

ソーシャルブックマークに基づく情報発見

百田 信[†] 伊東 栄典[‡]

[†]九州大学大学院システム情報科学府 〒812-8581 福岡市東区箱崎 6-10-1

[‡]九州大学情報基盤研究開発センター 〒812-8581 福岡市東区箱崎 6-10-1

E-mail: [†] makoto.momota@i.kyushu-u.ac.jp, [‡] itou@cc.kyushu-u.ac.jp

あらまし blog や wiki, SNS, ソーシャルブックマークなどの様々なソーシャルサービスによって Web の質が変化し、これまで情報資源に対して傍観者であった利用者が作成者へと変貌を遂げた。これにより Web における情報資源は日々激増することになった。このようなソーシャルサービスには個人の嗜好に基づくデータが蓄積されており、そこから有益な情報を発見するシステムが求められている。本研究ではソーシャルブックマークに着目し、そこから情報を発見する二つの手法を提案する。一つは、自分の興味に近い利用者やページを発見する手法である。この手法では、利用者が保持するブックマークを個人のプロフィールとみなし、類似利用者および興味の近いページを発見する。もう一つは、新着情報をいち早く推薦する手法である。この手法は、人気の高いページを早期に見つけるブックマーカーを発見し、そのブックマーカーの見つけたページを推薦する手法である。本手法は早さだけではなく、分野の階層関係、自分とブックマーカーとの興味分野の近さも考慮している。上記の手法をシステムとして実装した。その際、ソーシャルブックマークデータとしてはてなブックマークのデータを利用した。またデータを用いて提案手法の評価を行った。

キーワード ソーシャルブックマーク, フォークソノミー, 情報発見, ブックマーカー

Information discovery based on Social Bookmarks

Makoto MOMOTA[†] Eisuke ITO[‡]

[†] Dept. Informatics, ISEE, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka-shi, Fukuoka, 612-8581 Japan

[‡] Research Institute for IT, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka-shi, Fukuoka, 612-8581 Japan

E-mail: [†] makoto.momota@i.kyushu-u.ac.jp, [‡] itou@cc.kyushu-u.ac.jp

Abstract The quality of Web was changed by recent social services such as blog, wiki, SNS, and Social bookmark. The user was shifting to creator of the information resource from onlooker. The amount of the information resources on the Web are rapidly increasing day by day. Social services accumulate folk's trends or individual preferences, then social bookmark data can use be an important data mining resource, and it may be possible to discover valid data for recommendation, marketing, and trend analysis. The authors propose two methods for information discovery from a social bookmark. The first method discovers similar users, and also discovers prefer pages which are preferred by a user. In this method, one user's bookmarks are considered as the profile of the user, and calculate similarity between users using profiles. The authors also propose page recommendation using user similarity. The second method recommends newly arrived information. This method discovers the alpha-bookmarkers who are an early detector of popular pages and also a spreader of the pages. This method considers not only earliness but also a hierarchy of interests. Especially, this method calculates neighborhoodness between a user's and an alpha-bookmarker in the specific interest. The authors implemented the two methods and examine using real data retrieved from "HATENA Bookmark". Examination shows that proposed methods are effective.

Keyword Social bookmarking, Folksonomy, Information discovery, Alpha-bookmarker

1. はじめに

近年, blog や wiki, SNS, ソーシャルブックマーク(以下 SBM)などの様々なソーシャルサービスによって Web の質が変化し、これまで情報資源に対して傍観者であった利用者が作成者へと変貌を遂げた。これにより Web における情報資源は日々激増することにな

った。また、情報資源はテキストのみならず画像や音声、動画など多岐にわたっており情報が溢れ返っている。このような情報爆発時代では利用者が求めている情報を素早く探ることが重要となっているが、日々増え続ける情報資源の中から探すのは困難であり利用者は探すことに明け暮れている現状がある。マルチメデ

ィア化する情報資源を効率的に検索するために情報資源を体系化する様々な試みがされている。その中に近年注目されている Folksonomy と呼ばれる体系化手法がある。Folksonomy は利用者がタグと呼ばれるキーワードを情報資源に付与することによって利用者のリアルタイムな要望を反映するという特徴を持つ。また、特定の情報資源に対して複数のタグを付与することも可能であり、これらから利用者が予想しえない未知の情報資源を発見することに長けている。

我々はそのような特徴を持つ Folksonomy に基づく SBM のデータを用い、情報発見を支援する2つの手法を提案する。一つは、自分の興味に近い利用者やページを発見する手法である。この手法では、利用者のブックマークをプロフィールとみなし、類似利用者および興味の近いページを発見する。もう一つは、今後流行する可能性の高い情報をいち早く発見する手法である。この手法は、人気の高いページを早期に見つけるブックマーカーを発見し、そのブックマーカーの見つけたページを推薦する手法である。本手法は早さだけではなく、分野の階層関係、自分とブックマーカーとの興味分野の近さも考慮している。はてなブックマーク[7]のデータを利用し、上記の手法をシステムとして実装した。

本論文の全体の構成は以下の通りである。まず2章で Folksonomy と SBM についての説明を行い、3章で関連研究の紹介する。次に4章でそれらを用いた情報発見支援についての手法を説明し、5章にて提案手法の実験方法および実験結果に基づく評価および考察を行い、最後に7章にてまとめを行う。

2. ソーシャルブックマーク

Web における情報資源を体系化するために、近年 Folksonomy と呼ばれる方法が盛んに利用されている。本章では Folksonomy とそれに基づく SBM について触れる。

2.1. Folksonomy

Folksonomy は collaborative tagging や social tagging, social classification などとも呼ばれる[1]。Folksonomy は利用者が情報資源に対してタグと呼ばれる自由なキーワードを付与しその情報資源を共有・分類・管理する Web ベースのシステムである。タグは情報検索や情報へのナビゲーション、情報発見のために利用され、例えば利用傾向の発見など利用者直接的な利益を与える。Folksonomy という用語は人々を表す"folk"と分類を表す"taxonomy"からなる合成語であり、利用者によって生成される概念構造を表している。著名なサービスが Folksonomy の扱う情報資源によって存在する。

例えば、写真を共有する flickr¹やフォト蔵²、ブックマークを共有する del.icio.us³やはてなブックマーク⁴、論文を共有する CiteULike⁵や Connotea⁶、動画を共有する YouTube⁷やニコニコ動画⁸がある。

Folksonomy は従来の Taxonomy に変わる新しい分類手法と言われている。従来の Taxonomy と比較したときの Folksonomy の利点と欠点を以下に挙げる。

■ 利点

・ 包括性

従来の Taxonomy では個々のエンティティは厳密に一つの分類語に属さなければならない。それに対し Folksonomy では、利用者が選んだ自由な用語を分類語(タグ)として付与する。個々のエンティティはそれら複数の分類語に属するため包括的であり多面的である。また、あらゆる利用者がそれぞれの要望を直接的に表現した自由なタグを付与するため、文化的・社会的・政治的バイアスを除いたすべての利用者の要望を反映する。

・ 低コスト

Taxonomy は専門家によって分類スキームの決定がなされ個々のエンティティを分類している。Taxonomy のようなトップダウン型分類を行うためには多くの訓練や高い技能が必要となる。Folksonomy では利用者が自由なキーワードで分類する。このとき複雑な操作は必要なく利用者なりの分類をするためコストが低い。

・ 新鮮さと柔軟性

利用者はタグ付与のための特別な技能は必要としないため即座に自由なタグ付けが可能である。この容易性や柔軟性によって新しい情報や情報事態の変化に対して検討や熟考する必要がないため素早く柔軟な対応が可能である。

■ 欠点

・ 言語的あいまいさ

Folksonomy では利用者のタグ付け方法に関する制限が一切ない。そのため、同じ意味を指すが表記の異なる同義語タグが存在する。例えば、複数形と単数形や口語形、頭文字形などが挙げられる。また、同じ表記で複数の意味を指す多義語タグが存在する。(図1)

・ 認識の境界

利用者のエンティティに対する意識レベルの

¹ <http://www.flickr.com/>

² <http://photozou.jp/>

³ <http://del.icio.us/>

⁴ <http://b.hatena.ne.jp/>

⁵ <http://www.citeulike.org/>

⁶ <http://www.connotea.org/>

⁷ <http://www.youtube.com/>

⁸ <http://www.nicovideo.jp/>

違いによって選ばれるタグが変わる。例えば、Javascript に関するある web ページに対して利用者 A は "programming" を付け、利用者 B は "javascript" を付けるという状況が考えられる。

・ **非階層構造**

Folksonomy は、タグ間に親子関係や兄弟関係がないフラットな構造となっている。そのため、情報が多い場合には下位階層へ辿り情報を絞ることや逆に情報が少ない場合には上位階層へ辿り情報を広げていくことができない。

・ **検索精度の低さ**

Folksonomy における上述した言語的問題や非階層構造などによって検索精度が低いという問題がある。

以上のことから Folksonomy は、正確な情報を検索することに向かないが、利用者独自の自由なタグによって思いがけない情報を発見でき、また利用者のリアルタイムな要望を反映しているため新鮮な情報を発見することに向いていることが分かる。

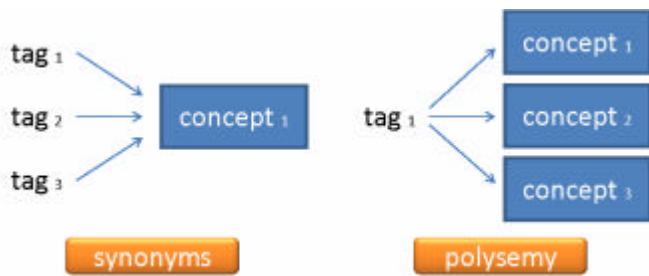


図1 タグの言語的問題

2.2. ソーシャルブックマーク

ブックマークとは、将来訪れる可能性のある Web サイトの URL をブラウザに保存する機能である。それに対し、Folksonomy の一つである SBM は、一つの Web サイト上で複数利用者のブックマーク情報を共有するサービスである。利用者はどんな URL でもブックマークとして保存でき、100 文字程度の短いコメントや自由なタグ付けが可能である。それぞれの利用者はタグによってブックマークを分類し管理、検索できる。タグによって複数利用者のブックマーク情報が関連付けられる。

2.3. ソーシャルブックマークのデータ構造

SBM における利用者集合を $USERS$ 、リンク URL 集合を $LINKS$ 、タグ集合を $TAGS$ と表すとき、SBM 情報は以下のような 4 つ組みで表現できる。

$$Bookmark(u, l, t, d)$$

ここで、 $u \in USERS, l \in LINKS, t \in TAGS, d$ は日付を表す。

3. 関連研究

Folksonomy についていくつかの議論や分析が行われている [2][3][4][5][6][8][9][10][11]。今までに行われた研究は主に、Folksonomy を情報体系化のための新しい分類法として着目している [2][4][10]。

Folksonomy に関する研究として長所を活かすアプローチと短所を補うアプローチがある。Folksonomy の短所として、タグ表記の自由さゆえの検索精度の低さが挙げられる。それに対して山家ら [5] や Hotho [8][9] らは検索精度向上を図っている。また、Rui [3] らや丹羽 [6] ら、Begelman [12] らはタグのクラスタリングを行うことによってタグが持つ言語的問題を解消している。本研究では長所である情報の新鮮さや意外性を活かすアプローチで情報発見の支援を行う。

4. ソーシャルブックマークからの情報発見

本研究では、SBM を使うことで情報発見の支援を行う。情報発見の指標として利用者に着目した。我々は 2 つの仮説を立てる。

仮説 1: ある利用者 A に類似した利用者 B が持つ情報は利用者 A の興味に近い情報である。

仮説 2: あるトピックに関する情報を他の利用者よりも早く発見できる利用者があるとき、その利用者が注目しているそのトピックに関する情報は将来的に流行する可能性の高い有用な情報である。

これら二つの仮説から情報発見を支援するための 2 つの手法を提案する。

4.1. 手法 1: 類似利用者手法

利用者は自分の興味に沿った情報を探している。ここで類似利用者が興味を示しているページは興味があると仮定する。そこで、ある任意の利用者 $u_a \in USERS$ についてのブックマークデータ $Bookmark(u_a, l, t, d)$ を利用者 u_a のプロフィールとみなし、利用者間の類似度を計算する。任意の利用者 u_a に対する類似利用者が保持するブックマーク URL で、 u_a がまだブックマークしていないものを後述する cosrank で順位付けて提示する。利用者 u が所有するブックマーク URL 集合を $link(u) = \{l_1, l_2, \dots, l_n\}$ とすると利用者間のコサイン類似度は以下ようになる。

$$\text{cos}(u_i, u_j) = \frac{|link(u_i) \cap link(u_j)|}{\sqrt{|link(u_i)| |link(u_j)|}}$$

u_a と他の利用者 $u_i \in USERS$ 間のコサイン類似度 $\text{cos}(u_a, u_i)$ においてその類似度の値で並べ替えたときの上位 N の利用者集合を $\text{sim}(u_a)$ とする。このとき提示候補 URL l_a は以下ようになる。

$$l_a \in \bigcup_{u \in \text{sim}(u_a)} link(u) \text{ and } l_a \notin link(u_a)$$

l_a に対して以下の方法でランキングを行う。

$$\text{cosrank}(l_a) = \sum_{u \in \text{sim}(u_a)} \cos(u_a, u)$$

$\text{cosrank}(l_a)$ の値の上位 k のブックマーク URL を提示 URL とする。

4.2. 手法 2: ブックマーカー手法

本研究では、利用者の中でも人気のある情報をいち早く入手できる利用者をブックマーカーと呼ぶ。ブックマーカーが持つ最新情報は将来的に人気がある可能性が高い。ここでは、あるトピックに関してのブックマーカーを探す手法を提案する。利用者によって自由なキーワードで付与されるタグはその利用者の興味を端的に示している。つまり、我々はタグがある種のトピックを表すと仮定する。

まず、タグの階層化を行い Folksonomy が持つ言語的問題の解消を試みる。それから、任意のトピックについてのブックマーカーを算出しそれに基づくブックマーク URL 提示アルゴリズムを提案する。

4.2.1. タグの階層化

Intersection Rate によってタグを階層化することによって Folksonomy の言語的問題を解決する。

本研究では term-document 行列より上位下位関係を求め単語間の階層構造を求める。単語 u, v の上下関係を u, v の出現する文書集合 $DF(u), DF(v)$ を用いた以下の Intersection rate の式で定義する。

u is upper term of v

$$\Leftrightarrow DF(u) \supset DF(v) \quad \text{and} \quad \frac{|DF(u) \cap DF(v)|}{|DF(v)|} > 0.5$$

本手法では単語をタグ t_i 、文書をタグ付けされたブックマーク URL $link(t_i)$ として階層構造を構築する。例えば "DB"-"Mysql" のような上位下位関係がある時、下位のタグ "MySQL" は上位のタグ "DB" よりも詳細なトピックを表している。逆に "DB" は "MySQL" よりも一般的であることになる。

4.2.2. ブックマーカーの算出

任意のタグ $t_{in} \in TAGS$ を選んだとき、 t_{in} 以下のタグ下位集合を $under(t_{in})$ とする。 $under(t_{in})$ はトピックとみなせる。その各ノード $t_i \in under(t_{in})$ が付与されたブックマーク URL 集合 $relink(t_{in})$ は以下のように表せる。

$$\text{relink}(t_{in}) = \bigcup_{t_i \in \text{under}(t_{in})} \text{link}(t_i)$$

すべての $l_i \in \text{relink}(t_{in})$ に関する $Bookmark(u, l_i, t, d)$ において早期にブックマークした利用者を探すため $Bookmark(u, l_i, t, d)$ を d で昇順に並べ替える。このとき、上位 $\%$ の利用者 u_i に対して重み $w(u_i, l_i)$ を付与する。利用者 u_i の度 $alpha(u_i)$ を以下のように定義する。

$$\alpha(u_i) = \sum_{l_i \in \text{relink}(t_{in})} \{w(u_i, l_i) - \text{rank}(u_i, l_i)\}$$

ここで、 $\text{rank}(u_i, l_i)$ は、時間軸で評価した利用者 u_i の URL l_i に対するブックマーク順位を表す。また、 $w(u_i, l_i)$ を $Bookmark(u, l_i, t, d)$ における $|u|$ とする。 $alpha(u_i)$ の大きい利用者上位 x 人を $under(t_{in})$ におけるブックマーカー $alphaBookmaker(t_{in}) \subseteq USERS$ とみなす。ブックマーカー $au_i \in alphaBookmaker(t_{in})$ が $Bookmark(au_i, l_i, t_{in}, d)$ を d で降順に並べ替えた上位 N の提示候補 URL 集合を $reclink$ とする。 $rl_i \in reclink$ に対して以下の方法で順序づけを行う。

$$\text{alpharank}(rl_i) = \sum_{u \in \text{alphaBookmaker}(rl_i)} \alpha(u)$$

ここで、 $alphaBookmaker(l)$ は URL l をブックマークとして保持しているブックマーカー集合を表す。 $\text{alpharank}(l_a)$ の値の上位 k のブックマーク URL を提示 URL とする。

5. 実験

提案手法について実際に存在する SBM に適用して実験を行った。実験 1 では、手法 1 の類似利用者手法についての評価を行い、実験 2 では手法 2 のブックマーカー手法についての評価を行った。

5.1. データ収集

今回は実験対象のデータとしてはてなブックマークを選んだ。はてなブックマークはこのサービスは現在日本最大規模の SBM サービスであり利用者数は 2006 年 10 月付で約 60,000 人⁹、登録 URL 数は 2007 年 1 月現在で約 7,000,000 ページ¹⁰が存在している。2007 年 5 月から 2007 年 6 月の間、以下の方法で収集した。

(1) Tag cloud から人気タグの抽出

"<http://b.hatena.ne.jp/>" は Tag Cloud ページになっており、そこから使用頻度の高い人気タグが取得できる。このページより人気タグを取得し DB へ格納した。

(2) タグ t_i が付与された URL 集合 $link(t_i)$ の取得

(1) で取得したそれぞれの人気タグが付与された URL を収集する。"http://b.hatena.ne.jp/t/t_i?sort=count" は、タグ t_i が付与された URL を被ブックマーク数の多い順に表示している URL 一覧ページである。そのページを収集し、ブックマーク URL とその URL のタイトルを抽出し DB へ格納した。

(3) URL l_i に関するブックマーク $Bookmark(u, l_i, t, d)$ における u, t, d の取得

(2) で収集した URL についてブックマークしている利用者、その利用者がその URL に対して付与したタグ名、

⁹ <http://d.hatena.ne.jp/naoya/20061020/1161314770>

¹⁰ <http://b.hatena.ne.jp/entry/7000001>

その利用者がブックマークした日付情報を取得する。
["http://b.hatena.ne.jp/entry/](http://b.hatena.ne.jp/entry/)より l_i をブックマークしているユーザーとそのタグ、日付情報のリストが得られる。ここで、(1)で収集した人気タグだけでなく他の使用頻度の低いタグも収集可能となる。

これらの結果として得られ詳細を表 1 に示す。

表 1 収集したデータ

項目	USERS	LINKS	TAGS	Bookmark
サイズ	42,753	350,279	108,753	6,648,994

5.2. 実験 1

類似利用者手法 (手法 1) を評価するために収集したデータの中で 2006 年 12 月 31 日以前のものだけを使用し手法 1 によるシステムを構築した。

2007 年 1 月 1 日以降にデータが存在する利用者の中から無作為に 1000 人の利用者 $RandomU$ $USERS$ を選び、その各利用者 ru_i $RandomU$ に対して手法 1 に基づいて k を $k = \{10, 20, \dots, 100\}$ と変化させ URL を提示した。その中で一つでも 2007 年 1 月 1 日以降のデータ内でその利用者によってブックマークされているものがあれば成功とみなした。本研究では、100 回 $RandomU$ を選び直し実験を行った。それらの平均をとったものを図 2 に示す。比較として $RandomU$ に対して無作為に選んだ URL を提示した結果も示している。x 軸は、システムが提示した URL 数 k を表し、y 軸は、成功人数を $|RandomU|$ で割った値を示している。成功率は 3% から 15% 程度ということが分かる。すべての k において、無作為な提示に比べ本手法が優れているため、本手法が有効であることが分かる。また、情報提示数 k と成功率に正の相関が見られる。これは情報推薦数が増加すると解集合の中に占める推薦集合の割合が増加するためと考えられる。

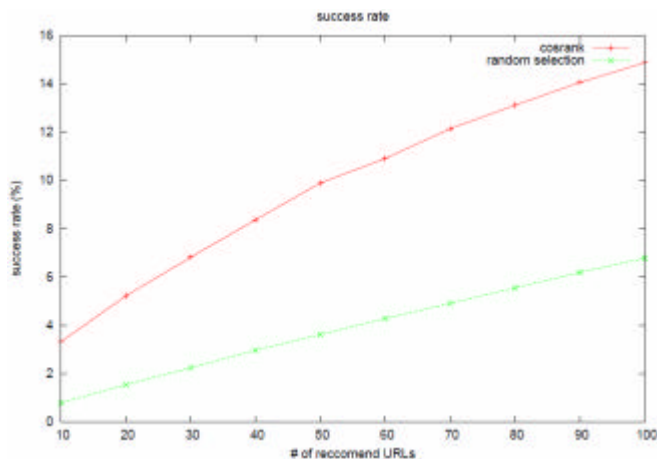


図 2 実験 1 の結果

5.3. 実験 2

収集したデータに対し、Intersection rate を計算しタグの階層化を行った。ここで、 $|TAGS|$ を N 、 $|LINKS|$ を M とするとすべてのタグ t_i $TAGS$ に対する Intersection rate を計算するとき、計算量は $O(MN^2)$ であり計算量が非常に大きいため、頻度が 50 以下のタグを除いて計算を行った。これにより N は 108753 から 6467 になった。図 4 にタグ "db" の下位となるタグ集合を graphviz によって出力したものの一部を示す。"db" の下位集合には "mysql", "sql", "RDB" などすべてデータベースに関連するタグが現われている。また、その中で "mysql" に注目してみる。"mysql" の下位集合をみると "LAMP" や "phpmyadmin" など mysql に関するより詳細なタグとなっている。多少定性的な評価ではあるが、図 5 に関して階層化は適切に行えている。またその他のタグについても同様に適切に行えていることを確認した。(タグの階層化に関する試作システムより¹¹)



図 3 "db" の下位集合の一部

ブックマーカー手法におけるパラメータを $\alpha=10\%$, $N=200$, $x=100$ に設定した。実験 1 同様に 2006 年 12 月 31 日までのデータを用いて手法 2 に基づいた情報提示を行った。入力タグ t_{in} を TF 値の上位から 20 個選び ("web", "javascript", "ネタ", "design", "google", "AJAX", "blog", "2ch", "css", "web2.0", "tips", "YouTube", "tool", "business", "Programming", "flash", "book", "perl", "社会", "あとで読む"), それらについて手法 2 を適用した。タグを入力する利用者はそのタグに興味がある、または、タグについての情報を求めていると考えられる。そこで、それぞれの t_{in} について、

¹¹ <http://shiva.cc.kvushu-u.ac.jp/~momota/hatena/hic/graph.cgi>

データ全体の中で t_{in} をタグ付けしている利用者の中から無作為に 1000 人の利用者 $RandomU(t_{in})$

$USERS$ を選んだ。それぞれの ru_i $RandomU(t_{in})$ について URL 数 k を $k = \{10, 20, \dots, 100\}$ と変化させ情報提示をした。その中で一つでもその利用者によってブックマークされているものがあれば成功とみなした。

ブックマーク手法の評価では、利用者によって 2007 年 1 月 1 日以前にもブックマークされていれば成功とする。これは、提示された情報が、すでにブックマークしてある場合もその利用者に対して興味分野のマッチングができているとみなせるためである。類似利用者手法と同様に 100 回の試行の平均を図 4 に示す。どの入力 t_{in} に対しても幅はあるものの成功率と情報提示数の関係には似たふるまいが観られる。

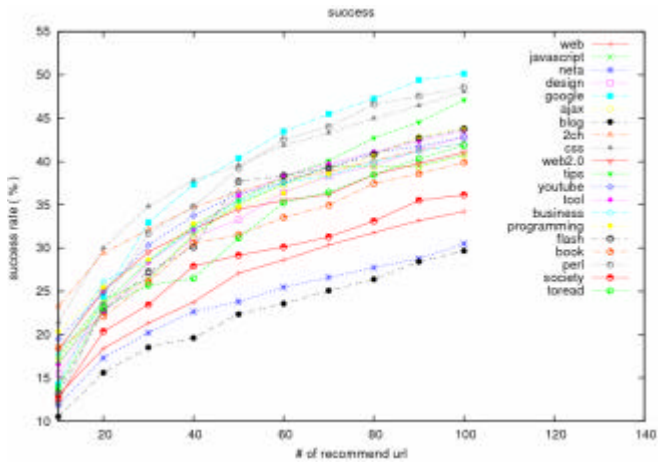


図 4 実験 2 の結果

また、実験 2 の結果の平均成功率と実験 1 の結果を比較したものを図 5 に示す。手法 1 に比べ、手法 2 は 3 から 4 倍の成功率を示している。つまり、トピックを限定したときの情報発見についてはブックマーク手法の方が優れていることが分かった。類似利用者手法は、利用者のプロフィールを使うので利用者がすでに保持しているブックマークを提示時に排除できるため、成功率が低くなったと考えられる。

6. おわりに

我々は Folksonomy の一つ SBM に着目し、利用者を軸とした情報発見支援手法を提案した。一つは、自分の興味に近い利用者および興味の近いページを発見する手法である。もう一つは、新着情報をいち早く推薦する手法である。この手法は、人気の高いページを早期に見つけるブックマークを発見し、そのブックマークの見つけたページを推薦する手法である。上記の手法をシステムとして実装した。その際、SBM データとしてははてなブックマークのデータを利用した。またデータを用いて提案手法の評価を行った。その結

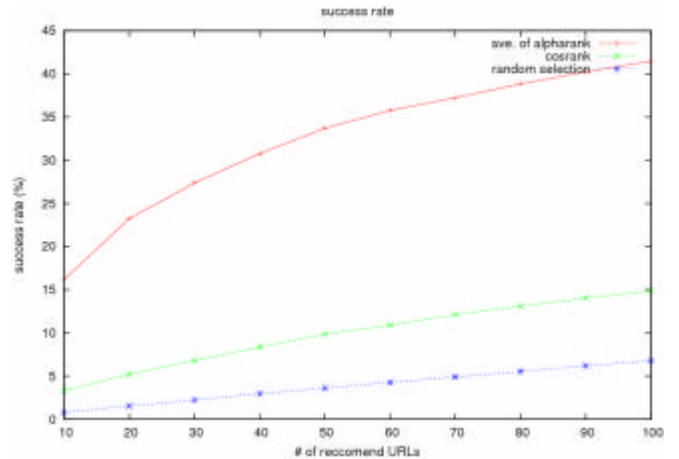


図 5 2つの手法の比較

果、トピックを限定した際には類似利用者手法よりもブックマーク手法の方が優れており効果的に情報発見の支援を行っていることが分かった。

文 献

- [1] “Folksonomy”: Wikipedia, the free encyclopedia, <http://en.wikipedia.org/wiki/Folksonomy>, (accessed 2007.12.26)
- [2] Golder S Huberman BA, “The Structure of Collaborative Tagging Systems”, Journal of Information Science 2006, 2006.
- [3] Rui Li, Shenghua Bao, Ben Fei, Zhong Su, Yong Yu, “Towards Effective Browsing of Large Scale Social Annotations”, Proceedings of WWW2007, 2007.
- [4] Emanuele Quintarelli, “Folksonomies: power to the people”, ISKO Italy-UniMIB meeting, <http://www.iskoi.org/doc/folksonomies.htm>, 2005.
- [5] 山家雄介, 中村聡史, Adam Jatowt, 田中克己, “Web 検索におけるランキング精度向上のためのソーシャルブックマークの利用”, DEWS 2007, 2007.
- [6] 丹羽智史, 土肥拓生, 本位田真一, “Folksonomy マイニングに基づく Web ページ推薦システム”, 情報処理学会論文誌, Vol. 47, No. 5. 2006, pp. 1382-1392.
- [7] はてなブックマーク, <http://b.hatena.ne.jp/>
- [8] HOTHO, A., JÄSCHKE, R., SCHMITZ, C. and STUMME, G, “Information Retrieval in Folksonomies: Search and Ranking”, ESWC 2006, 2006.
- [9] Robert Jäschke, Leandro Balby Marinho, Andreas Hotho, Lars Schmidt-Thieme, and Gerd Stumme. “Tag recommendations in folksonomies”. PKDD 2007, pp. 506-514, 2007.
- [10] A. Mathes. “Folksonomies - Cooperative Classification and Communication Through Shared Metadata”, Computer Mediated Communication, LIS590CMC (Doctoral Seminar), 2004.
- [11] Wu Harris, Zubair Mohammad, Maly Kurt, “Harvesting Social Knowledge from Folksonomies”, HYPERTEXT 06, Proceedings of the seventeenth conference on Hypertext and hypermedia,
- [12] Grigory Begelman, G. Keller, P. and Smadja, F., “Automated Tag Clustering: Improving search and exploration in the tag space”, Collaborative Web Tagging Workshop, 15th International World Wide Web Conference, Edinburgh, Scotland, 2006.