

D-043

ソーシャルブックマークにおける方向性を持った推薦の検討

A Webpage Recommendation by Specifying A Target User on Social Bookmark Service

中根 慶恵† 山田 剛一† 絹川 博之†

Yoshie Nakane Koichi Yamada Hiroshi Kinukawa

1 はじめに

本研究では、ソーシャルブックマークを用いたWebページ推薦システムの構築を目指している。ユーザの過去のブックマーク履歴などユーザの暗黙の嗜好だけでなく、ユーザが明示的に指定する目標に沿った推薦を行う、方向性を持った推薦を検討している。

今回は、国内シェアの約75%を占めるソーシャルブックマークサービス「はてなブックマーク」[1][2]を行った。

2 方向性を持った推薦

協調フィルタリング[3][4]な推薦では、あるユーザに対して、そのユーザと似た行動パターンや嗜好を持つユーザが重要視される。これに対し、方向性を持った推薦では、ユーザに対して、そのユーザが目標としているユーザとの間に立つユーザ(これをパス上のユーザと呼ぶ)を選択し、そのユーザがブックマークしているWebページを優先して推薦する。

2.1 方向性を持った推薦の必要性

「はてなブックマーク」には、自分がブックマークしているWebページの情報に基づいて、自分と同じWebページを多くブックマークしているユーザを推薦する機能がある。しかし、この推薦機能の場合、今までのブックマークと全く違った方向性のブックマークを行っているユーザや、Webページを推薦することは不可能である。新しく興味を持った分野に関するWebページをブックマークする時等に、目標となるユーザ指定して推薦を受けたいことがある。

2.2 パス上のユーザ

方向性を持った推薦を行う場合、目標となるユーザではなくパス上のユーザの嗜好を優先した推薦をするため、まずパス上のユーザを探す必要がある。今回はパス上のユーザの中でも、自分と共通のWebページをブックマークしている、自分に最も近いパス上ユーザを推薦に用いることを検討する。さらに、そのパス上のユーザが目標のユーザと共通のWebページをブックマークしていることが多いため、図1で表現されている構造を持つ場合を想定している。

3 実現方法の検討

自分が目標とするユーザを指定するWebページ推薦システムの実現方法を検討した。

3.1 パス上のユーザの抽出

ユーザ (自分と同じURLをブックマークしている)

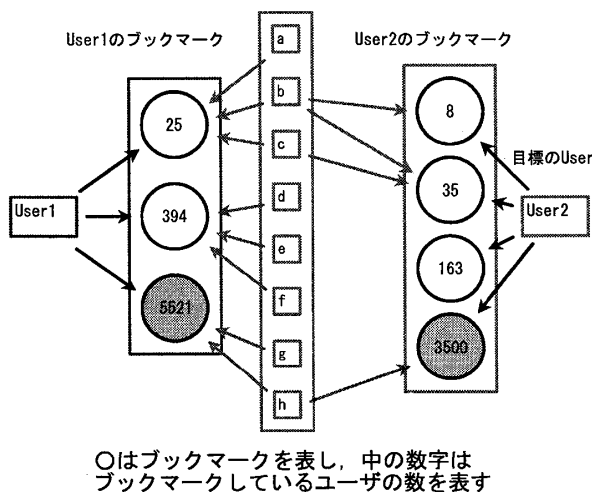


図1. パス上のユーザの位置付け

前述のようにパス上のユーザが自分(User1)と目標ユーザ(User2)の双方と共通のブックマークを持つと仮定する(図1)。

「はてなブックマーク」から配信される、ユーザがブックマークしているエントリー一覧のRSSと、あるURLをブックマークしているユーザー一覧のRSSを用いて、自分がブックマークしているURLと同じURLをブックマークしているユーザの中から、目標のユーザと同じURLをブックマークしているユーザを抽出し、パス上のユーザとする。この段階でパス上のユーザが見つからない場合、本来は再帰的にユーザとそのブックマークを抽出し、目標のユーザのエントリーとの比較を繰り返す必要があるが、今回は再帰的にたどることは行わない。

図1のUser b, c, h がパス上のユーザとなる。

3.2 探索範囲の絞り込み

目標のユーザと自分との橋渡し役となる実質的なパス上のユーザを選択するために、以下に挙げる制約条件を設ける。これにより探索範囲を絞ることにもなる。

- (1) GoogleやYahoo!のトップページのようなメジャーなURLは、ブックマークしていることがその人の特

†東京電機大学大学院情報メディア学専攻
Graduate School of Science and Technology for Future
Life, Tokyo Denki University

徴を表しているとは言い難く、パスとして参考にならないため、あまりにも多くのユーザからブックマークされているURLは探索しない。

図1のUser1のブックマークとUser2のブックマークの1番下のものに関しては、ブックマークしているユーザが非常に多く、メジャーなURLだと考えられるため、その先の探索は行わない。よって、User h はパス上のユーザとして選択されない。

(2) あまり古いウェブページは情報としての価値が低く、また、ユーザの趣向が変化している可能性があるため、ブックマークされた日付が古いエントリの探索はしない。

ブックマークされている数が少ないエントリを介した繋がりであるほど、より特徴的で価値の高い繋がりと考えられるので、パス上のユーザが複数発見された場合、ブックマークされている数が少ないエントリを持つパス上のユーザを優先して選択する。図1のUser b と c であれば、bの方がブックマークしているユーザが少ないエントリで目標のユーザと繋がっているため、User b を優先して選択する。

4 評価実験

パス上のユーザが、実際にパスにふさわしい嗜好を持っているのか確認するための実験を行った。一般に、パス上のユーザは自分と目標のユーザの中間に位置する嗜好を持つことが望ましく、ブックマーク傾向としては、目標のユーザに近い分野のページをある程度ブックマークしていることが期待される。

4.1 手順

- (1) 自分と目標のユーザの、計2名のユーザを選ぶ。
- (2) 自分がブックマークしているエントリのRSSから、同じURLをブックマークしているユーザを抽出する。
- (3) 抽出されたユーザの中で、目標のユーザと同じWebページをブックマークしているユーザを探す(今回は2階層以上ユーザをたどることはしない)。
- (4) パス上のユーザの中でも、ブックマーク数が少ないエントリをパス上に持つユーザを選択し、そのユーザがブックマークしたエントリのカテゴリを確認する。

4.2 結果の取得方法

抽出されたパス上のユーザの中から、1回の実験につき無作為に5名を選ぶ。パス上のユーザのブックマーク新着100件のうち、何件が目標のユーザとのパス上にあるエントリと同じカテゴリに属しているかを見る。

この時、「はてなブックマーク」のユーザはもともと「コンピュータ・IT」関係に強い関心を持っているため、「コンピュータ・IT」カテゴリのブックマークは除外して考察する。

4.3 結果

実験にあたり、ブックマークを公開している、自分と同じエントリをブックマークしていないユーザ4名を無作為に選び、目標のユーザに設定した。目標のユーザのうち

2名は「コンピュータ・IT」カテゴリのWebページを特に多くブックマークしていて、他の2名はそれぞれ「社会」と「スポーツ・音楽・芸能」カテゴリのWebページを多くブックマークしていた。また、自分は「スポーツ・音楽・芸能」カテゴリのWebページを多くブックマークしている。

今回の実験では目標のユーザを4名設定したので、サンプリングを行ったパス上のユーザは計20名である。20名のユーザそれぞれの新着エントリ100件を調べ、何件が目標のユーザとのパス上にあるエントリと同じカテゴリに属しているかを確認した。「コンピュータ・IT」以外のカテゴリで共通するエントリを持つユーザは14名で、パスとして抽出されたカテゴリと同じカテゴリに属するエントリの数は100件のうち平均15.31件あった。

「コンピュータ・IT」以外のカテゴリでは、パスとして抽出されたエントリのカテゴリが、目標のユーザが多くブックマークしているカテゴリと違う事が度々あった。

4.4 考察

パス上のユーザのブックマークが自分にとって参考になるか確認したところ、サンプリングしたパス上のユーザの中でも、パスとして抽出されたURLが少なく、特に分野にこだわり無く最新のニュースや流行の動画を次々とブックマークしているユーザは、パス上のユーザとして選択されても参考になるエントリが少ない事が多かった。逆に、パスとして多数のURLが抽出され、分野を絞ってブックマークしているユーザは、参考になるエントリが多く、目標とするユーザとのパスの役割を十分に果たしているユーザがほとんどであった。

5 おわりに

今回の調査実験から、自分のブックマークしている分野とは異なる分野のブックマークをしているユーザを目標とする際に、その足がかりとなるブックマークを持つユーザの存在を確認することができた。パス上のユーザを用いることが、目標とするユーザを指定する推薦において有効であると期待される。

今後は、パス上のユーザのエントリを優先する仕組みを協調フィルタリングの枠組みに組み入れ、ユーザによる評価実験を行っていきたい。

参考文献

- [1] はてなブックマーク, <http://b.hatena.ne.jp/>
- [2] MarkeZineニュース, 「はてなブックマーク、リニューアルの成果はいかに?」, (2008): <http://markezine.jp/article/detail/6080>
- [3] 石橋 智幸, 顧 優輝, 脇屋 達, 真部 雄介, 菅原 研次, 「ソーシャルブックマークを用いたWebページ推薦システムの開発(「社会的インタラクションにおける知」及び一般)」, 電子情報通信学会技術研究報告, 109(211), pp.7-12 (2009)
- [4] 土方 嘉徳, 「嗜好抽出と情報推薦技術(嗜好抽出・情報推薦の基礎理論, <特集>利用者の好みをとらえ活かす-嗜好抽出技術の最前線-)」, 情報処理 48(9), pp.957-965 (2007)