

# インターネットコミュニティにおけるユーザの 振る舞いに及ぼす社会的関係と行動類似性の影響

津田洋平\*1

熊野雅仁\*2

木村昌弘\*2

Yohei Tsuda, Masahito Kumano, and Masahiro Kimura

## 1. はじめに

インターネットコミュニティにおいてユーザの将来行動の予測を可能とする手法の確立は、次世代の高度推薦システムで期待されるコア技術の一つである。その有効な予測法の構築を目指す上で、本研究では、社会学で議論されている「社会的関係」と「行動類似性」のどちらが、ユーザの将来行動に強い影響を及ぼすかについて調べる。ここに、社会的関係とは、現実世界での友達同士の関係を指し、行動類似性とは、類似的な行動をする人同士の関係を指す。最近、CrandallらはWikipediaとLiveJournalを用いて影響分析を行い、その影響の違いを示す結果を得ている[1]。本研究では、高度推薦システムの実現に向けた基礎データとして、長期間及び短期間の予測結果から、オンラインコミュニティユーザの振る舞いに、社会的関係と行動類似性がどのように影響するかについて分析する。

## 2. 影響分析法

オンラインコミュニティ上での社会的関係は、現実世界と同様に、友達同士のユーザ関係を示すものと考えられる。そこで、オンラインコミュニティの機能を用いてユーザ同士が繋がったネットワークを社会的ネットワークと見なす。

一方、行動類似性は、ユーザが映画を視聴する、商品を買うといった振る舞いが類似しているユーザの関係を示すものと考え、各ユーザに対して、その振る舞いが近いものから上位  $n$  人を繋げたネットワークを行動類似ネットワークとする。 $n$ とは、社会的ネットワークでの各ユーザの次数と同値である。また、行動類似性の算出には、類似度を用いる。その際、ユーザ  $i$  または  $j$  が見た映画もしくは買った商品を要素とするベクトル  $w$  を  $w_i, w_j$  とし、また、 $w_i$  を  $w_i$  の 0 でない要素の集合、 $w_j$  を  $w_j$  の 0 でない要素の集合として、式(1)、または式(2)により類似度を算出する。

$$\cos(i, j) = \frac{w_i \cdot w_j}{|w_i| \cdot |w_j|} \quad (1)$$

$$Jaccard(i, j) = \frac{|w_i \cap w_j|}{|w_i \cup w_j|} \quad (2)$$

社会的ネットワークにおいて、学習期間に隣人  $k$  がある振る舞いを行った場合、その後の評価期間に個人  $u$  が特定の振る舞いを採用する確率  $P(k)$  を測定する。

一方、行動類似ネットワークでは、評価期間に個人  $u$  と個人  $k$  が同様な振る舞いを行った場合、その後の評価期間に個人  $u$  が特定の振る舞いを採用する確率  $P(k)$  を測定し、社会的関係と行動類似性の影響度の違いを調査する。

## 3. 分析データ

本研究では、URL を入力すると、それを起点としてネットワークリンクを幅優先探索で辿り、分析データを構築するソフトウェアを開発し、登録ユーザからランダムに選んだ 3 人を起点としてデータを収集した。その際、99.5%が同一 URL で構築されていたため、その一つの近傍ネットワークを用いて分析を行った。

### 3.1 映画データ

映画データとして、映画情報サイト yahoo 映画を用いる。yahoo 映画では、ユーザが映画に対するレビューを記述する機能や、ユーザ同士を繋ぐファン機能が存在する。その機能に着目し、レビュー機能から行動類似ネットワーク、ファン機能から社会ネットワークを構築した。2008年6月から2009年6月までの期間を取得し、映画総数 14438 作品、ユーザ数 5524 人となった。

### 3.2 購買データ

購買データとして、口コミサイト@cosme を用いる。@cosme では、ユーザが商品に対する口コミを記述する機能や、ユーザ同士を繋ぐお気に入り機能が存在する。その際、口コミは商品とそのブランドに対して記述される。そこで、口コミそれぞれに対して行動類似ネットワークを構築し、ファン機能から社会ネットワークを構築した。2008年12月から2009年12月までの期間を取得し、商品総数 48548 品、その際のブランド数 7139 社、ユーザ数 45024 人となった。

\*1 龍谷大学 理工学部 電子情報学科

\*2 龍谷大学大学院 理工学研究科 電子情報学専攻

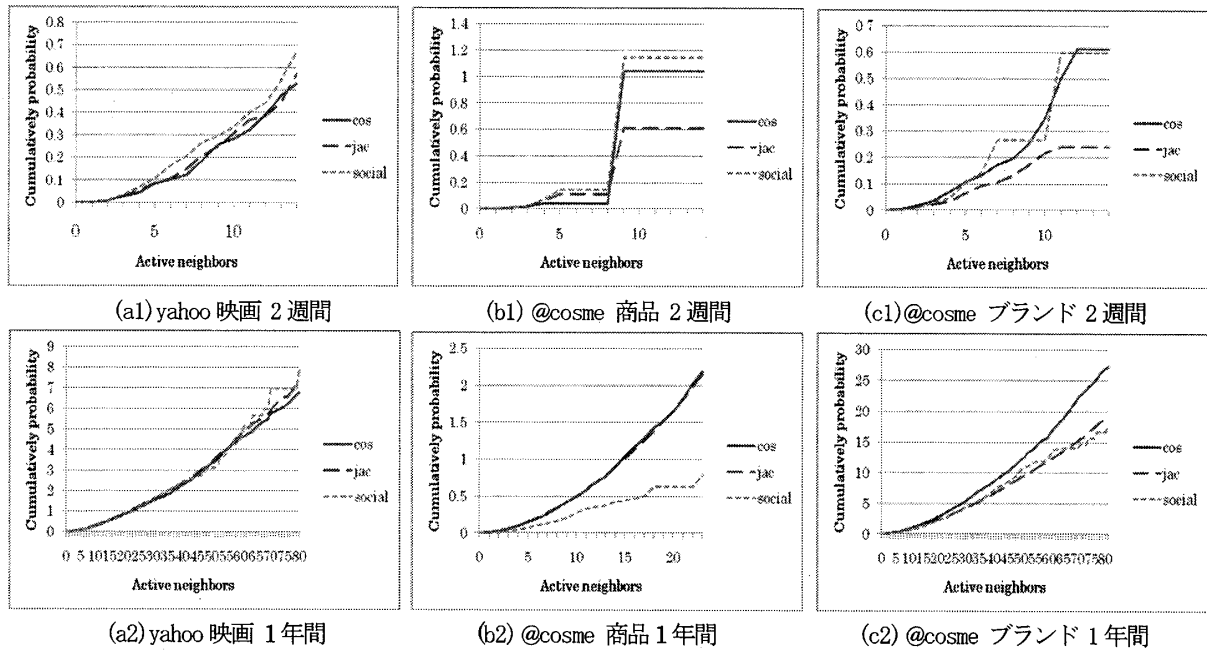


図1 社会的ネットワーク (social) と行動類似ネットワーク (cos, jac) の影響度比較

#### 4. 分析結果

前述した社会的ネットワークと行動類似ネットワークに対して、短期間及び長期間での影響度を比較した。図 a1~c2 の横軸は個人  $u$  に影響を与えた影響寄与者の数、縦軸は隣人の振る舞いが採用された確率である。

図1の(a1)はyahoo映画において1週間を学習期間、残りの1週間を評価期間とした短期間の結果であり、(a2)は6カ月を学習期間、残りの6カ月を評価期間とした長期間の結果である。短期間では、社会的ネットワークが行動類似ネットワークより高い値を示している。長期間では、影響寄与者が少ないとき Jaccard 類似度の行動類似ネットワークが高くなったが、影響寄与者が増えるに従い社会的ネットワークが優位になる傾向があった。cos 類似度と Jaccard 類似度を比較すると、Jaccard 類似度の方が高い値を示した。また、社会的ネットワークの次数分布は power law であった。クラスター係数は 0.116 で、平均ノード間距離は 3.38 であった。

@cosme において、図1(b1), (b2)は、商品の購買行動、(c1), (c2)はブランド嗜好の購買行動に基づく行動類似ネットワークであり、期間は映画データと同一である。短期間では共に行動類似ネットワークが高い値を示した。商品の短期間では、社会的ネットワークと、cos 類似度の行動類似ネットワークが高い値となった。ブランド嗜好の短期間では、cos 類似度の行動類似ネットワークのみが高い値となった。cos 類似度と Jaccard 類似度を比較すると、cos 類似度が高い値を示した。また、社会的ネットワークのクラスター係数は 0.0396 で、平均ノード間距離は 4.17 となった。以上の結果から、

短期間では、yahoo映画と@cosme共に社会的ネットワークが高い結果になった。しかし、長期間では、@cosmeの社会的ネットワークのみが落ち込む傾向にあった。行動類似ネットワークを見ると、yahoo映画は jaccard 類似度が、@cosmeは cos 類似度が高くなった。yahoo映画では、1作品に対して2度以上見ることが少ないことから Jaccard 類似度が、@cosmeでは、1商品に対して何度も購入することから cos 類似度の影響力が強くなったと考えられる。@cosmeでの商品とブランドでの影響度は、ほとんど変わらない傾向が見られた。

#### 5. まとめ

本稿では、オンラインコミュニティのユーザの振る舞いに社会的関係と行動類似性がどのように影響するかについて分析した。映画データ、購買データ共に短期間では社会的関係の影響度が強かった。長期間になると、映画データでは社会的関係と行動類似性共に似た影響度を示したが、購買データでは社会的関係が落ち込む傾向が見られた。また、映画データでは Jaccard 類似度を用いた行動類似性が高く、購買データでは cos 類似度を用いた行動類似維持性が高くなった。これらを考慮し、高度な推薦システムを確立に適用することが、今後の課題である。

#### 参考文献

- [1] David Crandall, Dan Cosley, Daniel Huttenlocher, Jon Kleinberg, Siddharth Suri. Feedback Effects between Similarity and Social Influence, KDD 2008.