本 和幸†
Kazuyuki Matsumoto

吉田 稔†
Minoru Yoshida

北 研二†
Kenji Kita

1. はじめに

急速な利用者の増加により発展してきた SNS やマイクロブログといったソーシャルメディアにおいて、多くのユーザが情報発信源として多くの情報を発信している。そのため、有益な情報を効率良く発見する方法がこれまで以上に必要とされている。有益な情報を発見する手段の一つが、情報発信者が持つ影響力を指標とすることである。Twitter 上では影響力の高いアカウントは有益な情報を発信しているという仮定に基づき、アカウントごとの影響力の推定手法を提案する。しかし、影響力を定量的に評価することは一般に困難であるため、影響力は影響を与えた範囲と定義する。

SNS やマイクロブログなどのソーシャルメディアでは、一般人から企業が運営する公式アカウントまで非常に多くのユーザが情報を発信している。それらの情報発信源から情報を得たいユーザはそれらの情報源との間にリンクを張る (Twitter 上ではフォローするという) ことでそれらの情報源から効率的に情報を得ることができるようになる。

近年、Twitter や Facebook といったソーシャルメディアのユーザ数が増加しており、特に Twitter は、2014 年 6 月時点では国内のユーザ数が 1980 万人[1]といわれている。そのため、膨大な情報源から、自身が必要としている有益な情報を探索するための指標が必要となる。先行研究ではフォローの前後関係から有益な情報を発見する能力に長けたアカウントを探す研究などが存在するが、アカウントの影響力自体を定量的に表す研究に乏しい。また、ソーシャルメディアの影響力を推定するサービスは複数存在するが、いくつかの問題点がある。

本研究では Twitter アカウントの情報を利用して、影響力を推定するサービスとの比較をおこないながら影響力の推定手法を提案する。

2. 既存の影響力推定サービス

急速に発展する SNS やマイクロブログといったソーシャルメディアにおける発言者の影響力を測定するサービスは複数存在する。例えば、Qrust[2]や Social Leaders[3], Klout[4]といったサービスである。これらのサービスはソーシャルメディアの情報を基に 100 点満点で影響力のスコアを計算し、提示するサービスである。例えば、Twitter の影響力を求めるならば、アカウントのフォロワー数、フォロワー数、最近のツイートのリツイート数やお気に入り数などを参考に影響力が導きだされる。以下のグラフは先に述べた 3 サービスを 2015 年 2 月 3 日時点の日本においてフォロワー数上位 4 アカウントを対象にサービスを利用した際の影響力スコアを比較したものである。

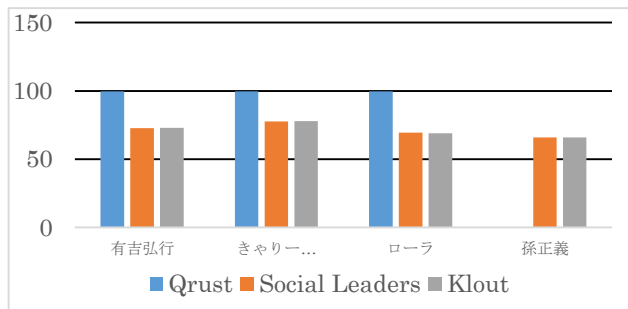


図 1 影響力推定サービスの比較

図 1 のグラフをみると、影響力推定サービスの問題点がいくつか浮かび上がる。

- 各アカウントはフォロワー数や RT 数において大きな差があるにもかかわらず、同水準のスコアとなっている
- Qrust における孫正義氏のスコアが過小評価されている。
- Social Leaders, Klout における孫正義氏のスコアが半年以上投稿していないにもかかわらず他のアカウントと同水準のスコアとなっている。

以上の問題点から影響力推定に改善の余地が考えられる。本研究では、これら問題点を是正した推定手法を提案する。

3. Twitter API による情報取得

本研究で影響力の推定をおこなう上で必要な情報を Twitter API を用いて取得する。2015 年 2 月 3 日時点の日本におけるフォロワー数上位 4 アカウントを対象にフォロワー数および一日で最もリツイートされたツイートを取得した。



図 2 取得したツイート

† 徳島大学, Tokushima University

4. 影響の範囲を考慮した影響力推定手法

本研究では、影響力は影響を与えた範囲と定義している。よって Twitter 上で必要となる情報はアカウントのフォロワー数およびツイートに対して RT された数および RT したアカウントのフォロワー数となる。以上のフォロワー数、すなわちツイートを閲覧したアカウントの数が影響を与えた範囲となる。しかし、Twitter API の仕様上 RT したアカウントが全て取得できないため RT 者のフォロワー総数は取得できたアカウントのみから計算した近似値となる。Twitter の情報伝播の特性として、最初に情報を受け取るのはアカウントのフォロワーである。情報を受け取ったフォロワーがさらにリツイートすることでリツイート者のフォロワーに情報が伝播する。この繰り返しで影響を与えた範囲が増え続け、最終的な影響力となる。図 3 にて影響を与えた範囲と影響力の関係を示す。

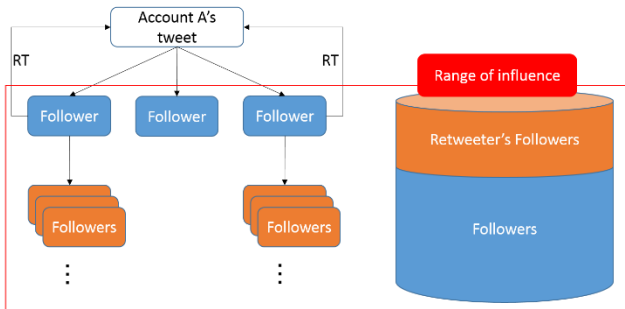


図 3 影響を与えた範囲

式(1)においてアカウント i の一日の全ツイート K からリツイート数が一番多いツイートを x とする。式(2)では x を基にフォロワー数とリツイートの結果としてツイートを閲覧したフォロワー数を算出した結果をアカウント i の影響力とする。

$$x = \operatorname{argmax}_{k \in K_i} F_{RT}(k) \quad (1)$$

$$Influence_i = N_{Fi} + \left(\frac{1}{|Y|} \sum_{y \in Y} N_{Fy} \right) * F_{RT}(x) \quad (2)$$

- K_i : アカウント i の一日の全ツイート
- x : 一日で最も RT されたツイート
- $F_{RT}(x)$: x がリツイートされた数
- Y : x にリツイートしたアカウントの集合
- $|Y|$: $F_{RT}(x)$ 内で取得できたアカウント数
- N_{Fy} : y のフォロワー数

5. 実験

5.1 実験設定

実験データとして 2015 年 2 月 3 日時点の日本においてフォロワー数上位 4 アカウントのツイートを取得した。取得したツイートで使用するデータを表 1 に示す。

表 1 実験データ

名前	リツイート数	フォロワー数
有吉 弘行	4418	4010924
きゃりーぱみゅぱみゅ	1509	2901406
ローラ	4144	2604000
孫正義	0	2332652

5.2 実験結果

取得したツイートを基に Twitter API よりリツイート者のフォロワーを取得し、提案手法の数式を適用した結果を次のグラフに示す。

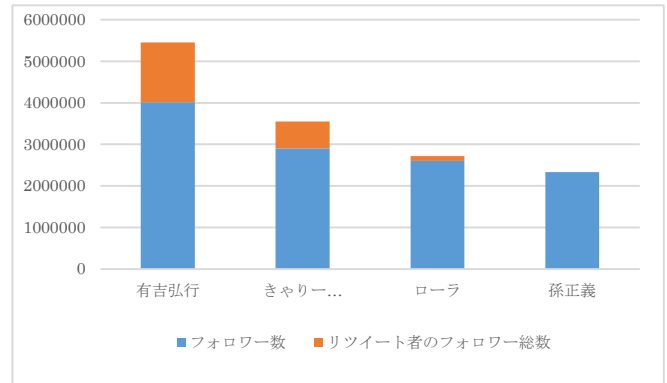


図 4 実験結果の比較

6. 考察

実験結果より、既存の影響力推定サービスの問題点であった、同水準のスコアとなる現象および、過小評価されていた点を共に解決できた。結果のグラフはツイートが影響を与えた範囲を考慮した場合に、この程度の差が生じることを表している。

本研究の問題点として、評価が困難な点が挙げられる。結果のグラフは影響を与えた範囲を基に既存サービスの問題点を是正し、アカウント間の影響力の差を示したものであるが、評価の手法がなく、影響力として妥当かどうかの判断が難しい。また、一日のみのツイートを対象としているため、アカウントのリツイート数に大きな差が出てしまう。今後は評価手法の模索と影響力推定の期間を変更できるようなデータを取得し改善する必要がある。

7. まとめと今後の課題

本稿では、Twitter による情報探索の際に新たな指標として影響力を挙げ、影響力は影響を与えた範囲であると定義した。この定義を基に影響力推定手法を提案した。また、既存の影響力推定サービスと比較し、問題点を解決したことを実験結果により示した。今後はデータ取得の基準を定めることで課題の改善に取り組んでいきたい。

参考文献

- [1] GaiaX SocialMedia Lab.
<http://gaiax-socialmedialab.jp/socialmedia/368>
- [2] Qrust.
<http://qru.st/>
- [3] Social Leaders
<http://socialleaders.jp/>
- [4] Klout
<https://klout.com/home>