

ブックマークの時系列情報を利用した ソーシャルブックマークにおける注目度予測

毛受 崇[†] 吉川 正俊^{††}

[†] 名古屋大学大学院 情報科学研究科 〒464-8601 名古屋市千種区不老町

^{††} 京都大学大学院 情報学研究科 〒606-8501 京都市左京区吉田本町

E-mail: [†]menjo@dl.itc.nagoya-u.ac.jp, ^{††}yoshikawa@i.kyoto-u.ac.jp

あらまし ソーシャルブックマークは、ウェブページのブックマーク情報をウェブ上で管理でき、また、不特定多数の人間と共有できるサービスである。従来はローカルのウェブブラウザに蓄積されていたブックマーク情報が集約、共有されることで、ソーシャルブックマークはウェブの新たな情報源として台頭し始めている。我々はソーシャルブックマーク利用者、および利用者がブックマークに付与する“タグ”の両者の時系列情報を利用し、新しく投稿されたページの注目度を予測する手法を提案する。

キーワード ソーシャルブックマーク, トレンド予測, ランキングアルゴリズム

Trend Prediction in Social Bookmark Service Using Time Series of Bookmarks

Takashi MENJO[†] and Masatoshi YOSHIKAWA^{††}

[†] Graduate School of Information Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8601

^{††} Graduate School of Informatics, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto, 606-8501

E-mail: [†]menjo@dl.itc.nagoya-u.ac.jp, ^{††}yoshikawa@i.kyoto-u.ac.jp

Abstract Social bookmark is a web-based service which enables its users to manage and share their bookmarks. Many bookmarks aggregated and shared on social bookmarks, so they become useful news sources now. In this paper we propose a trend prediction method of newly-posted pages, using sequential data of users and “tags” annotated to bookmarks.

Key words Social Bookmark, Trend Prediction, Ranking Algorithm

1. はじめに

近年、ウェブ上の新たな情報源としてソーシャルブックマークと呼ばれるサービスが台頭している。ソーシャルブックマークの利用者はウェブページのブックマーク情報をウェブ上で管理でき、また、不特定多数の人間と共有できる。従来であればローカルのウェブブラウザに保存されていたブックマーク情報が集約、共有されるため、ソーシャルブックマークは人手によって厳選された有用な情報が発見できる場として注目されている。代表的なサービスとして del.icio.us [1]、はてなブックマーク [2] がある。

ソーシャルブックマークにおいてウェブページは“タグ”によって分類される。タグは利用者がブックマークしたページに対して自由に付与できる短い文字列で、多くの場合はページの内容に関連する単語が用いられる。サービス提供者が天下り的に分類を行うウェブディレクトリなどとは対照的に、利用者自

身が情報を分類できるこのような仕組みはフォークソノミーと呼ばれており、ブログなども含めた消費者発信型メディアの情報分類の手段として広く用いられている。

現在の主要なウェブ検索エンジンでは PageRank [3] に代表されるリンク解析アルゴリズムに基づいてページの重要度が測られ、順位付けが行われる。しかしリンク解析は計算コストが大きく、また新たに公開されたページがインリンクを獲得するまでに時間がかかることから、新しいページの重要度を測るには向かないと考えられる。一方、多くのソーシャルブックマークではウェブページのブックマーク数を利用することで、迅速なページの順位付けを行っている。例えば、ソーシャルブックマークのトップページには、最近多くの利用者によってブックマークされた、あるいはされ始めたページとして“人気ページ”や“注目ページ”などと呼ばれる一覧が表示されている。利用者はこれらの一覧から現在話題のページを容易に知ることができる。

しかし、ソーシャルブックマークにおけるウェブページの順

位付けには、まだ利用されていない情報が多い。例えば、日頃有用なページを多くブックマークする利用者、一般的な利用者、スパム利用者、これらはウェブページの順位付けの際にはいずれも一人と数えられる。だがブックマークの振る舞いを考慮し、利用者ごとに適切な重みを与えることができれば、前者がブックマークしたページを重視するような順位付けが可能になる。また、タグについても同様で、利用者の興味、関心や現在流行の話題を把握し、順位付けに反映することが考えられる。そこで本論文では、利用者およびタグを考慮したソーシャルブックマークにおけるウェブページの順位付け手法を提案する。この手法は新しく投稿されたページが今後どの程度の注目を集めるかを予測し、その度合いによって順位付けるものである。手法において我々が用いるのは利用者およびタグの時系列情報である。ここからまず個々の利用者およびタグを評価し、その評価を利用してページの注目の度合いを予測する。

本論文の以下の流れを簡単に示す。2. 節では関連する研究を挙げ、本研究との差異について述べる。3. 節では我々が提案する手法の詳細を説明する。4. 節では3. 節で提案した手法の有効性を評価する実験を行った結果を示し、考察する。最後に5. 節で本論文をまとめ、今後の課題を挙げる。

2. 関連研究

ソーシャルブックマークは、現在活発に研究が行われている分野であるが、特に利用者に対するブックマークの推薦に関する研究が盛んに行われている [4] [5] [6] [7] [8]。本研究は推薦ではなくページの順位付けを目的としている点で異なる。

Hotho らはソーシャルブックマークにてブックマークされたページの順位付けに PageRank を応用したアルゴリズムを提案している [9] [10]。また、山家らはソーシャルブックマークと PageRank を融合したウェブ検索のランキング手法を提案している [11]。ブックマークを用いたウェブページの重み付けという点では本研究と共通しているが、Hotho らがソーシャルブックマークに投稿された全てのページを、山家らが既存のページも含めたあらゆるページを対象としているのに対し、本研究は特にソーシャルブックマークにてブックマークされた新しいページを対象としている点で異なる。

3. ウェブページの注目度予測手法

3.1 ブックマークおよびブックマーク列のモデル化

図1は2007年2月6日にはなブックマークに投稿されたあるページのブックマークの推移を、初めて投稿された時刻から一週間後まで、6時間単位で集約して示したものである。破線はページが各時区間に獲得したブックマーク数を、一点鎖線は隣接する時区間同士のブックマーク数の差分を示す。これらはそれぞれ、ブックマーク増加の速度、および加速度と見なせる。また、実線はその時区間までの累積ブックマーク数を示す。図の破線を見ると、ページが最初に投稿されてから一両日中にブックマーク数の増加はピークを迎え、その後の増加は次第にゆるやかになることが分かる。Biddulph [12] や Golder ら [13] が指摘するように、この傾向は del.icio.us においても見られる。ま

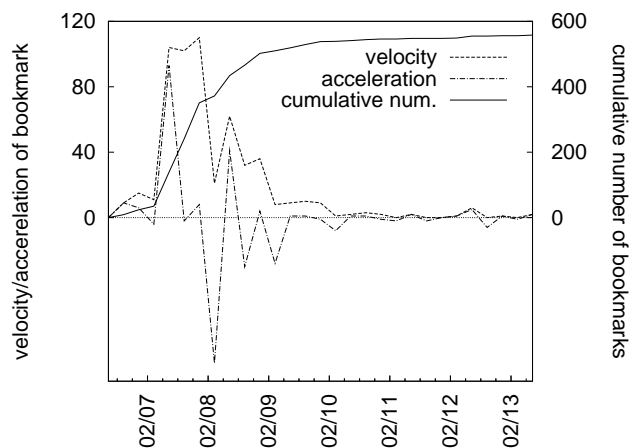


図1 ブックマーク数の時間推移

た、増加のピークが始まる区間と終わる区間には加速度の大きな極大値と極小値が現れており、累積ブックマーク数はこの間に大きく伸びていることが分かる。そこで我々は、ブックマーク増加の加速度が大きな極大値をとる直前にページをブックマークしていた利用者を、そのページの重要性を発見した利用者として見なし注目する。

ここまでの議論を定式化する。まず、我々は一つのブックマーク b を四つ組によって

$$b = (u, p, t, L)$$

とモデル化する。ここで u は利用者、 p はウェブページ、 t はブックマーク時刻、 L は付与されたタグの集合であり、一つの四つ組は“誰が、どのページを、いつ、どんなタグを付与して”ブックマークしたかを表す。 u は p を多くとも一度しかブックマークできない。また、ブックマークにはタグを付与しなくてもよく、その場合は $L = \emptyset$ となる。

あるページ p が最初にブックマークされた時刻から時間 T 経過後までのブックマークを時刻順に並べた列を

$$B_p^T = [(u_1, t_1, L_1), (u_2, t_2, L_2), \dots, (u_n, t_n, L_n)]$$

とする。この列をある単位時間 τ (ただしある正整数 l に対して $T = l\tau$) ごとに集約して得られるブックマーク数の列を

$$V_p^{(T, \tau)} = [v_1, v_2, \dots, v_l]$$

とする。すなわち、

$$v_i = \left| \left\{ (u_k, t_k, L_k) \mid (u_k, t_k, L_k) \in B_p^T \wedge t_1 + (i-1)\tau \leq t_k < t_1 + i\tau \right\} \right|$$

である。各 v_i は時刻 $t_1 + i\tau$ におけるブックマーク増加の速度である。また、

$$a'_i = v_{i+1} - v_i \quad (i = 1, 2, \dots, l-1)$$

は時刻 $t_1 + i\tau$ におけるブックマーク増加の加速度であり、その列を

$$A_p^{(T, \tau)} = [a'_1, a'_2, \dots, a'_{l-1}]$$

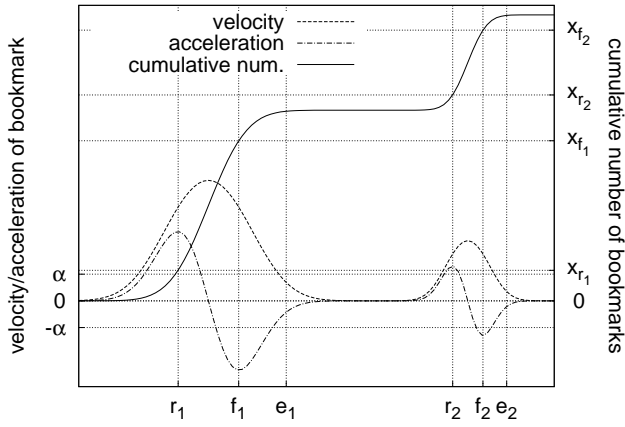


図2 理想的なブックマーク数の時間推移

とする。ただし $A_p^{(T,\tau)}$ には雑音が多く含まれるため、 a'_i を以下の式を用いて平滑化する。

$$\begin{aligned} a_1 &= (3a'_1 + a'_2) / 5 \\ a_i &= (a'_{i-1} + 3a'_i + a'_{i+1}) / 5 \quad (i = 2, 3, \dots, l-2) \\ a_{l-1} &= (a'_{l-2} + 3a'_{l-1}) / 5 \end{aligned}$$

これ以降は平滑化後に得られる列

$$A_p^{(T,\tau)} = [a_1, a_2, \dots, a_{l-1}]$$

を加速度と見なして用いる。

同様に、時刻 $t_1 + i\tau$ における累積ブックマーク数 x_i と、その列 $X_p^{(T,\tau)}$ を

$$x_i = \sum_{j=1}^i v_j \quad (i = 1, 2, \dots, l),$$

$$X_p^{(T,\tau)} = [x_1, x_2, \dots, x_l]$$

とおく。

3.2 ブックマークの伸びとその評価

3.1 節で述べたように、我々はブックマーク増加の加速度を用いてウェブページを発見した利用者を決定するが、本節以降ではこれら利用者の決定の詳細および評価について述べる。

図1のようなブックマークの推移を理想化したものとして図2を考える。図の横軸には添字 i を、縦軸には速度 v_i 、加速度 a_i 、累積ブックマーク数 x_i を取っている。両軸とも実際には離散的であるが、図では簡単のため連続的に示している。この図では2箇所ブックマーク数が伸びている。Biddulph [12] も指摘しているように、ブックマークはある時点においてのみ大きく伸びるのではなく、大小の伸びが複数回あるのが一般的である。それぞれの伸びは加速度が極大、極小となる時刻に対応する添字 $i = r_j, f_j$ (図では $j = 1, 2$) によって規定される。以下ではこの区間を j 番目の伸び区間と呼ぶ。

我々は、ページの重要性を発見した利用者は、伸び区間の直前にページをブックマークしていた利用者であると見なす。そして次の考えを元にこれらの利用者进行评估する。すなわち、ページが他の多くの利用者に素早く広まるほど、そのページの重要

性を発見した利用者は高く評価されるべきである、という考えである。図2に戻れば、伸び区間の長さ $f_j - r_j$ が小さいほど、またブックマークの伸び $x_{f_j} - x_{r_j}$ が大きいほど、利用者を高く評価すべきとも言い換えられる。

ここまで述べてきたような利用者の評価を行うために、以下では準備としてブックマークの伸びの評価を行う。まず、閾値 $\alpha, \beta (> 0)$ 、および正整数のパラメータ γ を設定し、ブックマークの全ての伸び区間を次の手順によって得る。

(1) $j = 1, s = 1$ とする。

(2) $i \geq s$ において、 $a_i > a_{i-1}$, $a_i > a_{i+1}$, かつ $a_i > \alpha$ を満たす最小の i を r_j とする。そのような i が存在しない場合は計算を終了する。

(3) $i > r_j$ において、 $|a_k| < \alpha$ ($i \leq k < i + \gamma$) かつ $\sum_{k=i}^{i+\gamma-1} |a_k| < \beta$ を満たす最小の i を e_j とする。そのような i が存在しない場合は $e_j = l$ とする。

(4) $r_j < i \leq e_j$ において、 $a_i < a_{i-1}$, $a_i < a_{i+1}$, かつ $a_i < -\alpha$ を満たす最小の i を f_j とする。そのような i が存在しない場合は $a_i < a_{i-1}$ かつ $a_i < a_{i+1}$ を満たす最小の i を f_j とする。そのような i も存在しない場合は $f_j = r_j + 1$ とする。

(5) $s = e_j + 1$ とし、 j を1増やして(2)に戻る。

α はブックマーク数の伸び始めや伸び終わりで見えないような絶対値の小さい極値を無視するために、 β と γ は(3)においてブックマーク増加が無くなる時刻に対応する添字 e_j を得るために用いる。

こうして得られた r_j と f_j を用いて、 p の j 番目の伸びを

$$g\text{-score}(p, j) = \frac{x_{f_j} - x_{r_j}}{f_j - r_j} \quad (1)$$

と評価する。

3.3 利用者の評価値を用いたウェブページの注目度予測

3.2 節で求めた(1)式を元に、利用者の評価を行う。

まず、個々の伸びにおいて利用者进行评估する。ページ p の j 番目の伸びについて、伸び以前の時間 $\gamma\tau$ 内に p をブックマークした利用者の集合を $U_{p,j}$ とする。すなわち、 p が最初にブックマークされた時刻を t_1^p として、

$$U_{p,j} = \left\{ u_k \mid (u_k, t_k, L_k) \in B_p^T \wedge t_1^p + (r_j - \gamma)\tau \leq t_k < t_1^p + r_j\tau \right\}$$

である。この式と(1)式を用いて、 p の j 番目の伸びにおける利用者 u の評価値 $u\text{-score}_{p,j}(u)$ を

$$u\text{-score}_{p,j}(u) = \begin{cases} g\text{-score}(p, j) / |U_{p,j}| & (u \in U_{p,j}) \\ 0 & (u \notin U_{p,j}) \end{cases} \quad (2)$$

と定義する。

次に、ある区間にブックマークされた全てのページについて利用者の評価値を集約し、全体としての利用者进行评估を行う。ここで同時に考慮したいのは利用者のブックマーク精度である。これは各利用者がブックマークした全てのページのうち、伸びたページの割合が大きい利用者と小さい利用者では、前者を高く評価したいという考えである。さらに、ブックマークしたペー

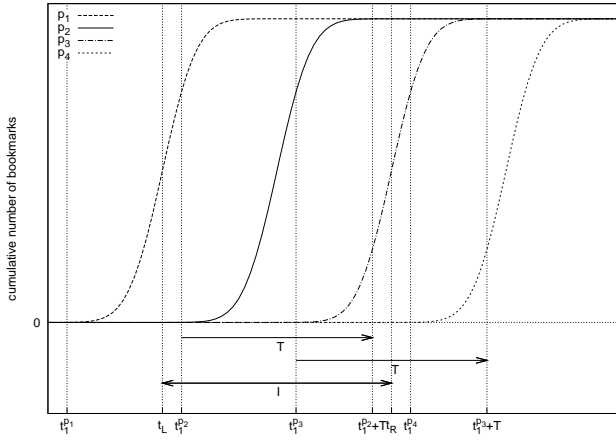


図3 T と I の関係

ジ数が少ない利用者は評価するための情報が不足していると考え、この段階で足切りを行う。

利用者 u が時区間 $I = [t_L, t_R]$ にブックマークした全てのページの集合を P_u^I とする。すなわち、

$$P_u^I = \{p \mid (u, p, t, L) \wedge t \in I\}$$

である。また、 I に初めてブックマークされ、かつ u がいずれかの伸びの直前にブックマークしていたページの集合を

$$P_u^{I,r} = \left\{ p \mid u \in \bigcup_j U_{p,j} \wedge t_1^p \in I \right\}$$

とする。これらの式と (2) 式を用いて、利用者の評価値 $u\text{-score}(u)$ を

$$u\text{-score}(u) = \begin{cases} \frac{|P_u^{I,r}|}{|P_u^I|} \sum_p \sum_j u\text{-score}_{p,j}(u) & (|P_u^I| \geq N) \\ 0 & (|P_u^I| < N) \end{cases}$$

と定義する。ただし N は足切りのための正整数パラメータである。

ここで T, I , および $P_u^{I,r}$ について補足する。図3はある四つのページ p_1, p_2, p_3, p_4 について T と I の関係を示したものである。 T は各ページの初ブックマーク時刻からどれだけの期間のブックマーク情報を利用するかを定めた値であり、 I は利用者の評価に用いるページを規定する区間である。 I 内に初めてブックマークされたページ(図では p_2, p_3) であれば、例え I 外のブックマーク情報であっても、 T で定められた範囲のものをを用いて利用者の評価を行う。ある利用者 u がこれら全てのページを伸び直前にブックマークしていた場合、 $P_u^{I,r} = \{p_2, p_3\}$ となる。

最後に、1. 節でも述べたように、利用者の評価値を用いてページの“注目度”，すなわちあるページ p が今後どの程度の注目を集めるかを次式で予測する。

$$\text{hotness}(p) = \frac{1}{t_{\text{now}} - t_1^p} \sum_{u \in U_p} u\text{-score}(u) \quad (3)$$

ここで U_p は p をブックマークしている利用者の集合、 t_{now} は現在時刻である。総和の前の分数により、注目度は時間の経過に伴って小さくなる。

3.4 利用者とタグの突出度を用いた注目度予測

本節ではタグを導入して 3.3 の注目度手法を拡張する。タグの導入は、現在流行の話題を考慮し、注目度予測に反映させるためである。ただし、一般的な意味を持つ単語のタグはソーシャルブックマークでは頻繁に使われているため、使用頻度が多いタグほど流行の話題を表しているとは一概には言えない。そのため、我々はタグの使用頻度ではなく“突出度”を考える。

右の端点が等しい、あるふたつの区間 $I_1 = [t_{L1}, t_R], I_2 = [t_{L2}, t_R]$ (ただし $t_{L1} < t_{L2}$) が与えられたとき、タグ l の突出度 $\text{prominency}(l)$ を

$$\text{prominency}(l) = \begin{cases} \frac{|I_1| |B_l^{I_2}|}{|I_2| |B_l^{I_1}|} & (|B_l^{I_1}| \neq \emptyset) \\ 0 & (|B_l^{I_1}| = \emptyset) \end{cases}$$

と定義する。ただし B_l^I は I において l が付与された全てのブックマークの集合、すなわち

$$B_l^I = \{(u, p, t, L) \mid t \in I \wedge l \in L\}$$

である。 l の突出度は、 l の I_1 に対する I_2 での l の使用頻度であり、 l が I_2 でのみ使用されている場合に最大値を取る。

ウェブページの注目度は、タグ集合 L のスコアを

$$\text{prominency}(L) = \begin{cases} \max_{l \in L} \text{prominency}(l) & (L \neq \emptyset) \\ 1 & (L = \emptyset) \end{cases}$$

として次式で予測する。

$$\text{hotness}(p) = \frac{1}{t_{\text{now}} - t_1^p} \sum_{u \in U_p} u\text{-score}(u) \cdot \max\{\text{prominency}(L_{u,p}), 1\} \quad (4)$$

(4) 式の $\max\{\cdot\}$ の項は最低でも 1 になる。これは突出度が高いという意味で良いタグを付与した利用者は高く評価し、そうでない利用者も最低限その利用者自身程度には評価するという考えに基づく。

4. 評価実験

3. 節で提案した手法の有効性を評価する実験を行った。

提案手法の (3) 式または (4) 式により、時刻 t_{now} における注目度の予測値が高いページの順位付き一覧を作ることができる。これを既存手法による一覧と比較することで提案手法の有効性を評価した。

比較対象は、はてなブックマークの注目エントリ一覧^(注1)とした。これは指定された閾値以上のはてなブックマーク利用者にブックマークされたページが、その最初のブックマーク日時で逆時系列順に上位から順位付けされた一覧である。閾値の既定値は 3 であり、今回の評価にもこの値を用いた。

評価手法は以下の通りである。まず現在時刻として想定したある t_{now} における両一覧を作成した。次に 3.2 節の手法を用

(注1): <http://b.hatena.ne.jp/entrylist?sort=hot>

表 1 パラメーター一覧

変数名	値
T	7 日
τ	1 時間
α	2
β	5
γ	6
N	28
t_{now}	03/01/07 00:00, 03/02/07 00:00, 03/03/07 00:00
I	$[t_{\text{now}} - 28 \text{ 日}, t_{\text{now}})$
I_1	$[t_{\text{now}} - 7 \text{ 日}, t_{\text{now}})$
I_2	$[t_{\text{now}} - 1 \text{ 日}, t_{\text{now}})$
K	100
T'	7 日

いて既にブックマークの伸びがあったページ集合を求め、両一覧からそれらのページを除外した。さらに我々の手法による一覧からはブックマーク数が 3 に満たないページを除外した。こうして各一覧に残った上位 K ページの集合を M とし、以下の四つの式で評価した。

$$\text{score}_1(M) = \sum_{p \in M} \text{bookmark}(p)$$

$$\text{score}_2(M) = \sum_{p \in M} \frac{\text{bookmark}(p)}{\text{rank}_M(p)}$$

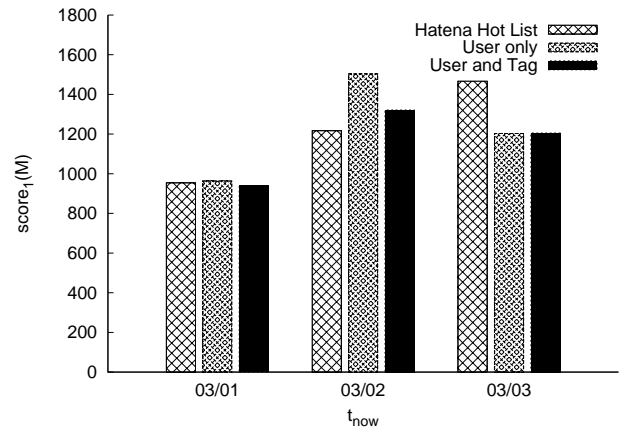
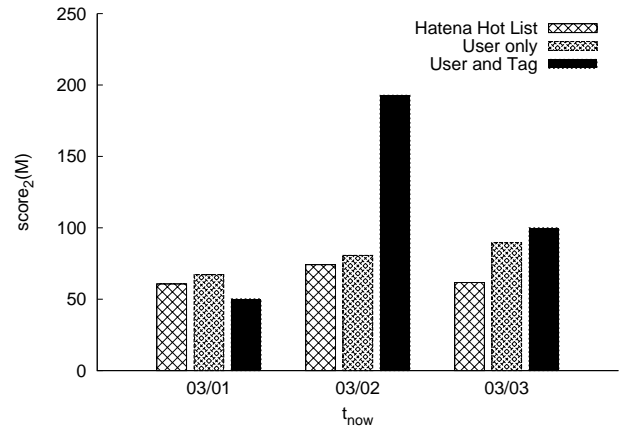
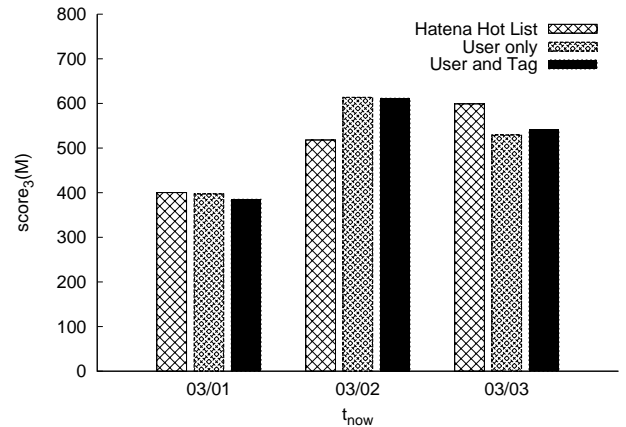
$$\text{score}_3(M) = \sum_{p \in M} \frac{\text{bookmark}(p)}{\log_{10} \text{rank}_M(p) + 1}$$

$$\text{score}_4(M) = \sum_{p \in M} \frac{\log_{10}[\text{bookmark}(p) + 1]}{\log_{10} \text{rank}_M(p) + 1}$$

ここで $\text{rank}_M(p)$ は M における p の順位、 $\text{bookmark}(p)$ は T' をパラメータとして区間 $[t_{\text{now}}, t_{\text{now}} + T')$ に p が獲得したブックマーク数である。いずれの式も、作成した一覧に含まれるページが実際に t_{now} 以降にどの程度ブックマークを集めたかによって手法を評価するものであり、値が大きいほど良い結果と言える。それぞれの式の違いはページの順位を考慮するか否か、考慮する場合、どのような重みを与えるか、である。

データセットとしてははてなブックマークを用い、パラメータを表 1 のようにして実験を行った。まず、2007 年 2 月 1 日から 3 月 2 日までに初めてブックマークされたウェブページ 200380 件を取得し、それぞれの初ブックマーク時刻から 7 日 ($= T$) 分のブックマーク情報を取得した。次に、表 1 に示した三つの t_{now} に対応する区間 I について、3. 節の手法を用いて利用者の評価値およびタグの突出度を求めた。最後に、はてなブックマークの注目エントリー一覧と 3.3 節および 3.4 節で述べた二つの提案手法によるページ一覧を生成し、 t_{now} から 7 日 ($= T'$) 分のブックマーク情報を取得して、本節で述べた四つの評価手法それぞれによって一覧を評価した。

実験結果を図 4 から図 7 までに示す。“Hatena Hot List” ははてなブックマークの注目エントリー、“User only” は 3.3 節の手法、“User and Tag” は 3.4 節の手法による結果である。図より、はてなブックマークの注目エントリーと 3.3 節の手法を比較すると、多くの場合 ($9/12 = 75\%$) において我々の手法に

図 4 $\text{score}_1(M)$ での実験結果図 5 $\text{score}_2(M)$ での実験結果図 6 $\text{score}_3(M)$ での実験結果

よる一覧が良い結果であることが分かる。同様に、はてなブックマークの注目エントリーと 3.4 節の手法を比較しても、我々の手法が良い結果を出している ($8/12 \approx 67\%$)。また、提案した二つの手法同士を比較すると、タグの突出度を用いない方が用いるよりも結果が良い場合が多く、タグを予測に反映させる手法にはまだ改善すべき点がある。

ここで、この評価手法は我々にとって不利なものであることを強調しておきたい。はてなブックマーク注目エントリー一覧

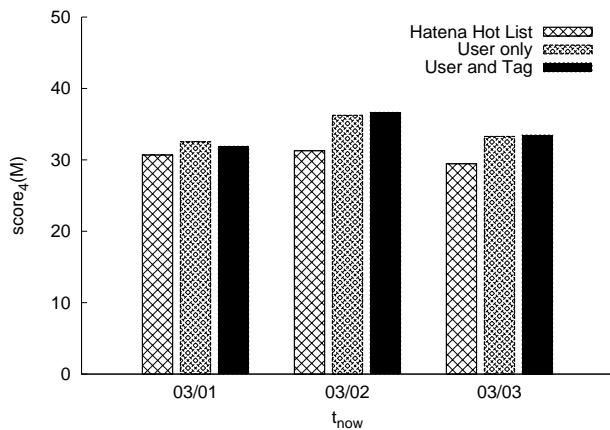


図 7 $score_4(M)$ での実験結果

の上位ページは実際にははてなブックマークのトップページに表示されるため、多くの利用者がそのページを見ることになるため、ブックマーク数も増えやすい。一方で我々の手法によるリストはそのような人目には晒されていない。にもかかわらず、提案手法がはてなブックマークに劣っていないことは我々の手法が既存手法と比較しても通用し、また、実際に本手法をサービスとして提供することでより良い結果が得られる可能性があることを示していると考ええる。

5. おわりに

近年台頭を見せるソーシャルブックマークは、利用者によって厳選された情報が集まるサービスとして注目されている。このサービスでは、ウェブページのブックマーク数によってページを順位付けし、利用者に提示している。しかし個々の利用者のブックマークの振る舞いやタグの流行など、まだ考慮されていない情報が多い。

本論文で、我々はブックマークの時系列情報を用いてウェブページの注目度を予測する手法を提案した。具体的には、個々のページについてブックマークの伸びというものを考えた。そして伸びの直前にページをブックマークしていた利用者をそのページの重要度を発見した利用者とし、みなして評価し、ウェブページの注目度予測に利用した。また、流行の話題をタグから捕らえるために、さらにタグの突出度を用いる予測手法を提案した。はてなブックマークのデータセットによる評価実験では、はてなブックマークの注目エントリー一覧と提案手法とを比較し、本手法の優位性を確認した。

今後の課題として、利用者の興味、関心を捕らえるための手法の提案がある。一般に一人の利用者には興味深い分野と浅い分野があり、それに応じて情報収集のふるまいも異なると考えられる。我々は個々の利用者がブックマークに付与したタグを用いてこの違いを捕らえることができると考えており、今後は手法の具体化と実験による評価を行う予定である。

文 献

- [1] “del.icio.us”. <http://del.icio.us/>.
- [2] “Hatena Bookmark”. <http://b.hatena.ne.jp/>.
- [3] L. Page and S. Brin: “The Anatomy of a Large-Scale Hy-

- pertextual Web Search Engine”, *Computer Networks and ISDN Systems*, **30**, 1-7, pp. 107-117 (1998).
- [4] Y. Xu, L. Zhang and W. Li: “Cubic Analysis of Social Bookmarking for Personalized Recommendation”, *Frontiers of WWW Research and Development – APWeb 2006: 8th Asia-Pacific Web Conference*, Harbin, China, January 16-18, 2006, Proceedings, pp. 733-738 (2006).
- [5] S. Niwa, T. Doi and S. Honiden: “Web Page Recommender System based on Folksonomy Mining for ITNG’06 Submissions”, *Third International Conference on Information Technology: New Generations (ITNG’06)*, pp. 388-393 (2006).
- [6] 丹羽, 土肥, 本位田: “Folksonomy マイニングに基づく Web ページ推薦システム (エージェント応用システム, 特集 マルチエージェントの理論と応用)”, *情報処理学会論文誌*, **47**, 5, pp. 1382-1392 (2006).
- [7] A. Sasaki, T. Miyata, Y. Inazumi, A. Kobayashi and Y. Sakai: “Web Content Recommendation System based on Similarities among Contents Cluster of Social Bookmark”, *DBWeb 2006*, pp. 59-66 (2006).
- [8] K. Yajima and U. Inoue: “Recommendation of Similar Documents and Users by Document Analysis on Social Bookmark”, *Proceedings of Data Engineering Workshop (DEWS) 2007* (2007).
- [9] A. Hotho, R. Jäschke, C. Schmitz and G. Stumme: “Information Retrieval in Folksonomies: Search and Ranking”, *The Semantic Web: Research and Applications; 3rd European Semantic Web Conference, Eswc 2006 Budva, Montenegro, June 11-14, 2006, Proceedings*, pp. 411-426 (2006).
- [10] A. Hotho, R. Jäschke, C. Schmitz and G. Stumme: “Trend Detection in Folksonomies”, *Semantic Multimedia: First International Conference on Semantic and Digital Media Technologies, SAMT 2006 Athens, Greece, December 6-8, 2006 Proceedings*, pp. 56-70 (2006).
- [11] Y. Yanbe, A. Jatowt, S. Nakamura and K. Tanaka: “Can Social Bookmarking Enhance Search in the Web?”, *Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2007)*, pp. 107-116 (2007).
- [12] M. Biddulph: “Introducing del.icio.us”, *XML.com* (2004). <http://www.xml.com/pub/a/2004/11/10/delicious.html>.
- [13] S. A. Golder and B. A. Huberman: “The Structure of Collaborative Tagging Systems”, *Information Dynamics Lab, HP Labs* (2005).