

レビュー, ブログ, マイクロブログサイトの 高次数ユーザの自己中心トライアド分析

Egocentric triad analysis of high degree users on review, blog and microblog sites

西可南子[†]
Kanakano Nishi

加藤翔子[†]
Shoko Kato

大久保誠也[†]
Seiya Okubo

齊藤和巳[†]
Kazumi Saito

1. はじめに

ソーシャルメディア上には, 多様なソーシャルネットワークが構築されている. このようなネットワークの構造分析を目的として, 小出らは [1, 2], 高次数ユーザの隣接関係の類似度をデンドログラムなどで分析した.

これに対し本研究では, 小出らと同様に, レビュー, ブログ, マイクロブログの各サイトの高次数ユーザに着目するが, これらユーザの自己中心ネットワークのトライアド分布により各サイトの特徴を分析する. 詳細には, 中心ユーザと入出力リンク関係を持つユーザ群から自己中心ネットワークを構築して各ユーザのトライアド分布を求め, これら分布間の類似度に基づく k -近傍グラフを可視化して評価する. レビュー, ブログ, マイクロブログの各サイトとして @cosme, Ameblo, Twitter を用いた実験結果では, @cosme と Ameblo には, 高次数のみで特徴付けられ, 読み手側に限定されたユーザが多数存在するのに対して, Twitter にはこのようなユーザは存在せず, 相互結合が特徴的なユーザが支配的となったことを示す.

2. 分析手法

2.1. トライアド

トライアドとは, 有向グラフにおいて, 任意に選択した連結3ノードの接続パターンを表すもので [3], 図1に示すように, 可能なパターン数は13となる. 本研究では, 3.1節で説明する3つのネットワークにおいて, 次数が上位 M 位以内のノードとその隣接ノードのトライアドを分析する.

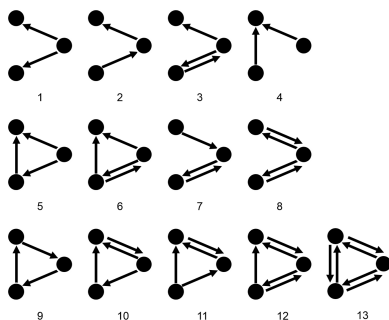


図1: トライアドの種類

2.2. k -近傍法

k -近傍法は, パターン認識の分野でよく用いられる単純な分類アルゴリズムである. 本稿では, 抽出した

コア構造をノードとし, トライアドの頻度ベクトルのコサイン類似度が高い関係のあるコア間のみリンクを張ることでネットワークを作成し, バネモデルを用いて可視化する. ネットワーク作成の詳細な手順を以下に示す.

1. 全ノードペア間のコサイン類似度を計算する
2. 各ノードに対して, 自身との類似度で降順に相手ノードを並び替える
3. $k \leftarrow 1$ とする
4. 各ノードは, 自身との類似度が k 位タイのノード全てとリンクを結ぶ
5. 全ノードが単一連結成分になれば終了する. そうでなければ $k \leftarrow k + 1$ とし, 4に戻る

2.3. 提案分析法

ネットワークを $G = (V, E)$ とし, 分析対象とする高次数ノードの数を M とすれば, 提案分析法の処理手順は以下となる.

1. $m = 1$ から M まで以下の処理を繰り返す
 - (a) 第 m 位の高次数ノードと入出力リンク関係を持つノード集合を $V_m \subset V$ とし, V_m 間の全リンク集合 E_m を求め, 自己中心ネットワークを $G_m = (V, E_m)$ として構築する.
 - (b) ネットワーク G_m に出現するトライアドのパターンを数え上げ, 13次元のトライアド頻度ベクトル \mathbf{x}_m を構築する.
2. トライアド頻度ベクトルの集合 $\{\mathbf{x}_1, \dots, \mathbf{x}_M\}$ から k -近傍ネットワークを構築し, このネットワークをバネモデルで可視化する.

ここで各ノード m に対し, 高頻出トライアドパターンを $\arg \max_{1 \leq i \leq 13} \{x_{m,i}\}$ として求め, そのパターンに応じてマーカーと色を指定して可視化する.

3. 評価実験

3.1. データ概要

一つ目は, 化粧品の口コミサイト “@cosme”^{*} から2009年12月に取得したお気に入りネットワークである(以下, cosmeNW). サイトに登録しているあるユーザ X がユーザ Y をお気に入り登録をすると, ユーザ X からユーザ Y に対してリンクが張られる. このネットワークはランダムに選択したユーザから10段まで辿

[†]静岡県立大学

^{*}<http://www.cosme.net/>

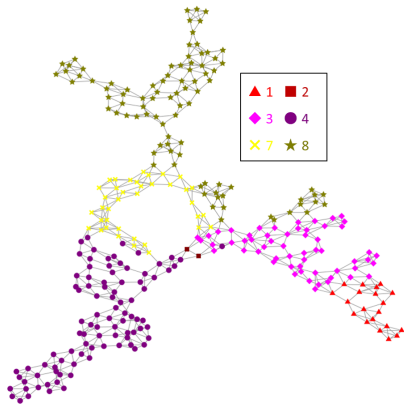


図 2: cosmeNW 可視化結果

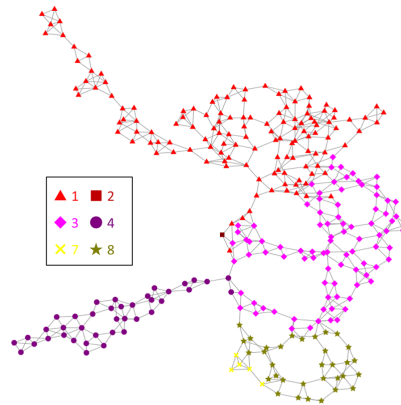


図 3: amebloNW 可視化結果

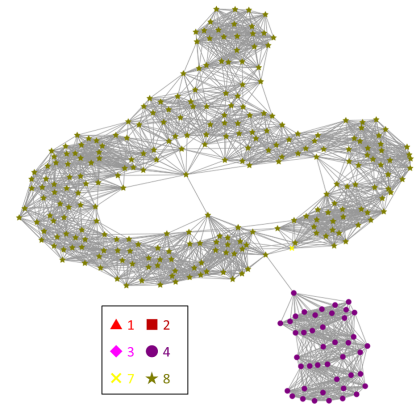


図 4: FollowNW 可視化結果

て収集したもので、ノード数は 45,024 ノードでリンク数は 546,930 本である。

二つ目は、日本のブログサービスサイト“Ameba”[†]の読者ネットワークである [15] (以下, amebloNW)。“Ameba”の各ブログには読者のブログへのハイパーリンクが張ってある読者リストがある。ここで、ブログ X からブログ Y への読者リンクは、ブログ X がブログ Y をお気に入りブログとして登録した時に生成される。このネットワークのノード数は 56,604 ノード、リンク数は 1,071,080 本である。

三つ目は、Twitter 検索[‡]で収集された日本語で書かれたツイートを開示しているユーザのフォローネットワークである。ここで、ユーザ X からユーザ Y へのリンクは、ユーザ X がユーザ Y をフォローしたときに生成される。本稿では、2011 年 1 月 31 日時点において累積 200 回以上ツイートしているユーザのフォロー関係から、ユーザをノード、フォロー関係をリンクとした有向ネットワークを作成した。ネットワークのノード数は 1,079,986 ノード、リンク数は 157,371,341 本である (以下 FollowNW)。さらに、発言内容の判定に用いたツイートデータは、2011 年 3 月 7 日 0:00:00 から 11 日 12:00:00 までの期間に、ネットワーク内に存在する 1,079,986 ノードのツイートを収集したもので、この期間内の総ツイートは 57,625,889 ツイートであった。

3.2. 分析結果

分析対象ネットワークでは、ユーザ X がお気に入り、読者登録、フォローにより、ユーザ Y にリンクを張るので、ユーザ Y は書き手で、ユーザ X は読み手となる。よって、赤三角のトライアド 1 が多いノードは、一方的な読み手ユーザで、逆に紫丸のトライアド 4 だと、一方的な書き手ユーザになることを意味する。一方、オリブ星のトライアド 8 なら、読み手かつ書き手ユーザであり、ピンク菱形のトライアド 3 は、トライアド 1 と 8 の中間、黄クロス of トライアド 7 は、トライアド 4 と 8 の中間などとなる。

トライアド 1 の主として読み手ノードについては、ある程度数 cosmeNW や amebloNW で存在するが、fol-

lowNW ではほとんど存在しない。すなわち、@cosme と Ameblo には読者のみの高次数ユーザが存在するのが特徴となる。なお、トライアド 4 の主として書き手ノードについては、どのネットワークでもある程度存在する。一方、トライアド 8 は、FollowNW で大多数を占め、Twitter での特徴となっている。すなわち、@cosme と Ameblo には、高出次数のみで特徴付けられ、読み手側に限定されたユーザが多数存在するのに対して、Twitter にはこのようなユーザは存在せず、相互結合が特徴的なユーザが支配的となったことがわかった。

4. おわりに

本研究では、レビュー、ブログ、マイクロブログの各サイトの高次数ユーザに着目するが、これらユーザの自己中心ネットワークのトライアド分布により各サイトの特徴を分析した。その結果、@cosme と Ameblo には、高出次数のみで特徴付けられ、読み手側に限定されたユーザが多数存在するのに対して、Twitter にはこのようなユーザは存在せず、相互結合が特徴的なユーザが支配的となったことを確認した。

今後は、本分析で得られた知見を用いて、さらに大規模なソーシャルメディアデータや自己中心ネットワークのトライアド分布の分析法を評価していく。

謝辞

本研究は、科研費 (C)(No.23500312) の助成を受けた。

参考文献

- [1] 小出明弘, 齊藤和巳, 風間一洋, 鳥海不二夫, “ネットワーク分析による Twitter ユーザのフォロー形成に関する一考察,” 情報処理学会論文誌, 数理モデル化と応用-ネットワークが創発する知能論文特集-, Vol.6, No.2, pp.164-173, 2013.
- [2] 小出明弘, 齊藤和巳, “マイクロブログ, ウェブログ, レビューサイトのネットワーク構造比較,” 第 97 回 知識ベースシステム研究会 (SIG-KBS), 2012.
- [3] S.Wasserman, K.Faust, “Social Network Analysis,” pp566, 1994.

[†]<http://ameblo.jp/>

[‡]<http://yats-data.com/yats/>