

Folksonomy マイニングに基づく Web ページ推薦システム

丹羽 智史[†] 土肥 拓生[†] 本位田 真一^{†,‡}

協調フィルタリングを用いた商品推薦システムを応用して Web ページ推薦システムを構築しようとする試みは従来から行われてきたが、十分な量のデータソースを確保することが困難なことや推薦対象である Web ページの数が大きすぎるなどから、その用途は非常に限定されたものだった。本論文では近年急速に普及し始めたソーシャルブックマークと Folksonomy を利用して、インターネット上の Web ページ全体を対象とした Web ページ推薦システムの構築手法を提案する。

Web Page Recommender System Based on Folksonomy Mining

SATOSHI NIWA,[†] TAKUO DOI[†] and SHINICHI HONIDEN^{†,‡}

There have been many attempts to construct web page recommender systems using collaborative filtering, but the domains these systems can cover are very restricted, because it is very difficult to assemble data of user preference about web pages, and the number of web pages on the Internet is too large. In this paper, we propose the way to construct new type of web page recommender system covering all over the Internet, by using folksonomy and social bookmarks which are getting very popular in these days.

1. はじめに

近年のインターネットの爆発的な普及にともない、我々は日々発信される膨大な Web 情報の中から自分にとって必要な情報を取捨選択する作業に多大な労力を強いられている。こうした状況を解決するために、膨大な Web ページ情報の中からユーザの要求にあったものだけを自動的に選択し、推薦してくれる Web ページ推薦システムが数多く開発されてきた。Web ページ推薦システムは用途に合わせて様々な種類のもが開発されているが、いまだ発展途上の分野である。

従来の Web ページ推薦システムの多くはシステムの核となる部分に協調フィルタリングを用いている¹⁾⁻³⁾。協調フィルタリングは主に商品推薦システムを構築する際に用いられる手法で、おおまかに以下の手順からなる⁴⁾。

(1) ユーザの購買履歴をマイニングして、ユーザと商品購買パターンが類似しているユーザ群を抽出する。

(2) ユーザ群に共通して好まれる商品群を抽出し、ユーザに推薦する。

基本的に協調フィルタリングにおいては 2 ユーザ間で共通の商品を購入していないと類似度が 0 と見なされてしまうため、商品数 M に対して十分に大きいテストユーザ数 N の購買履歴を取得しなければならない。 M が N に対して大きすぎる場合、ユーザの嗜好をより抽象化して見かけ上の M を圧縮する必要がある。

従来の多くの Web ページ推薦システムでは「商品購買履歴」を「Web ページアクセス履歴」に置き換えて応用している。しかし Web ページのアクセス履歴はサーバごとに分散しているため、Web ページ推薦システムを構築する際に特定のサイトやサイト群に推薦対象を絞らざるをえなかった。またアクセス履歴以外のデータソースを用いてインターネット全体を対象とした推薦システムを構築する場合も、対象となる Web ページ数 M があまりにも大きいために同一ページを評価しているユーザ群が十分な数集められず、この場合ユーザの Web ページの嗜好をいかにして抽象化して表現するかという問題を解決する必要があった。

本論文では近年急速に普及し始めたソーシャルブックマーク上のデータをユーザの Web ページ嗜好データとして利用することで、従来の Web ページ推薦システムにおける「ユーザのページ嗜好データの不足」を解決し、インターネット上の Web ページ全体を対象とする Web ページ推薦システムを構築する手法を

[†] 東京大学

The University of Tokyo

[‡] 国立情報学研究所

National Institute of Informatics

示した。その際 Folksonomy によって与えられるタグ情報をマイニングすることで「ユーザの Web ページ嗜好の抽象化表現」と「Folksonomy のタグ表記のゆれ」に対する解決アプローチを提案した。

本論文の全体の構成は以下のとおりである。まず 2 章で Folksonomy とソーシャルブックマークについての紹介を行う。次に 3 章でそれらを用いた Web ページ推薦システムの構築手法を説明し、4 章でシステムの実験方法について述べる。5 章で実験結果の考察と関連研究比較を行い、最後に 6 章でまとめを行う。

2. Folksonomy とソーシャルブックマーク

近年、セマンティック Web に代表される Web2.0 構想の概念要素の 1 つである Folksonomy (フォクソノミー) と、Folksonomy を利用したサービスであるソーシャルブックマークが注目を浴び、国内外で急速に普及し始めている。

2.1 Folksonomy

Folksonomy⁵⁾ は従来の Taxonomy に代わる新しい分類手法といわれている。たとえば Web ページの分類に関して述べるのならば、従来の Taxonomy においては Yahoo ディレクトリなどによって誰かがあらかじめ決定した分類木に基づいて各 Web ページが分類されていくという方法をとっていた。一方、これに対し Folksonomy では多数のエンドユーザが各々の Web ページに対し「タグ」と呼ばれるキーワードを付与していく。「タグ」は個々のユーザにとってそのページの性質や分類を表すキーワードであり、そのユーザの価値観に基づいて自由にどんなキーワードでも選ぶことができる。また、1 つの文書に複数のタグを付けることもできる。図 1 に Folksonomy によるタグ付けの一例を示した。こうして生成された大量のタグ情報を利用することで、様々な分類処理を行うことができるのが Folksonomy の特徴である。

トップダウンな Taxonomy に対して Folksonomy は Web ページ利用者自らがページ分類に参加するボトムアップアプローチであり、これによってユーザの実際の興味や利便性をリアルタイムに反映した分類情報を構築することが可能である。

一方 Folksonomy を扱ううえで解決しなければならない問題として「タグ表記のゆれ」の問題がある。Folksonomy ではユーザのタグ付け方法に関していっさいの制限を行わないため、同じ意味で表記の異なるタグが氾濫してしまうという事態が起こる。たとえば数学に関するある Web ページに対してユーザ A は「数学」タグを付け、ユーザ B は「math」タグを付

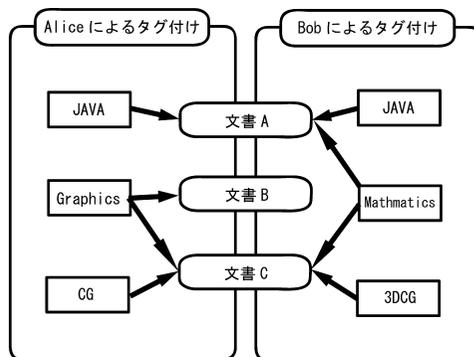


図 1 Folksonomy によるタグ付け
Fig. 1 Tagging in Folksonomy.

け、ユーザ C は「mathematics」タグを付ける、という状況が Folksonomy の世界では日常茶飯事である。Folksonomy を処理するシステムを構築する場合はこの問題を考慮する必要がある。

2.2 ソーシャルブックマーク

ソーシャルブックマーク (以下 SBM) は 1 つの Web サイト上で複数ユーザのブックマーク (お気に入りの Web ページ) 情報を共有するサービスであり、del.icio.us⁶⁾ が火付け役となって 2004 年後期から急速に普及し始めている。ユーザはブックマークを登録する際に Folksonomy に基づいてページに自由にタグ付けすることができ、タグによって複数ユーザのブックマーク情報が関連付けられる。

3. システムの仕組み

3.1 システムの概要

1 章で示したとおり従来の多くの Web ページ推薦システムは協調フィルタリングを用いている。したがって、あらかじめユーザ同士の嗜好の類似度を計算しておいた後に、嗜好の類似するユーザの選択したページ群を推薦する、というのが一般的な流れである。しかし推薦対象の Web ページを限定しない今回のシステムの場合、たとえデータソースに SBM を利用してもユーザ数 N に対する Web ページ数 M が圧倒的に多いという事実は変わらないため、ユーザ同士の嗜好の類似度を比較する際に「いくつかのブックマークページを共有しているか」という単純な方法ではほとんどの 2 ユーザ間の類似度は 0 になってしまううまく働かない。

この問題を解決するため、本論文ではユーザの嗜好を「ユーザと各タグとの親和度」で表現することにした。これによってユーザの嗜好表現が抽象化され、ブックマークページどうしを直接比較するよりもユーザ間の嗜好類似度が算出しやすくなる。

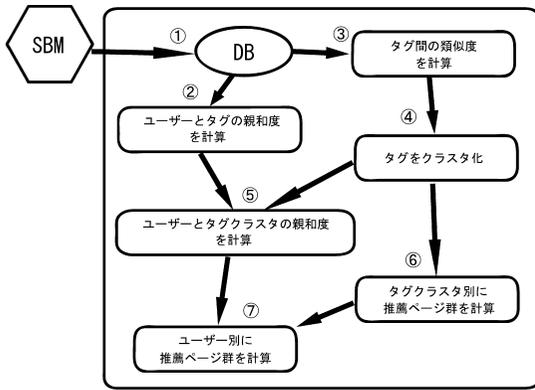


図2 全体の処理の流れ
Fig.2 Entire system chart.

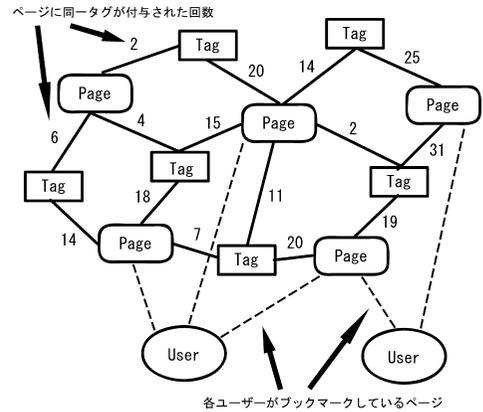


図3 SBMの内部データのモデル
Fig.3 Data model of Social Bookmark.

2.1 節で触れた「Folksonomy のタグ表記のゆれ」に対して本論文では類似タグをクラスタリングすることで解決を試みている．さらにこのタグクラスタリングを利用して「ユーザと各タグクラスタとの親和度」を計算することによって，ユーザの嗜好表現をよりいっそう抽象化させている．

また，4.3.1 項で後述するが，本推薦システムを利用するためにユーザは必ずしも SBM に加入している必要はない．普段使っている Web ブラウザのブックマークなどを用意することで，あらゆるユーザがこのシステムを利用することができる．

これらのことをふまえ，本論文で提案するシステムの全体的な処理の流れは図 2 のようになる．はじめに (1)～(7) の 7 つのステップについて簡単に紹介して全体の流れを示したあと，それぞれのステップの詳細を説明してゆく．

(1) データ収集

SBM 上で公開されているブックマークデータを収集し，DB に格納する．

(2) ユーザとタグの親和度を計算

Step(1) で構築したグラフをもとに，各ユーザと各タグ間の親和度を求める．

(3) タグ間の類似度を計算

Step(2) ではユーザとタグ間の親和度を求めたが，このステップでは同様のアルゴリズムを用いて各タグ間の類似度を計算する．

(4) タグのクラスタ化

Step(3) で求めたタグ間の類似度をもとにタグをクラスタ化する．

(5) ユーザとタグクラスタの親和度を計算

Step(2) と Step(4) の結果をもとに，今度は各ユーザと各タグクラスタの親和度を計算する．

(6) タグクラスタ別に推薦ページを計算

Step(4) で求めたタグクラスタ 1 つ 1 つがそれぞれ「話題」や「テーマ」の単位を構成していると見なし，タグクラスタごとに推薦ページ群を計算する．

(7) ユーザ別に推薦ページを計算

Step(5) で求めたユーザとタグクラスタの親和度と Step(6) で求めた各タグクラスタの推薦ページ群を組み合わせると，各ユーザに対する推薦ページ群を計算する．

3.2 各ステップの詳細

3.2.1 データ収集

del.icio.us に代表される SBM 上では，主に次の 2 種類の情報が公開されている．

- ユーザ A がブックマークしているページ群，およびユーザ A がそれらのページに付けたタグ群
- ページ P をブックマークしているユーザ群，およびページ P がそれらのユーザから付けられたタグ群

上の 2 種類の情報は実質的には同じものを表している．これらの情報は図 3 のように，ユーザ（ユーザ ID），Web ページ（URL），タグの 3 種類のオブジェクトが相互に結びつくグラフにモデル化できる．

図 3 において，タグとページが重み付き 2 部グラフを構成している．各辺に付与された重みは，異なるユーザによって特定のタグとページが関連付けられた回数である．さらに図 3 では「どのユーザがどのページ群をブックマークしているか」に関する情報も付加されている．

以降，このグラフにおけるページ P とタグ T の多重度を $w(P, T)$ と表し，ユーザ A のブックマークしているページ群を $bookmark(A)$ と表すことにする．

3.2.2 ユーザとタグの親和度を計算

このステップでは図3のグラフをもとに各ユーザと各タグの親和度を計算する．親和度はユーザ A とタグ T の関連度を相対的に表す値であり、「ユーザ “Bob” はタグ “car” と関連性が高い」といったような各ユーザの嗜好を表現する．また、タグ T が希少であればあるほど親和度は高くなりやすい．以降、ユーザ A とタグ T の親和度を $rel(A, T)$ というスカラー値で表すことにする (“rel” は “relation” の略)．

$rel(A, T)$ を計算するための前段階として、ページ P とタグ T の親和度 $rel(P, T)$ を次のように計算する．

$$rel(P, T) = TF(P, T) \times IDF(T)$$

$$TF(P, T) = \frac{w(P, T)}{\sum_{T_i \in TAGS} w(P, T_i)}$$

$$IDF(T) = \log \frac{\sum_{P_j \in PAGES} \sum_{T_i \in TAGS} w(P_j, T_i)}{\sum_{P_j \in PAGES} w(P_j, T)}$$

ページとタグの親和度を計算するアルゴリズムは $TF \cdot IDF$ をベースにしている．この場合、 $TF(P, T)$ はページ P に関連付けられているすべてのタグに対してタグ T が占める割合を示し、 $IDF(T)$ は全体のページにおけるタグ T の希少性を表す値である．

以上のアルゴリズムによりページとタグの親和度が求まったら、 $rel(A, T)$ を以下のように定義する．

$$rel(A, T) = \sum_{P_i \in bookmark(A)} rel(P_i, T)$$

3.2.3 タグ間の類似度を計算

Step(2) と同様のアルゴリズムを用いて、今度はタグ間の類似度を計算する．類似度も親和度と同様に2タグ間の関連度を相対的に表す値であり、「タグ “macintosh” とタグ “apple” は関連性が高い」といった事象を表現する．タグ T_1 にとってのタグ T_2 の類似度を $rel(T_1, T_2)$ と定義し、以下のように計算する．なお、この値は T_1 を基準とした値なので必ずしも (T_2, T_1) とは一致しない．

$$rel(T_1, T_2) = \sum_{P_i \in PAGES} w(P_i, T_1) \times rel(P_i, T_2)$$

3.2.4 タグのクラスタ化

Step(3) で求めたタグ間の類似度をもとに、話題やテーマとして類似するタグどうしをクラスタ化する．まず、前段階として各タグに対して自分と最も相関度が高いタグを表す「親タグ」を次のように定義する．

タグ T に対して $rel(T, T_i)$ の値が最も大きいタグ T_i を選び、 $rel(T, T_i)$ が閾値 V_{limit} を超えていたら T の親タグとし、超えていない場合は T 自身を T の親タグ

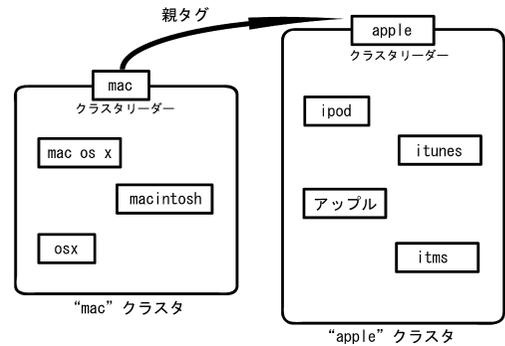


図4 クラスタリングの途中段階の例
Fig. 4 An example of tag clustering.

とする．2つのタグ T_1, T_2 が互いにとって最も類似度が高いタグである場合は、 $rel(T_1, T_2)$ と $rel(T_2, T_1)$ を比較して前者が大きければ T_2 を、後者が大きければ T_1 を双方の親タグとする．

以下にクラスタリングの基本的な手順を示す．

(1) すべてのタグがそれぞれサイズ1のタグクラスタを作っている状態から開始する．また、各々のタグを各クラスタの「クラスタリーダ」に設定する！「クラスタリーダ」とはその名のとおりそのクラスタを代表する最も重要度の高いタグである．

(2) 各クラスタ C に対し、クラスタリーダの親タグの属するクラスタを D とし、C と D をマージする．ただしマージ後のクラスタサイズが、あらかじめ決められている最大クラスタサイズを超えるようだったらマージを断念する．マージ後のクラスタのリーダは D のクラスタリーダに設定する．

図4はクラスタリング処理の途中段階の例である．タグ “mac” の親タグは “apple” なので、もし両方の合計クラスタサイズが C_{max} を超えなければ、“mac” クラスタは “apple” クラスタに吸収されてしまう．

親タグを決定するための閾値 V_{limit} や最大クラスタリングサイズ C_{max} を変化させることで、最終的に生成されるクラスタの粒度を調節することができる．

3.2.5 ユーザとタグクラスタの親和度を計算

このステップでは Step(2) によって計算されたユーザとタグの親和度をもとに今度は「ユーザとタグクラスタの親和度」を計算することで、ユーザの嗜好をより抽象的に表現できるようにする．ユーザ A とタグクラスタ C の親和度を $rel(A, C)$ と定義し、以下のよう求める．

$$rel(A, C) = \sum_{T_i \in C} rel(A, T_i)$$

図5はこのステップの具体例を表している．このようにユーザの嗜好をタグレベルからタグクラスタレ

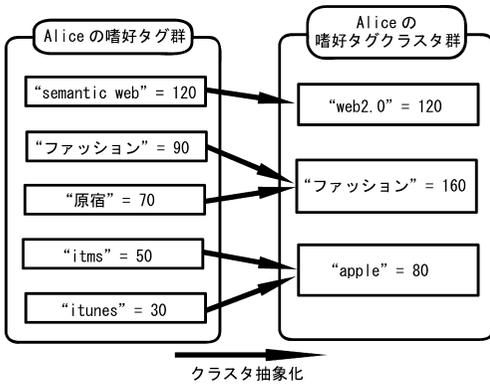


図5 ユーザの嗜好表現のクラスタ抽象化

Fig.5 Abstraction of user's preference.

ベルへと抽象化することにより Folksonomy の表記のゆれを吸収すると同時に、適度な大きさの話題の単位ごとにユーザにページを推薦することができるようになる。

3.2.6 タグクラスタ別に推薦ページを計算

このステップでは Step(4) で求めたタグクラスタ 1 つ 1 つがそれぞれ「話題」や「テーマ」の単位を構成していると見なし、タグクラスタごとに推薦ページ群を求める。クラスタ C から各ページ P に付与される「推薦ポイント」を $point(C, P)$ とおき、以下の式によって値を計算する。

$$point(C, P) = \sum_{T_i \in C} w(P, T_i)$$

この推薦ポイントが高いページから順番に、クラスタ C の推薦ページ群とする。

3.2.7 ユーザ別に推薦ページを計算

Step(5) で求めたユーザとタグクラスタの親和度と、Step(6) で求めたクラスタから各ページへの推薦ポイントを用いて、今度はユーザ別の推薦ページ群を求める。ユーザ A からページ P に付与される推薦ポイントを $point(A, P)$ とおき、以下の式によって値を計算する。

$$point(A, P) = \sum_{C_i \in CLUSTERS} rel(A, C_i) \times point(C_i, P)$$

この推薦ポイントが高いページから順番に、ユーザ A に対する推薦ページ群とする。

4. 実験

前章で示したシステムを Java で構築し、実際に存在する SBM に適用して実験を行った。

実験は、SBM の大量のデータをもとに推薦精度を測定する実験 (A) と、ユーザに実際に推薦ページを

評価してもらう実験 (B) の 2 種類を行った。実験 (A) ではどれくらいのクラスタリング粒度で最適な結果を出すのかという内部的な相対評価を行うのが目的である。そのため SBM 上の大量のユーザデータをもとに、タグのクラスタサイズ別にシステムの擬似的な推薦精度を計測し比較を行った。それに対し実験 (B) では、実験 (A) から導き出された最適なパラメータ設定を行った際、システムの推薦するページが実際にユーザにどのくらい受け入れられるのかというシステムの絶対評価を行うのが目的である。

4.1 実験概要 (実験 (A)・(B) 共通部分)

4.1.1 データ収集

今回は実験対象の SBM として“はてなブックマーク”⁷⁾を選んだ。同サービスは 2005 年 8 月現在日本最大規模の SBM であり、500,000 人以上の登録ユーザと 1,000,000 ページ以上のブックマークページデータが存在する。

今回ははてなブックマーク上で公開されているユーザブックマークデータのうち、現在もアクティブにブックマークを行っているユーザのデータを約 5,800 人分収集した。そのうち 5,000 人を訓練ユーザとし、訓練ユーザによってブックマークされたページデータ約 57,000 ページ分を収集した。さらに訓練ユーザによってそれらのページに付与されているタグ約 9,300 個分のデータを収集した。

その結果として得られたタグ ページ間のリンク総数 (図 3 の太線の数) は約 84,000 本、ユーザ ページ間のリンク総数 (図 3 の破線の数) は約 210,000 本だった。

4.1.2 タグのクラスタリング

Step(4) のタグのクラスタリングフェーズにおいて親タグを決定するための閾値 V_{limit} や最大クラスタリングサイズ C_{max} を変化させることで最終的に生成するクラスタの粒度を調節することができる。今回の実験においては図 6 のように 5 段階のクラスタリングの粒度でそれぞれ実験を行った。なお、SBM から収集した約 9,300 種類のタグのうち実際にシステム構築に利用したのは少なくとも合計 6 回以上出現したタグのみで、その数は 2,175 個だった。

ケース A はクラスタリングをまったく行わなかった場合である。すべてのクラスタのサイズは 1 で、全クラスタ数は実験に用いた全タグ数と同じ 2,175 個である。ケース A から B, C, D, E の順番でクラスタの粒度が大きくなり、クラスタ数は減っていく。表中に参考データとして“english”タグ、“java”タグ、“blog”タグをそれぞれリードとするクラスタサイズの比較を

	A	B	C	D	E
全クラスタ数	2175	1430	877	512	412
平均クラスタサイズ	1.00	1.52	2.48	4.25	5.28
Step(7)における1ユーザーあたりの平均処理時間(秒)	5.4	2.8	2.2	1.8	1.6
"english"クラスタのサイズ	1	3	5	7	10
"java"クラスタのサイズ	1	11	15	30	34
"blog"クラスタのサイズ	1	28	51	64	70

図 6 各クラスタリングケースの比較

Fig. 6 List of 5 clustering cases.

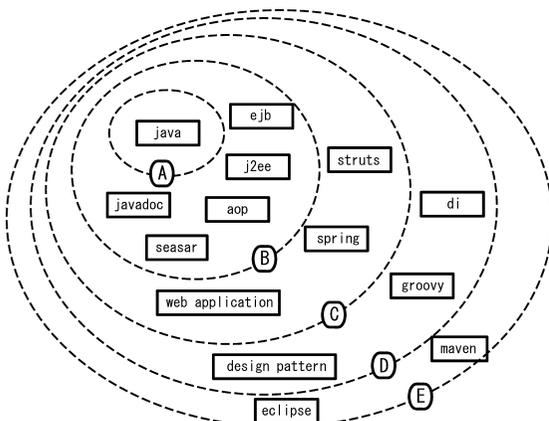


図 7 各クラスタリングケースにおける“java”クラスタの中身

Fig. 7 Members of “java” cluster in each case.

載せた。クラスタの粒度が大きくなるほどこれらのクラスタサイズも大きくなっているのが分かる。

また、各ケースごとの Step(7) における 1 ユーザあたりの平均処理時間を図 6 中に示した。これを見ると処理時間とクラスタ数がおおまかに比例関係になっていることが分かる。ちなみに Step(1) を除いた Step(2) ~ Step(6) における処理時間は合計して約 1~2 時間程度だった。本システムにおいては Step(6) までを前処理としてあらかじめ行った後、ユーザからのクエリに対して Step(7) のみを実行するのでユーザの体感速度に影響するのは Step(7) の処理のみである。

本システムが正しい推薦結果を返すためにはタグのクラスタリングが適切に行われている必要がある。クラスタリングがねらいどおりに適切に行われたか、具体例を見ながら考察する。

図 7 は、ケース A ~ E における “java” クラスタの例である。実際には “java” クラスタにはもっと多くのタグが属するが、図 7 では主なタグ（出現数が大きいタグ）のみ載せている。B や C の “java” クラスタ中には “java” に対して関連性や依存性が高いタグが並ぶ。一方、D や E で現れている “design pattern” や “eclipse” は “java” と関連性が高いものの、それ自

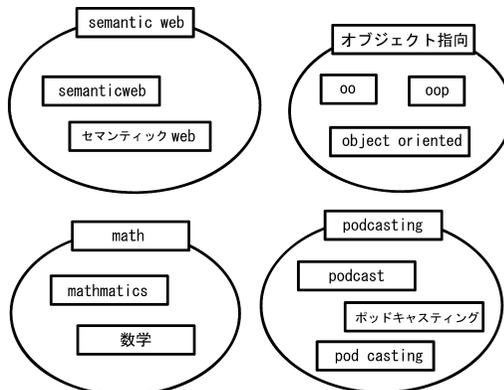


図 8 ケース B, C に見られた同義語クラスタの例

Fig. 8 Examples of “Synonym Tag Cluster” in case B and C.

体独立性が高い話題であり、それなりの大きさのクラスタを作っていたのが “java” クラスタに吸収されたと見ることができる。多少主観的な評価であるが、図 7 に関してはクラスタリングは適切に行われているといえる。同様に他のクラスタの中身も一般的に関連度が高いと見せるタグの集合で構成されていた。

また、ケース B, C においては図 8 に示したようないわゆる「同義語クラスタ」が数多く見られた。図 8 を見て分かるように今回の実験によって英語表記、日本語表記、略語など多様な表記のタグがうまくクラスタリングされており、これによって Folksonomy を扱ううえで大きな課題だった「表記のゆれ」の問題が解決されているといえる。

4.2 実験 (A)

4.2.1 実験 (A) の評価方法

実験 (A) の目的はクラスタリング粒度とシステムの推薦精度の関係を調べることである。そのため以下のような方法で SBM 上の大量のユーザデータをもとに、タグのクラスタサイズ別にシステムの擬似的な推薦精度を計測し比較を行った。図 9 は本章の実験と評価の全体的な流れを示している。

まず、Step(1) で SBM からデータを収集する際、ユーザを訓練ユーザとテストユーザの 2 種類にランダムに分け、訓練ユーザに関連するデータのみを収集する。こうして収集された訓練ユーザに関連するデータを用いて Step(2) ~ Step(6) を実行し、ページ推薦システムを構築する。

次に、各テストユーザに対して以下の処理を行い、システムの評価を行う。

テストユーザ A のブックマークページ群を “クエリ用ページ群” と “テスト用ページ群” の 2 種類に分け、

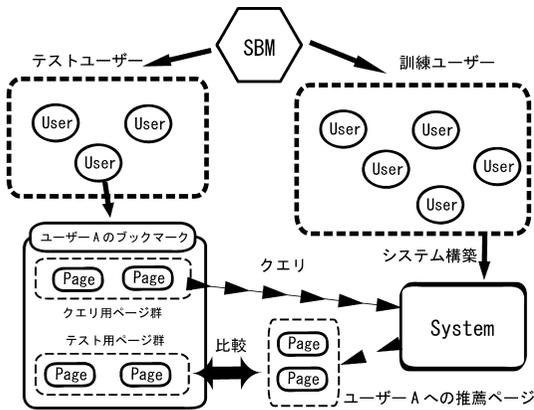


図 9 Web ページ推薦システムの性能評価手順

Fig.9 Evaluation of the Web page recommender system.

“クエリ用ページ群”をシステムにクエリとして投げる．そうするとシステムはユーザ A に対して推薦ページ群を返してくるので，この推薦ページ群とユーザ A のテスト用ページ群の一致度を比較することで，本当にシステムがユーザの好むページを推薦してくれるのかを評価する．

推薦ページ群とテストページ群の一致度には，以下の 2 種類の評価指標を用いる．ここで推薦ページ数を R，テストページ数を T，両者の一致したページ数を H とする．

(1) 再現率 = H / T

この値は推薦ページ数 R が大きくなるほど増加する

(2) 適合率 = H / R

この値はテストページ数 T が大きくなるほど増加する

4.2.2 実験 (A) の結果

上で述べた評価方法に基づき，A～E の 5 種類のクラスタリングケースごとにシステムの推薦精度の評価を行った．1 ユーザのブックマークページ数が再現率や適合率に大きな影響を与えるため，ユーザをブックマークページ数の大きさに応じてクラス分けし，クラスごとに集計を行った．また，テストユーザのブックマークのうちランダムに選んだ 50% をクエリ用ページ群，残りの 50% をテスト用ページ群として用いた．システムの返す推薦ページ群はつねに 30 ページに固定した．

図 10，図 11 はそれぞれクラスタリングケース別の平均再現率と平均適合率をまとめたものである．横軸はブックマーク数の大きさに応じて分けられたユーザクラスを表している．

図 10，図 11 を見てみると，ブックマーク数が 15 未満のライトユーザに対してはクラスタリングケース B の推薦精度が最も良いことが分かる．またこれらのラ

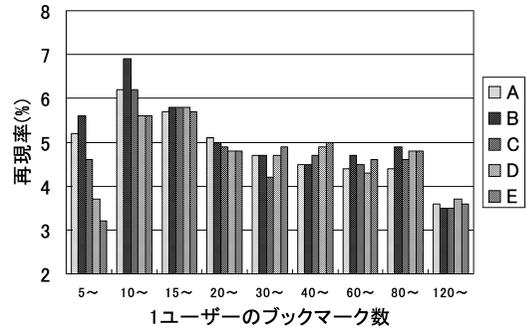


図 10 実験 (A) : 各クラスタリングケースごとの再現率の比較
Fig. 10 Reflecting rate in each case.

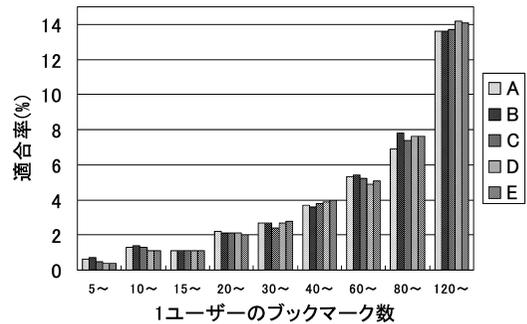


図 11 実験 (A) : 各クラスタリングケースごとの適合率の比較
Fig. 11 Precision rate in each case.

イトユーザに対しては D や E のような粒度の大きいクラスタリングは逆効果になっていることが分かる．

これについて考察すると，そもそもタグのクラスタリングは図 5 のように与えられたテストページ群のタグを話題ごとに集計するという意味があるのに対し，与えられたテストページが少なすぎて集計効果が働いてないことが原因としてあげられると思われる．

その中でもケース B が有効に働いているのは，図 8 で示したような同義語クラスタリングの効果が顕著に表れているためだと思われる．ブックマーク数の少ないユーザの場合，Step(2) で算出される親和度の高いタグ群がマイナーなタグばかりで構成されてしまう可能性が高く，この場合は結果としてマイナーなページばかり推薦されてしまう．いわゆる「表記のゆれ」の弊害である．これらのマイナータグを同じ意味のもっとメジャーなタグのクラスタに吸収させることで，多くの人の嗜好を反映させたよりメジャーなページが推薦される．

一方，ブックマーク数が 60 を超えるようなヘビーユーザに対しては D や E などの粒度の高いクラスタリングが有効に働いているのが分かる．これはブックマーク数が多い場合，図 5 で示したようなタグの話題

ごとの集計，抽象化が狙いどおりうまく働いているからだと思われる．

結果として，推薦精度は全体的にケース B が最も優れていた．ケース C, D, E の推薦精度があまり上がらなかった原因として，話題を抽象化しすぎたために個人の細かい嗜好が反映されず，一般的に人気のあるページばかり推薦されてしまうという問題が起こっているとされる．

なお，実験 (A) で求めたこれらの再現率や適合率はあくまで「ユーザが実際にブックマークしたページが推薦されたか」を測る指標にすぎず，実際の推薦精度を示しているわけではない．実際のシステムの推薦精度は実験 (B) のユーザテストによって計測する．

4.3 実験 (B)

4.3.1 実験 (B) の評価方法

実験 (B) では実際のユーザにページ推薦システムを利用してもらい，推薦ページの内容を評価してもらうことで Web ページ推薦システム本来の評価指標である「ユーザが実際に興味を持つ Web ページが推薦されたか」を測る．なお，実験 (B) におけるパラメータ設定（クラスタリング粒度）は実験 (A) で相対的に最も良い結果を出したもの（クラスタリングケース B）を用いた．

今回は SBM の会員ではない 10 人のユーザに実験に参加してもらった．ユーザが SBM の会員である場合は SBM 上のそのユーザのブックマークページをそのまま計算に用いればよいが，ユーザが SBM の会員でない場合にはユーザの嗜好するページ群を代わりに用意してシステムに別途インプットする必要がある．インプットされたページ群に対しては SBM のユーザのブックマークとまったく同様の処理がされ，ユーザに対する推薦ページ群が計算される．今回は各ユーザが普段使っている Web ブラウザのブックマークをシステムにインプットしてもらった．この場合，インプットされたページ群のうち SBM にも登録されているページのみが計算に用いられる．このように本推薦システムは SBM に加入していないユーザに対しても，SBM のデータの学習結果を利用して推薦ページを計算することができる．

まず，各ユーザは普段使っている Web ブラウザのブックマークデータを HTML インタフェースからアップロードし，システムに入力させる．システムはそのブックマークデータを解析し，出力としてそのユーザと最も親和度が高いタグ（Step(5) 参照）上位 30 個と，ユーザへの推薦ページ上位 30 ページを HTML インタフェースから表示する．

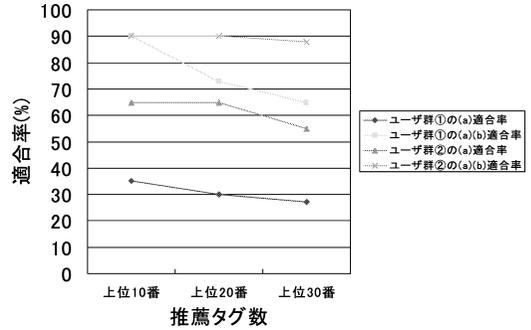


図 12 実験 (B) : 推薦タグの適合率の比較

Fig. 12 Reflecting rate of recommended tags.

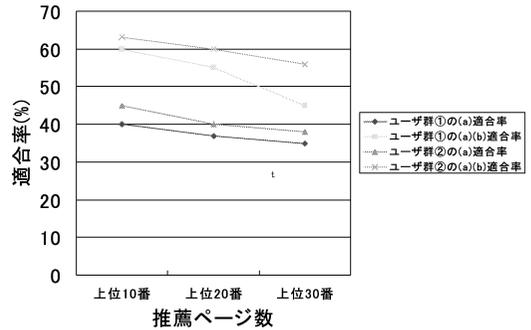


図 13 実験 (B) : 推薦ページの適合率の比較

Fig. 13 Reflecting rate of recommended pages.

ユーザは出力された 30 個の推薦タグと 30 個の推薦ページに対して，それぞれ内容を以下の 3 段階で主観評価する．

- (a) とても興味があり，内容が自分と関係が深い．
- (b) 興味はあるが，とくに自分と関係が深いわけではなく，万人向けの内容である．
- (c) 興味はない．内容も自分とは関係ない．

この評価結果を集計し，実験 (B) では Web ページ推薦システムの絶対的な適合率を求める．

4.3.2 実験 (B) の結果

上で述べた評価方法に基づきシステムの性能評価を行った．実験 (A) と同様に 1 ユーザのブックマーク数の大小が適合率に大きな影響を与えるため，ここでもユーザのクラス分けを行った．ここでは有効ブックマーク数が 50 ページ未満のユーザをユーザ群 (1) に，50 ページ以上のユーザをユーザ群 (2) に振り分けた．有効ブックマーク数とはユーザのブックマークページのうちで，実際にシステムで解析に用いられたブックマークページの数である（いい換えると step(1) で SBM から回収されたページ群に含まれるページの数である）．

図 12，図 13 はそれぞれユーザクラス別に推薦タ

グと推薦ページの適合率を比較したものである．横軸は上位何件で適合率を計算するかを示している．また図中の“(a) 適合率”とは (a) ~ (c) の選択肢のうち (a) が選択された割合を示し，“(a)(b) 適合率”とは (a) が (b) が選択された割合を示す．

図 12, 図 13 ではともにユーザ群 (1) とユーザ群 (2) にユーザをクラス分けして比較しているが, 期待どおりブックマーク数の多いユーザ群 (2) の方が適合率が高い．また, 当然のことではあるが推薦ページ・タグを上位に絞れば絞るほど高い適合率が出ている．テストユーザ数が決して多くないにもかかわらず安定して理にかなった結果が出ていることが分かる．

(a) 適合率の場合で平均 40%, (a)(b) 適合率の場合で平均 60%というのがこの Web ページ推薦システムの総合評価といえる．

また, 実際に推薦ページの計算に使われたのはインプットされたブックマークページのうち SBM に登録されているページだけだったが, その割合はインプットのうち 3 割を超えていた．このことから近年の SBM がいかに幅広い範囲の Web ページをカバーしているかということも分かる．

5. 考察と関連研究

1 章で述べたとおり, 従来の多くの Web ページ推薦システムはユーザのアクセス履歴をもとに推薦ページを導き出しているのに対し, 我々の推薦システムは複数ユーザのブックマークデータをもとに推薦ページを導き出している．まずは, ブックマークをもとにユーザの興味のある Web ページを推薦することの妥当性について考察する．経験的に, ユーザがブックマークするページの性質はおおまかに以下の 3 種類に分類できる．

- (a) 検索サイト, ニュースサイト, ポータルサイト, ショッピングサイトなど実用性が高く大規模で汎用的な, 誰しもがブックマークするようなサイト (例: Google, Yahoo!, Amazon, CNN)
- (b) ユーザの趣味, 仕事, 生活に関連する情報を持つページ (例: Java の入門ページ, 野球チームのファンサイト, ペットに関するサイトなど)
- (c) ユーザの実生活のコミュニティに関連するページ (例: 仲間内の掲示板, 研究室やプロジェクトの Wiki, 知人のブログ, SNS など)

この 3 種類のうち, (b) こそがユーザの趣味, 嗜好を直接的に反映している「推薦されるべきページ」といえる．(c) に分類されるページは通常 SBM 上には登録されないため, 本システムにおいて推薦対象となる

ことはない．また 3.2.2 項で示したように本システムでは TF・IDF を用いることで誰しもがブックマークするような汎用的なページに対しては低い重み付けがなされるため, (a) よりも (b) に分類されるページの方が推薦されやすいという性質を持つ．実際, Google や Yahoo! は多くのユーザがブックマークしていたにもかかわらず, 実験 (B) においてユーザへの推薦ページとして選出されることはなかった．以上のことから, 本システムを用いてブックマークをもとにユーザの興味ある Web ページを推薦することには妥当性があると思われる．

次に, 計算量的な観点から本システムの実現可能性を考察してみたいと思う．今回の実験では SBM に登録された 57,000 の Web ページを対象にしたが, 現在の del.icio.us の規模は数千万ページのオーダーであり, 今後 Folksonomy が SBM の枠組みを越えて広く使われるようになる可能性を考慮すると, 数億~数十億ページのオーダーを扱う可能性もある．これは現在の検索エンジンに匹敵する規模である．本システムのアルゴリズムは基本的に「図 3 の 3 部グラフにおいて各ノードからリンクを 1 段, もしくは 2 段たどる」しくみなので, その計算量はたかだか $O(n^2)$ である．実際には 1 つのノードから出るリンクの数は n よりもはるかに小さいので, 計算量はむしろ n の線形に近いと思われる．また, Step(7) はユーザブックマークというインプットを入れてから推薦ページというアウトプットを出すまでにかかるいわゆるシステムの体感時間に関わる重要なフェーズだが, 図 6 で示したようにこのフェーズの計算時間はタグのクラスタリング粒度を大きくすることによって短縮させることができる．また, Step(1) ~ Step(7) のすべてのフェーズは容易に計算を分散化させることができるので, 計算資源を投入することによって計算時間を大幅に短縮させることができるだろう．これらのことから, 本システムは現在の検索エンジンのようにインターネット規模の Web 文書を扱うことも可能であると考えられる．

実験 (A) の結果より, タグのクラスタリングによる「話題の抽象化」自体は適切に行われているが推薦精度に直接的に与える影響は少なく, クラスタリング粒度に比例して計算時間が短縮できることが分かった．また, タグのクラスタリングにより Folksonomy におけるタグ表記のゆれの問題を解決することができた．実際のシステム運用時にクラスタリング粒度を決定する際には図 6, 図 10, 図 11 の結果をもとに推薦精度と計算時間のトレードオフで決定するのが望ましいと思われる．

また、実験 (B) のユーザテストにより本システムの推薦精度 (適合率) として 40~60% という値が導き出された。後述するが、インターネット全体を対象とした既存の Web 推薦システムに比べてこれは突出しているとはいえないまでもまったく遜色がない値である。インターネット全体を対象とした既存の Web 推薦システムの多くがユーザに何らかの作業を要請するのに対し、本システムにおいてはユーザは普段使っているブラウザのブックマークを用意する以外に何の追加作業もする必要がない。

繰返しになるが、従来の多くの Web ページ推薦システムはユーザのアクセス履歴を主な情報源として協調フィルタリングを用いている。Li ら¹⁾ はアクセス履歴に加え Web ページの内容マイニング、リンク構造マイニングを組み合わせてユーザのサイトナビゲーションパターンを抽出し、Web ページ推薦システムを構築している。しかしこの手法においてもアクセス履歴が主要なデータソースであることに変わりはなく、推薦対象となるページは特定のサイト群に限定されてしまっている。

推薦対象を特定のサイト群に限定しない Web ページ推薦システムの数少ない有名な例としては Zhu らの “Web ICLite” がある⁸⁾。Web ICLite はユーザの Web 閲覧履歴をユーザサイドで記録し、検索行動などにおけるナビゲーションパターンを解析することでユーザの好むページやキーワードを導き出す。さらにこれらの解析情報を 1 カ所まで収集し、ユーザの好むキーワードに基づいてページ推薦を行う。Web ICLite はユーザが嗜好する話題単位で推薦ページを選出している点が我々のシステムと似ている。基本的な協調フィルタリングでは嗜好の類似する “ユーザ” 単位で推薦ページを選出しているが、話題単位の推薦の方が精度が上がる。

Web ICLite のユーザテストの結果は、本論文でいうところの (a) 適合率が 30~50%、(a)(b) 適合率が 60%~70% である。ユーザテストはユーザの主観的な判断によるところが大きいため単純には比較できないが、この値は我々のシステムの結果と非常に近い。しかし Web ICLite を使うには普段から Web 閲覧の際にシステムをユーザマシン上で稼働させ、Web ナビゲーションパターンを解析させる必要がある。また、相当の数のユーザにシステムを使ってもらい訓練データを収集しなければ推薦システムが成立しない。これに対し我々のシステムでは訓練データは既存の SBM を利用し、ユーザデータは普段使っているブラウザのブックマークを利用しているので、ユーザが事前に何

の準備をする必要もなく、非常に手軽にシステムを利用することができる。

6. 結 論

本論文では SBM (ソーシャルブックマーク) 上で公開されている大量のユーザのブックマーク情報をもとに、従来は実現が困難だったインターネット全体を対象とした Web ページ推薦システムを構築した。また、その過程として Folksonomy に基づく大量のタグ情報をマイニングして各ユーザの嗜好をタグレベルで表現し、さらに類似タグをクラスタリングすることでユーザの嗜好をタグクラスタレベルへ抽象化して表現した。タグのクラスタリングによって推薦精度を下げることなく推薦ページの計算時間を短縮し、さらに従来から Folksonomy に対していわれてきた「タグ表記のゆれ」の問題を解決する方法を提案した。

参 考 文 献

- 1) Li, J. and Zaiane, O.: Combining Usage, Content and Structure Data to Improve Web Site Recommendation, *Proc. WebKDD-2004 workshop on Web Mining and Web Usage* (2004).
- 2) Kazienko, P. and Kiewra, M.: Integration of relational databases and Web site content for product and page recommendation, *International Database Engineering and Applications Symposium, (IDEAS 2004)* (2004).
- 3) Golovin, N. and Rahm, E.: Reinforcement Learning Architecture for Web Recommendations, *Proc. International Conference on Information* (2004).
- 4) A Mathes — Retrieved: Evaluating collaborative filtering recommender systems, *ACM Trans. Inf. Syst.* (2004).
- 5) Golbeck, J., Parsia, B. and Hendler, J.: Folksonomies-Cooperative Classification and Communication Through Shared Metadata (2004).
- 6) del.icio.us. <http://del.icio.us/>
- 7) はてなブックマーク. <http://b.hatena.ne.jp/>
- 8) Zhu, T., Greiner, R., Haeubl, G., Price, B. and Jewell, K.: A Trustable Recommender System for Web Content, *International Conference on Intelligent User Interfaces* (2005).
- 9) Ziegler, C.-N.: Semantic Web Recommender Systems, *Proc. Joint ICDE/EDBT, Ph.D. Workshop* (2004).

(平成 17 年 9 月 30 日受付)

(平成 18 年 3 月 2 日採録)

**丹羽 智史**

1982年生。2005年東京大学理学部情報科学科卒業。同年同大学大学院情報理工学系研究科コンピュータ科学専攻修士課程進学。現在に至る。Webマイニング、ネットワーク分析

に興味を持つ。ACM、日本ソフトウェア科学会等各会員。

**土肥 拓生**

1980年生。2003年東京大学理学部情報科学科卒業。同年同大学大学院情報理工学系研究科コンピュータ科学専攻修士課程進学。2005年同修了。同年同大学院同専攻博士課程

進学。現在に至る。マルチエージェントシステム、ソフトウェア工学に興味を持つ。日本ソフトウェア科学会会員。

**本位田真一（正会員）**

1978年早稲田大学大学院理工学研究科修士課程修了（株）東芝を経て2000年より国立情報学研究所教授。2004年より同研究所研究主幹を併任。現在に至る。2001年より

東京大学大学院情報理工学系研究科教授を併任。現在に至る。2002年5月～2003年1月英国UCLならびにImperial College客員研究員（文部科学省在外研究員）。2005年度パリ第6大学招聘教授。早稲田大学客員教授。工学博士（早稲田大学）。1986年度情報処理学会論文賞受賞。ソフトウェア工学、エージェント技術、ユビキタスコンピューティングの研究に従事。IEEE、ACM、日本ソフトウェア科学会等各会員。本学会理事。