

マイクロブログにおけるフォロー傾向に基づくユーザ分析

11371012 猪野 稚香子 (灘本研究室)

あらまし: マイクロブログサービスを代表する Twitter のフォローユーザの推薦 において, 既存のフォロイーの傾向によりフォロイーの推薦手法が異なることが判明している. そのため, フォロイー傾向の自動判定が必要である. そこで本研究では, フォロイー傾向の自動判定のはじめの一歩として, フォロイーの分類を行う.

1. はじめに

マイクロブログサービスを代表する Twitter は興味のあるユーザをフォローすることで, そのユーザを自分のタイムラインに表示し, そのユーザのツイートを観覧する機能がある. このフォローしたユーザを本論文では, フォロイーと呼ぶ. ユーザ毎にフォロイーの傾向は様々である. 例えば, あるユーザは同じ興味を持つユーザをフォローしたり, 他のユーザはリアルな友人をフォローしたりと様々である. 山本ら^[1]の先行研究より, フォロイー推薦において, 既存フォロイーの傾向により推薦手法が異なることが判明している. その為, フォロイー傾向の自動判定が必要である. そこで本研究では, フォロイー傾向の自動判定のはじめの一歩として, フォロイーの分類を行い, 該当ユーザがどのようなユーザをフォローしているのかを分類する. 具体的には, ユーザがどのようなフォロイーをフォローしているのかを表す関係をフォロー関係タイプ, ユーザがフォローしているフォロイー自身のタイプをフォロイータイプとし, 分析を行う.

2. フォロー関係の分類

2. 1. フォロー関係のタイプ

本研究では, 図 1 のように該当ユーザのフォロー関係をトピック指向と人指向の 2 つに分類する. まず, ツイートの話題の類似性を考慮してフォローを行うユーザをトピック指向に分類する. さらにトピック指向に分類されるユーザのなかでも, 話題自体を重視しフォローを行うフォロイーを話題重視のフォロイー, その話題に加え, その話題に対する感情を重視するフォロイーを話題と印象重視のフォロイーと分類する. また, 友人・知人等を良くフォローする等, ユーザ同士の繋がりを考慮しフォローを行うトピック指向でないフォロイーを人指向と分類する.

2. 2. フォロー関係の分類手順

フォロー関係の分類手順を以下に示す.

- (1) フォロイーの情報を取得する.
- (2) フォロイーのプロフィールページの description に該当ワードが存在すれば, そのフォロイーは企業アカウントであるとする. 該当ワードとは, 公式アカウント・公式 Twitter・株式会社・企業・事業部・公式サイト・最新情報・CD. LTD. ・カ

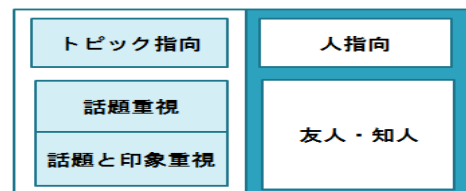


図 1: フォロー関係

スタマーサポートセンターである.

- (3) フォロイーの name, もしくは screen_name に bot という単語が存在すれば, そのフォロイーは bot であるとする.
- (4) フォロイーのフォロワーが閾値(5000 人)を超え, かつフォロイーのフォロイーが閾値(539 人)以下のとき, そのフォロイーは有名人であるとする.
- (5) Twitter API を用い, 該当ユーザとフォロイーのスクリーンネームを入力し, 両者のツイートを最新 200 件ずつ取得し, 統合する.
- (6) (5)で統合したツイートに対してクラスタリングを行い, トピックを 10 件抽出する. クラスタリングには Repeated Bisection^[2]を用いる.
- (7) 抽出したトピックのうち, 該当ユーザとフォロイー間で過半数を超えるトピックが同じの時, そのフォロイーはトピック指向のフォロイーとする.
- (8) トピック指向のユーザが話題を重視しているのか, 印象を重視しているのかを分類するため, (6)で抽出されたトピック毎に該当ユーザとフォロイーの感情ベクトルを算出する. そして, 該当ユーザとフォロイーの感情の類似性をコサイン類似度を用いて求める. この結果, 類似性が閾値 0.82 を超えるフォロイーをトピック指向印象重視ユーザと分類する. それ以外のユーザはトピック指向話題重視ユーザと分類する.
- (9) (8)までに分類されなかったユーザを人指向ユーザと分類する.

3. フォロイータイプの分類

3. 1. フォロイータイプ

フォロイーが Twitter を利用する目的にはそれぞれ特徴があると考え, 以下の 3 つのタイプに分類する. 主に情報発信を行うために Twitter を利用する情報発信型ユーザ, 主に情報収集を行うために

Twitter を利用する情報収集型ユーザ、ユーザ同士のコミュニケーションを目的としたコミュニケーション型ユーザである。企業のアカウントや有名人は情報発信型ユーザに分類される。

3. 2. フォロイータイプのカテゴリ分け

フォロイータイプのカテゴリ分けを以下に示す。

- (1) 情報発信型は、閲覧者数、フォロイー数が多いと考えられる。そこで、閲覧ユーザはフォロワー数がフォロイー数より多く、RT 数が少ないユーザを情報発信ユーザに分類する。
- (2) 情報発信型とコミュニケーション型以外のユーザは情報収集型ユーザとする。
- (3) コミュニケーション型は、相互フォロー数がフォロワー数の 8 割を超えており、且つリプライがツイートに対し 50%のユーザをコミュニケーション型とする。

4. 分析及び考察

実際に Twitter を利用している 2 名のユーザの分析を行った。

4. 1. データセット

ユーザ A はツイート数 1285、フォロワー数 86 人、フォロワー数 61 人、うち相互フォロー数 55 人のユーザである。直近 200 ツイートの RT の割合は 0.02、リプライの割合は 0.45 である。プロフィールには好きなクイズゲームのためのアカウントだと記載している。

ユーザ B はツイート数 1285、フォロワー数 142 人、フォロワー数 76 人、うち相互フォロー数 49 人のユーザである。直近 200 ツイートの RT の割合は 0.168、同じくリプライの割合は 0.402 である。プロフィールにオートバイや野球チームが好きであると記載している。

4. 2. ユーザ A の結果と考察

ユーザ A のフォロイーは図 2(1)より、話題指向ユーザが全体の 87%であった。全体に対するその内訳は、話題重視が 54%、話題と感情重視が 33%であった。これにより、ユーザ A は感情よりも話題そのものに重きを置く、トピック指向話題重視のユーザであると考えられる。そして、実際にフォロイーを確認すると該当のクイズゲームが好きなユーザに加え、トピック指向のユーザであり、情報発信を多く行う bot や企業アカウントを多くフォローしている。さらに、図 2(2)より、ユーザ A が Twitter を該当のクイズゲームの情報収集のために利用しているからだと考えられる。以上より、ユーザ A はトピック指向話題重視型のフォロイー傾向を持つ、情報収集を目的に Twitter を利用するユーザであると考えられる。

4. 3. ユーザ B の結果と考察

ユーザ B のフォロイーは図 2(3)より、話題指向が 55%、人指向が 45%と同程度である。これは、ユーザ B が複数の趣味を持ち、様々なアカウントをフォローしているからではないかと考えられる。実際にユーザ

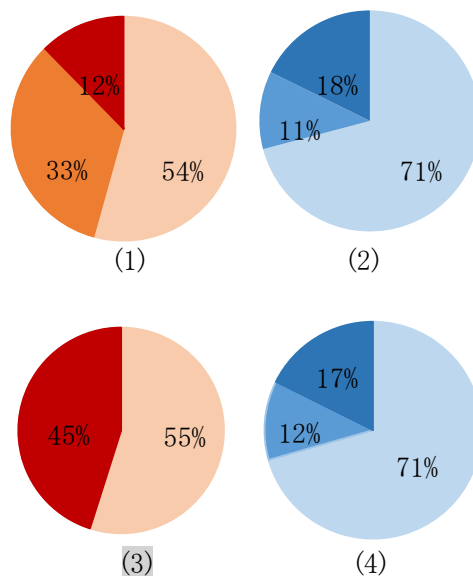


図 2: ユーザのフォロイー関係とフォロイータイプ

のフォロイーを見てみると、元野球解説者やキャスターなどが存在しており、様々なアカウントをフォローしているのがわかる。また、図 2(4)より、フォロイーの 71%が情報発信型であり、このうち 14%が企業アカウント、41%が有名人であった。以上より、ユーザ B はトピック指向且話題重視型のフォロイー傾向を持つ、情報収集を目的に Twitter を利用するユーザであると考えられる。

5. まとめと今後の課題

本研究では Twitter ユーザのフォロイーのフォロイー関係の傾向と、フォロイータイプのカテゴリ分けを行った。フォロイー関係は大きくトピック指向と人指向の 2 タイプに、また、トピック指向は話題重視、話題と印象重視に分類した。また、フォロイータイプは情報発信型、情報収集型、コミュニケーション型の 3 タイプに分類した。今後の課題はユーザ実験の回数を増やし、正解データの作成と検証を行うことである。

参考文献

- [1] 山本湧輝, 熊本忠彦, 灘本明代, ”話題と感情に基づくフォロイー推薦と評価,” 第 8 回データ工学と情報マネジメントに関するフォーラム(DEIM2016), B4-1, 8 pages, 2016.
- [2] Ying Zhao and George Karypis. Comparison of agglomerative and partitional document clustering algorithms. Technical report, Department of Computer Science, University of Minnesota, Minneapolis, MN 55455, 2002.