

Twitter のトレンド機能におけるセレンディピティを考慮した ランキングの再評価手法

Re-Evaluation of the Ranking of Trends in Twitter by Using Their Serendipity

佐伯 卓郎

Takuro Saeki

法政大学情報科学部デジタルメディア学科

E-mail: takuro.saeki.3q@stu.hosei.ac.jp

Abstract

Twitter is one of the most popular social networking services in the world. The users of Twitter can post (tweet) messages about what they are interested in, and share (retweet) tweets posted by other users. Twitter has a function called "trend", which allows the users to find topics and words that are attracting attention and about which many users made tweets. However, it is still difficult to find topics in which a user is not interested but may get interested in the future. This paper proposes a method that re-evaluates the ranking of trends to give serendipity (the concept of attractiveness and unexpectedness to a user). The proposed method re-evaluates the ranking by adopting the previous method used to recommend users with serendipity. The proposed method constructs a function for calculating the serendipity of the trends based on the tweet information of the users recommended by the previous method and the users selected from those who made tweets about each trend. The serendipity of a trend is obtained by calculating the similarity of the tweet information. This paper shows the result of an experiment on the calculation of the correlation between the trends ranked by the proposed method and the trends ranked by subjective evaluation.

1. はじめに

Twitter は世界で最も人気を集めているソーシャルネットワークサービスのひとつである。Twitter のユーザは、自分が関心のあることについて日常的にミニブログ形式の投稿 (ツイート) をしたり、他人のツイートを自分のフォロワーに共有 (リツイート) したりすることができる。Twitter にはトレンドという機能があり、リアルタイムで注目され、関連する多くのツイートがされているようなトピックやワードを確認できる。トレンドの表示には、地域ごとに算出されるランキング形式のトレンドと、ユーザごとに推薦されるおすすめトレンドの 2 種類がある。しかし、これらのトレンド表示だけでは、ユ

ーザがまだ興味をもっておらず、将来的にそれを知ったときに興味を持つ可能性があるようなトピックを見つけることは難しい。

本研究では、Twitter のトレンド機能のランキング表示を再評価して、自分にとってセレンディピティ (ユーザにとって魅力的かつ意外性があるという概念) のあるように順位付けし直す手法を提案する。再評価のために、セレンディピティのあるユーザを推薦する既存手法 [1] を利用する。この手法によって推薦されたユーザと各トレンドから抽出されたユーザのツイート情報をもとに、ランキングのセレンディピティの評価値を算出する。セレンディピティの評価値はそれらのツイート情報の類似度によって算出する。提案手法による並べ替え結果と主観による並べ替え結果との相関を計算する実験を行い、提案手法の有効性を評価する。

2. 関連研究

セレンディピティは推薦システムにおいて近年重視されている評価指標である。安部ら [1] は Twitter におけるセレンディピティを考慮したユーザ推薦手法を提案した。この手法は、ユーザ間のフォロー関係と類似度に基づいて推薦ユーザを発見する。Ge ら [2] はセレンディピティをユーザにとって発見や予期ができずかつ興味のあるものと定義している。Herlocker ら [3] はセレンディピティのあるものはユーザの興味を広げるようなものであると定義している。

Twitter のトレンド機能とは異なるアルゴリズムで、独自のトレンド機能を提案する研究が行われている。例えば、Blei ら [4] によって提案された Latent Dirichlet Allocation (LDA) という手法がよく用いられる。LDA はドキュメントに隠されているトピックを確率に基づいて抽出してトピックの種類ごとに分類できる。

3. 準備

本研究では、安部ら [1] の提案した、セレンディピティを考慮したユーザ推薦手法を利用する。

3.1. 安部らの手法の概要

安部らの手法は、ユーザ間の類似度を計算することによってユーザを推薦する手法である。被推薦ユーザとそ

の直接のフレンド (1-hop 先のフレンド) との類似度と、1-hop 先のフレンドとさらにそのフレンドのフレンド (2-hop 先のフレンド) との類似度を考慮することによって、被推薦ユーザにとって未発見で興味のあるユーザを発見する (図 1)。より具体的には、1-hop 先のフレンドのうち類似度の低いユーザを抽出し、さらに 2-hop 先のフレンドのうち類似度の高いユーザを推薦ユーザとすることによってセレンディピティのあるユーザを推薦する。



図 1 安部らの手法の概要図 [1]

3.2. 類似度の計算方法

安部らの手法でユーザ間の類似度を計算するために、ユーザのツイート情報を用いる。ツイートに含まれている単語はそのユーザの興味を表していると考えられる。そのため、ユーザの最新のツイート 200 件を取得し、そこに含まれている単語の情報を抽出し、これをユーザの興味を示す特徴量とする。

ユーザのツイート情報から得られた単語の情報を用いて、ユーザ i, j 間の類似度は以下の通り表される。

$$\text{cos}_{\text{tfidf}}(i, j) = \frac{\sum_{k=1}^K w_{i,k} * w_{j,k}}{\sqrt{\sum_{k=1}^K w_{i,k}^2 * \sum_{k=1}^K w_{j,k}^2}}$$

ここで、 $w_{i,k}$ は tf-idf による単語の重みを考慮したユーザ i の単語 k の投稿頻度であり、ユーザ i による単語 k の投稿頻度 $\text{tf}_{i,k}$ 、全てのユーザ数 U 、単語 k の文書頻度 uf_k を用いることで以下の通り表される。

$$w_{i,k} = \text{tf}_{i,k} * \log \frac{U}{\text{uf}_k} + 1$$

4. 提案手法

本研究で提案する手法は、安部らの手法によって推薦されたユーザのツイート情報を利用してトレンドのランキングを再評価し、被推薦ユーザにとってセレンディピティのより高い順に並べ替える。推薦されたユーザにはセレンディピティがあるため、そのユーザのツイート情報を用いて評価されたトレンドのランキングにも同様にセレンディピティがあるという考えに基づく。

4.1. 安部らの手法の利用

本手法はリアルタイムのトレンドのランキングを扱う。ユーザの興味やフォロー関係はすぐに変化する可能性があるため、できるだけ短時間で必要なツイート情報を収集する必要がある。本手法で利用する安部らの手法では、実行速度の向上のため、あるユーザから探索するフレンドの数を最大で 20 人までとする。また、1-hop 先のフレンドのうち被推薦ユーザとの類似度の最も低い 10 人を抽

出し、2-hop 先のフレンドのうち抽出された 1-hop 先のフレンドとの類似度の最も高い 10 人を抽出する。最終的に抽出された 10 人のユーザを推薦ユーザとする。

4.2. トレンドのランキングの再評価

本手法の処理手順は以下の通りである。まず、トレンドのランキングを可能な限り取得する。トレンドのランキングは、Twitter API を用いることにより最大で上位 50 位まで取得できる。取得した各トレンドについて、そのトレンドのキーワードを含む最新のツイートを取得し、それらをツイートしているユーザを 10 人抽出する。そのユーザと安部らの手法によって推薦されたユーザとのツイート情報の類似度を計算する。この類似度の計算方法は安部らの手法内で定義されたものと同じとする。計算された類似度の平均が高い順にトレンドのランキングを並べ替える。

4.3. 実装

本手法の実装に Python を用いた。Twitter のデータの取得に Twitter API を使い、ライブラリ tweepy を用いた。ツイートを単語に分割するために形態素解析エンジン MeCab を用いた。

5. 実験

5.1. 実験データ

本実験では著者自身の Twitter アカウントを用いる。このアカウントの 1-hop 先のフレンド最大 20 人と 2-hop 先のフレンド最大 20 人を取得し、全ユーザについてツイート情報を 200 件ずつ取得する。対象とするトレンドのランキングは日本のものとし、上位 50 位までを対象とする。

5.2. 実験手順

提案手法によって並べ替えられたトレンドにセレンディピティがあるかを評価する。まず、事前に既存のトレンドのランキングについて、著者自身の主観によってセレンディピティが高い順になるように並べ替える。次に、同じトレンドのランキングについて、提案手法によってセレンディピティが高い順になるように並べ替える。これらの 2 種類の順位の数値をもとに、スピアマンの順位相関係数、平均適合率という 2 種類の評価指標を求め、相関を評価する。

著者自身の主観によってセレンディピティを考慮してトレンドの並べ替えを行うために Ge や Herlocker による定義を参考にしてセレンディピティを決定するための指標を設定する。まず各トレンドについて、以下の指標 A と B をそれぞれ 0 点から 4 点の 5 段階で評価する。

- A. そのトレンドが自分にとって意外性のあるものであるかどうか。
- B. そのトレンドが自分にとって興味のあるものであるかどうか。

各項目について、当てはまるならば 4 点、どちらかと言えば当てはまるならば 3 点、どちらとも言えないならば 2 点、どちらかと言えば当てはまらないならば 1 点、当てはまらないならば 0 点とする。

各トレンドについて、これらの点数を合計し、点数の合計のより高いものが上位となるように並べ替えを行う。点数の合計が同じになった場合の並べ替えの仕方に応じて、以下の3種類のランキングを作る。

- (1) 並べ替え前のトレンドのランキングでより下位だったものほど並べ替え後にはより上位となるように並べ替える。
- (2) 指標 A がより大きいものほど上位となるように並べ替える。指標 A が同じ場合は並べ替え前のトレンドのランキングでより下位だったものほど並べ替え後にはより上位となるように並べ替える。
- (3) 指標 B がより大きいものほど上位となるように並べ替える。指標 B が同じ場合は並べ替え前のトレンドのランキングでより下位だったものほど並べ替え後にはより上位となるように並べ替える。

ランキング(1)は指標 A と B を同程度に重視した並べ替え方、ランキング(2)は指標 A を重視した並べ替え方、ランキング(3)は指標 B を重視した並べ替え方である。

5.3. 評価指標

スピアマンの順位相関係数 ρ は以下によって定義される。

$$\rho = 1 - \frac{6 \sum D^2}{N^3 - N}$$

ここで、 D は対応するトレンドの順位の違い、 N は取得したトレンドの総数である。 ρ の値が 1 に近いほど 2 種類の順位の間にはより強い相関があると考えられる。

ある順位 T 位までの平均適合率は、25 位よりも上位のトレンドを正解、それ以外のものを不正解とし、 T 位より上位のすべての正解の順位について、その順位までの適合率を計算して平均値をとることで計算される。平均適合率が 1 に近いほど、セレンディピティの評価値が高いトレンドが並べ替え後のランキングで上位に固まっていると考えられる。

5.4. 実験結果

提案手法によって並べ替えたトレンドのランキングと、著者自身の主観によってセレンディピティが高い順になるように並べ替えた 3 種類のトレンドのランキング(1), (2), (3) とのスピアマンの順位相関係数の 3 日間分の計算結果を表 1 に示す。結果から、主観によって並べ替えられた 3 種類のランキングの中ではランキング(3)が最も相関係数の値が高くなる傾向があることが分かる。3 日間のトレンドのランキングについて主観評価によって並べ替えられたランキング(3)と提案手法によって並べ替えられたランキングの関係を図 2 に示す。

表 1 3 日間のランキングの相関係数

	ランキ ング(1)	ランキ ング(2)	ランキ ング(3)
2021 年 12 月 23 日 21 時	0.23	0.16	0.27
2021 年 12 月 27 日 13 時	0.49	0.44	0.54
2021 年 12 月 28 日 17 時	0.49	0.43	0.52

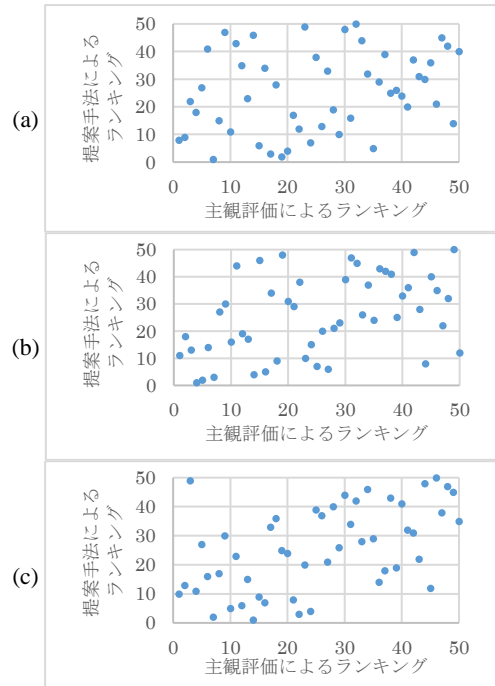


図 2 (a) 2021 年 12 月 23 日 21 時, (b) 12 月 27 日 13 時, (c) 12 月 28 日 17 時のランキングの関係

提案手法によって並べ替えたトレンドのランキングと、著者自身の主観評価によって並べ替えられたトレンドのランキング(3)について、 T 位までの平均適合率の 3 日間分の計算結果を表 2 に示す ($T = 10, 20, 25$ とする)。

表 2 3 日間のランキングの平均適合率

	$T = 10$	$T = 20$	$T = 25$
2021 年 12 月 23 日 21 時	0.93	0.87	0.82
2021 年 12 月 27 日 13 時	0.93	0.87	0.87
2021 年 12 月 28 日 17 時	1	0.96	0.93

6. 議論

6.1. 実験結果の考察

3 日間のトレンドのランキングのうち最もスピアマンの順位相関係数の値が高かったのは、2021 年 12 月 27 日 13 時のトレンドについての提案手法によるランキングと主観によるランキング(3)の組である。一方で最も値が低かったのは、2021 年 12 月 23 日 21 時のトレンドについての提案手法によるランキングと主観によるランキング(2)の組である。また、どの日においても、主観評価によるランキングと、トレンドがユーザにとってどの程度興味のあるものであるかを重視したランキング(3)との相関が最も高く、トレンドがどの程度意外性のあるものであるかを重視したランキング(2)との相関が最も低かった。これらから、提案手法によるランキングのセレンディピティは、ユーザにとってそのトレンドにどの程度意外性があるかよりも、どの程度興味があるかを重視した指標であると考えられる。

2021年12月23日21時の実験結果では、ランキング(1), (2), (3)のいずれの場合においても相関が低かった。これは、並べ替え前のランキングにおいて自分にとってセレンディピティのあるトピックが少なかったため、主観評価によるランキングの並べ替えが上手くできなかったためと考えられる。実際、並べ替え前の上位50位のうち指標AとBの合計が5以上となったトレンドは6件だけであった。

2021年12月27日13時と2021年12月28日17時の実験結果では、2021年12月23日21時の実験結果と比べて相関が全体的に高かった。並べ替え前の上位50位のうち指標AとBの合計が5以上となったトレンドの件数はそれぞれ16件と18件であり、主観評価による並べ替えが行いやすかったことが、相関係数の値の上昇に繋がったと考えられる。

2021年12月23日21時の実験結果について具体的に考察する。最も並べ替えが上手くいったトレンドは、並べ替え前の順位が22位の「無料60連」というトレンドであり、提案手法による並べ替えでは10位、主観評価による並べ替えでは11位となった。最も並べ替えが上手くいかなかったトレンドは、並べ替え前の順位が46位の「基本料金」というトレンドであり、提案手法による並べ替えでは47位、主観評価による並べ替えでは9位となった。前者は著者自身も興味のあるスマホゲームに関するトレンドで、安部らの手法による推薦ユーザの中にそのゲームのプレイヤーが多く含まれていたことから上位となったと考えられる。後者は携帯電話会社の料金プランに関する話題から生じたトレンドと考えられるが、トレンドの発生から時間が経っていたためトレンドに便乗する無関係なツイートが多く収集されてしまっていたことが原因と推測される。

3日間分の平均適合率はいずれも高い値を示している。よって著者自身による主観評価に基づいて並べ替えられたトレンドのランキング(3)において上位だったトレンドは、提案手法によって並べ替えたトレンドのランキングにおいても上位に固まって並べ替えられる傾向が高いといえる。

6.2. 課題

提案手法には比較対象となるような手法がまだないため、その性能を評価するためには主観によって並べ替えられたトレンドのランキングと提案手法によって並べ替えられたトレンドのランキングを比較する必要がある。そのため、提案手法の性能評価は主観評価に基づいた並べ替えの精度に大きく依存すると考えられ、主観評価による並べ替えの精度を上げることが求められる。しかし、実験により、主観評価による並べ替えには様々な問題点があり、十分な精度を保証することは困難であることが示された。具体的には、並べ替え前のトレンドのランキングに含まれるセレンディピティのあるトレンドの数、トレンドが発生した時間や現在の注目度、ワードの種類などが問題点である。これらの問題を解決するために、実験方法を改善する必要がある。具体的には、主観評価によってトレンドのランキングを並べ替えるためのセレンディピティの指標について、ユーザにとっての意外性

と興味の2種類からさらに増やすことや、各指標に対して割り与えられている点数を5段階よりもさらに細かく分けること、取得するツイートを時間やリツイート数などでフィルタリングすることが考えられる。

本研究においては被験者が著者自身の1人のみであったところを著者以外の複数人にする必要もある。ただし、セレンディピティは主観的な評価指標であるため、その感じ方には個人差があると考えられる。主観評価によるトレンドのランキングの並べ替えにおける指標A, Bのどちらを重視する傾向にあるかによって、提案手法の性能評価に差が生じる可能性があるため、被験者ごとに各指標の重みを設定することも検討する。また、安部らの手法の適用に関して、本研究でのフレンドの探索人数は最大で20までとしたが、被験者のフォロー状況によっては探索人数が不十分であり推薦されたユーザのうちセレンディピティのあるユーザの割合が低くなることも予想される。そのため、各被験者による安部らの手法の精度を検討する必要がある。

7. おわりに

安部らのユーザ推薦手法によって推薦されたユーザにセレンディピティがあることを利用し、セレンディピティを考慮して既存のトレンドのランキングを並べ替えて提示する手法を提案した。提案手法の性能については、提案手法によるランキングと主観評価によるランキングとの相関を2種類の指標によって評価した。2つの評価から、提案手法によるランキングのセレンディピティはユーザにとっての意外性よりも興味を重視した指標であり、提案手法は順位の正確さは保証できないものの、セレンディピティの高さを考慮してトレンドのランキングを大まかに並べ替えることができることを示した。また、性能評価における様々な困難性を考察した。この問題を解決するためには、セレンディピティの主観評価の基準をより細かくしたり、被験者の人数を拡大したりする必要がある。

文 献

- [1] 安部高城, 佐藤哲司, "ソーシャルメディアにおけるセレンディピティを考慮したユーザ推薦手法の提案," DEIM Forum, no. B2-6, pp. 1-6, 2015.
- [2] M. Ge and C. Delgado, "Beyond Accuracy: Evaluating Recommender Systems by Coverage and Serendipity," *Proceedings of the ACM Conference on Recommender Systems*, pp. 257-260, 2010.
- [3] J. L. Herlocker, J. A. Konstan, L. G. Terveen and J. T. Riedl, "Evaluating Collaborative Filtering Recommender Systems," *ACM Transactions on Information Systems*, vol. 22, no. 1, pp. 5-53, 2004.
- [4] D. M. Blei, A. Y. Ng and M. I. Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993-1022, 2003.