

## ソーシャルブックマークにおける登録共起関係を用いたブックマークの推薦

本間 正大<sup>2</sup>, 寺井 晃一<sup>1</sup>, Piyanuch Klaisubun<sup>2</sup>, 石川 孝<sup>1,2</sup>

1 日本工業大学, 2 日本工業大学大学院

**概要** 本論文では, あるブックマークに関連するブックマークをユーザに推薦することで情報発見の支援を行う推薦機能について述べる. ソーシャルブックマークはブックマークを検索する機能を提供しているが, ユーザはよりよい情報発見の支援を必要としている. そこで本論文では, ユーザの情報発見を支援するためにブックマークの登録共起関係を用いてブックマークを推薦する方法を提案し, それを実現した推薦機能を開発した. 本推薦機能の評価結果から, 本推薦機能は情報発見の支援に役立つという見込みが確認された.

### Recommendations of Bookmark Using Co-occurrence Relations of Posted Bookmark in

### Social Bookmarking Service

Masahiro Honma, Kouichi Terai, Piyanuch Klaisubun, Takashi Ishikawa  
Nippon Institute of Technology

**Abstract** The paper describes the recommendation function that supports the information discovery by recommending bookmark related to a certain bookmark to a user. The Social Bookmarking Service provides a function to search bookmark, but the user needs support of better information discovery. For this problem, we proposed a method to recommend bookmark by using co-occurrence relations of a posted bookmark for supporting information discovery and developed the recommendation function using the recommendation technique. The evaluation result of the recommendation function shows the possibility that the recommendation function is useful for supporting information discovery.

#### 1. はじめに

ソーシャルブックマーク (Social Bookmarking Service) は, ユーザが Web リソースをブックマークとして登録し, ユーザ間でブックマークを共有するオンラインサービスである. 一般にソーシャルブックマークは, 従来の検索エンジンなどの単なるキーワード検索のほか, ナビゲーションを提供している.

ソーシャルブックマークのナビゲーションは, ユーザが適切なブックマークを見つけるために, 次のどの Web ページを選択すべきなのかをユーザに案内する情報である. ナビゲーションが提供する情報は, 閲覧中のブックマークリストに関するユーザ名とタグ名の二つである. それらの情報はリンクで提供され, ユーザはそれらをクリックすることでユーザライブラリやタグにマッチするブックマークコレクションを表示することができる. キーワード検索がブル型情報サービスであるのに対して, ナビゲーションはプッシュ型情報サービスであり, 一種の推薦と

みなすことができる.

しかし, ソーシャルブックマークがこのような検索機能を提供している一方で, ユーザにとって情報発見を支援するような推薦システムの有用性が確認されている[10]. これは, ユーザが情報発見の支援に対して従来のブックマークを探すための機能だけでは満足しておらず, よりよい情報発見の支援を必要としていることを示している.

この要求に対して本研究では, 情報発見の支援のためにブックマークの登録共起関係を用いて関連性の高いブックマークをユーザに推薦する方法を提案する. ここでブックマークの登録共起関係とは, あるブックマークの登録が行われたとき, 同じユーザによって他のあるブックマークの登録も一緒に行われるという共起関係である. 本研究では, この共起関係をブックマーク間の関連とみなし, そのブックマークを関連ブックマークとしてユーザに推薦する機能を開発した.

本論文は, 第2章ではソーシャルブックマークに

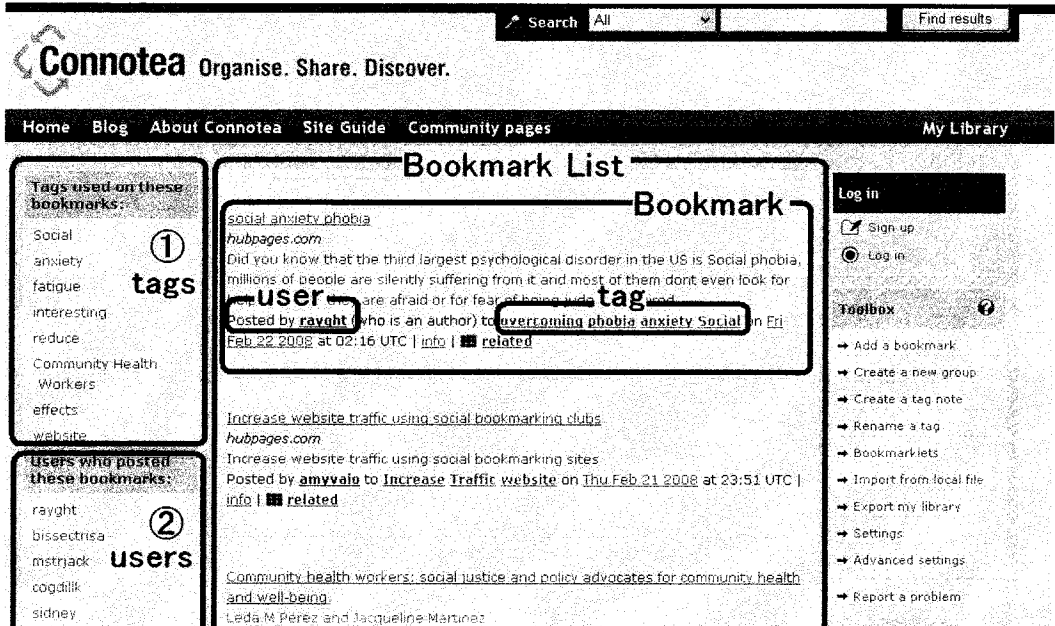


図1 Connoteaにおけるナビゲーション

におけるブックマークの推薦について説明し、第3章でアイテムの推薦に対して本研究と関連する研究について述べる。第4章では、情報発見の支援に対して登録共起関係を用いて関連ブックマークを推薦するアプローチと、その推薦アルゴリズムについて述べ、第5章では提案した関連ブックマークの評価実験の結果を示し、結果に対するの考察を述べる。最後に第6章では、まとめと今後の課題を述べる。

## 2. ブックマークの推薦

ソーシャルブックマークのナビゲーションにおいて、適切なブックマークをユーザに推薦することは、ユーザの情報発見の支援につながる。一般に、ソーシャルブックマークは、情報発見を支援するためにナビゲーションを提供している。ソーシャルブックマークのナビゲーションは、適切なブックマークを見つけるためにどの Web ページが選択されるべきかをユーザに案内するものである[7]。一般に、ソーシャルブックマークの Web ページは、特定の条件に対する検索結果、ユーザのブックマークライブラリ、タグにマッチするブックマークのコレクション、その他一般的な情報ページの4つのタイプに分類される。ブックマークに付けられるユーザ名とタグはどのページにも表示される。ユーザはそれらをクリッ

クすることでユーザライブラリやタグにマッチするブックマークコレクションを表示することができる。

学術文献のソーシャルブックマークサイト Connotea [4]でもナビゲーションを提供している(図1)。図1は、キーワード「Social Bookmarking」で検索した結果表示されたブックマークリストを示している。図中の中央に表示されている部分がブックマークリストである。ブックマークリスト中の各ブックマークには、そのブックマークを登録したユーザ名のリンクと、そのブックマークに付けられたタグのリンクが付けられている。図中①の部分「Tags used on these bookmarks」では、表示されたブックマークリスト中で使われているタグのリンクのリストが表示されており、図中②の部分「Users who posted these bookmarks」では表示されたブックマークを登録しているユーザ名のリンクのリストが表示されている。ユーザはこれらのリンクをクリックすることで、ユーザのライブラリやタグにマッチするブックマークコレクションを表示することができる。Connotea はここで例示したもの以外のナビゲーションも提供している。また、ナビゲーションはどのページにも表示されるため、ユーザはナビゲーションで提供されるリンクをたどってブックマークを探ることができる。

このように、ナビゲーションはユーザが受動的に情報収集できるプッシュ型情報サービスであり、一種の推薦とみなすことができる。

しかし、現在のナビゲーションは、あるブックマークに関するブックマークのコレクションといった、特定のブックマーク自体に着目したナビゲーションを提供していない。例えば、図1の①と②のナビゲーションは、閲覧しているブックマークのコレクションに対してブックマークのコレクションをナビゲートするが、特定のブックマークに対するブックマークのコレクションをナビゲートしていない。したがって、ユーザは特定のブックマークに関してしばらくでブックマークを探ることができず、情報発見に苦勞するため、ある特定のブックマークに関連するブックマークのコレクションをユーザに推薦する機能が必要である。

### 3. 関連研究

この章では、本研究に関連する研究の文献をレビューする。関連研究としては推薦システムがあげられる。推薦システムとは、利用者にとって有用と思われる対象、情報、または商品などを選び出し、それらを利用者の目的に合わせた形で提示するシステムである[9]。

#### 3.1 Amazon.comの推薦

Amazon.comの推薦技術[6]は、まず、一緒に購入された商品をデータベースに記録する。次に、オフラインでコサイン尺度によってアイテム間の類似度を計算してアイテムの類似度テーブルを構築する。そして、ある商品が表示されたときにその商品に類似している商品を推薦する。本研究で提案する推薦方法も、「一緒に購入される傾向に有るアイテムを類似しているアイテムとみなす」という考え方に基づいているが、Amazon.comがアイテムの類似度を計算するためにコサイン尺度を使うのに対して本研究は相関ルールのconfidenceを使う。また、本研究は類似度テーブルを構築せずにオンライン計算によってアイテムを推薦する方法を取る。

#### 3.2 データマイニング：相関ルール

データマイニング技術を用いた推薦システムは、相関ルール(association rule)を発見することでアイテムの推薦を行う[1, 3]。eコマースで一般的に用いられるデータマイニング技術は、同時購入されたアイテムの集合間の相関ルールを発見することに使われる。相関ルールのクオリティは、ルールにおけるパターンの出現頻度 support と相関の強さ

confidence によって評価される。そして、推薦システムはクオリティの高い相関ルールを用いて、ユーザにアイテムの推薦を行う。この推薦システムがアイテムを推薦することに、アイテムの集合間の相関を用いるのに対して、本研究が提案する推薦方法ではアイテム間の相関を用いる。

## 4. 関連ブックマーク

### 4.1 登録共起関係

本研究では、ブックマークを推薦する方法として、ブックマークの登録共起関係を用いてブックマークを推薦するアプローチをとる。ここでブックマークの登録共起関係とは、あるブックマークの登録が行われたとき、同じユーザによって他のあるブックマークの登録も一緒に行われるという共起関係である。

関連ブックマークの関連度 confidence は、相関ルールの confidence に基づいている。一般的に、 $m$ 個の製品の集合を  $P = \{P_1, P_2, \dots, P_m\}$  とし、1つのトランザクションの集合を  $T(T \subseteq P)$  とするとき、2つの製品の集合  $X$  と  $Y(X, Y \subseteq P, X \cap Y = \phi)$  の間の相関ルールは、トランザクション  $T$  に集合  $X$  の製品が存在するならば、同時に集合  $Y$  の製品も  $T$  に存在するという見込みを示している。一般に、相関ルールの confidence は次のように定義される[3]。

$$confidence = \frac{n(XUY)}{n(X)}$$

$n(XUY)$ : number of transactions containing XUY

$n(X)$ : number of transactions containing X

関連ブックマークの関連度 confidence は、この相関ルールの confidence に基づく。ただし、関連ブックマークは、ブックマークの集合  $X$  と  $Y$  が持つ要素を1つとし、 $n(XUY)$  はブックマークの集合  $X$  と  $Y$  の両方に含まれるブックマークを登録しているユーザの数、 $n(X)$  はブックマークの集合  $X$  に含まれるブックマークを登録しているユーザの数とした。そのため、confidence を以下のように再定義する。

$$Confidence = \frac{\text{number of users having both bookmark A and B}}{\text{number of users having bookmark A}}$$

### 4.2 関連ブックマークの推薦アルゴリズム

ここでは、関連ブックマークの推薦アルゴリズムについて述べる。関連ブックマークの推薦アルゴリ

← COPY

Social Bookmarking Tools (I): A General Review (info)

<http://www.dlib.org/dlib/april05/hammond/04hammond.html>

Posted by [wimvi](#) and 212 others to [Web 2.0 social bookmarking](#) on Thu Jan 03 2008 at 00:00 JT

Co-posteds with confidence = 137 / 283

①

← COPY

Folksonomies - Cooperative Classification and Communication Through Shared Metadata (info)

<http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>

Posted by [davidfowler](#) and 108 others to [lisistudy](#) on Sat Jan 05 2008 at 00:00 JT

Co-posteds with confidence = 47 / 283

②

図2 関連ブックマークのリスト

ズムは、以下のとおりである。

あるブックマーク A と推薦する数 N を入力とする  
ブックマーク A を登録したユーザ数を計算  
For サイトに登録された各ブックマーク B (A≠B)  
A と B の両方を登録したユーザ数を計算  
confidence を計算する  
ブックマーク B を confidence で降順にソート  
トップから N 個を関連ブックマークとして出力する

ただし、推薦する関連ブックマーク B は、ブックマーク A と B の両方を登録したユーザ数が二人以上のものとした。関連ブックマーク機能は、このアルゴリズムを用いて、N 個のブックマークをその confidence とともにユーザに推薦する。

#### 4.3 関連ブックマークのインタフェース

関連ブックマークは、ナビゲーションとして提供される。具体的には、ブックマークリスト中の各ブックマークに対して、関連ブックマークのリンク (Co-posteds) が表示される (図2)。ユーザは、そのリンクをクリックすることでそのブックマークに関連するブックマークのリストを表示することができる。ブックマークのリスト中の各関連ブックマークには、リンクをクリックしたブックマークと関連ブックマークとの関連度 confidence が表示される。図2は、「Social Bookmarking Tools (II): A Case Study - Connotea」というタイトルのブックマーク A に関連するブックマークリストの一部を示している。図

中①と②が、ブックマーク A と各関連ブックマーク B との関連度 confidence を示している。また、図中①と②では Co-posteds リンクが表示されており、ユーザはそれをクリックして、関連ブックマーク B に関連するブックマークリストを表示することができる。関連ブックマークのリストは confidence の高い順にソートされてユーザに提供される。

#### 4.4 関連ブックマークのメリット

関連ブックマークのメリットは、次の3つである。

- (1) 特定のブックマークに関するブックマークコレクション:従来のナビゲーションがブックマークのコレクションに関するブックマークのコレクションをナビゲートするのに対して、関連ブックマークは特定のブックマークに対してブックマークコレクションをナビゲートする。したがって、検索の範囲を特定のブックマークに絞ることができる。
- (2) 関連度というブックマーク選定の指標:今まではブックマークのタイトルや付けられたタグまたは登録したユーザなどがブックマーク選定の指標であったが、関連度という利用価値の高い指標が加わった。
- (3) 関連度によるランキング:従来のナビゲーションが提供するブックマークのリストは、登録日付の若い順にランキングされる。しかし、関連度の高いブックマークから順にランキングして表示することで、ユーザの情報の発見にかかるコストを削減できる。

表1 ブックマークの推薦機能の比較

Information Service Type	#Click Through	#Conversion	%Conversion Rate
<b>Keyword Search (pull-based information service)</b>			
Search	182	59	32%
<b>Navigation (push-based information service)</b>			
<b>Co-posteds</b>	<b>169 (119)</b>	<b>52</b>	<b>31% (44%)</b>
Users who used <tagname>	22	10	45%
Browsed users	11	4	36%
Users who posted the bookmarks	22	3	14%
Tag used in the bookmarks	29	15	52%
Related tags	36	15	42%

また、関連ブックマークの推薦アルゴリズムに登録共起関係を用いたメリットは次の二つである。

- (1) オンライン計算：単純な共起関係を使っているので計算コストがあまりかからず、協調フィルタリングなどの推薦システムなどでは提供できないオンラインでのリアルタイムな計算が可能になった。
- (2) わかりやすい推薦理由の説明：関連度は単純な式で定義されているので、何故そのブックマークが推薦されたのか、推薦の理由をユーザにわかりやすく説明することができる。

## 5. 評価実験

### 5.1 実験目的

実験の目的は、被験者が研究活動を行う上で参考にする文献を探すことにおいて、さまざまなナビゲーション機能の一つとして関連ブックマークの推薦機能を比較評価することである。

### 5.2 実験方法

実験の方法は次のとおりである。使用する実験システムは、実験参加者だけが利用できる学術文献のソーシャルブックマーク実験サイト ReMarkables2 [8] である。ReMarkables2 は Connotea Code [5] をベースに実現されている。使用するブックマークデータは、ソーシャルブックマークサイト Connotea [4] の 389,597 件(2004 年 12 月 9 日から 2008 年 1 月 10 日現在までに登録されたブックマークの件数)のブックマークから、被験者が探している文献のキーワードで絞り込んだ 76,536 件のブックマークデータである。今回は、Connotea の 39 万件のデータをすべて使うとシステムの応答速度が非常に遅くなるため、キーワードを使った絞り込みによるデータの削減によってシステムの負荷を軽減する処置を行った。情報発見

に対する効果の測定には、コンバージョンレートを用いる。コンバージョンは、ブックマークのコピー行動とする。

実験期間は 2008 年 2 月 9 日から同年 2 月 17 日である。被験者は、大学院生以上 7 名、学部生 4 名の計 11 名である。被験者には、サイトに登録されたブックマークから読みたい文献または既に読んだ文献のブックマークを探して、15 件程度コピーしてもらう。コンバージョンは、データベースに蓄積したアクセスログを用いて計算する。

### 5.3 実験結果

実験の結果は、表 1 にまとめられる。まず、各項目について簡単に説明する。

- **Search** : キーワード検索を提供する。ユーザはあるキーワードで、フリーマッチ検索やタグによる検索やユーザによる検索などを行うことができる。
- **Co-posteds (関連ブックマーク)** : 特定のブックマークに関連するブックマークを提供するナビゲーション。ブックマークにつけられた Co-posteds リンクをクリックすると、そのブックマークに関連するブックマークリストを表示することができる。
- **Users who used <tagname>** : ブックマークリスト中のブックマークに付けられたタグを使っているユーザ名のナビゲーション。このユーザ名のリンクをクリックすることで、そのユーザのライブラリのそのタグにマッチするブックマークのコレクションを表示することができる。
- **Browsed users** : 相互閲覧関係にあるユーザ名のナビゲーション。自分が閲覧したライブラリの所有者名や、自分のライブラリを閲覧し

たユーザ名, または相互閲覧の関係にあるユーザ名が表示される。閲覧関係は, 矢印の向きで表現される。

- **Users who posted the bookmarks** : ブックマークリスト中のブックマークを登録したユーザ名のナビゲーション。
- **Tag used in the bookmarks** : ブックマークリスト中のブックマークに付けられているタグのナビゲーション。このタグのリンクをクリックすることで, そのタグにマッチするブックマークのコレクションを表示することができる。
- **Related tags** : ブックマークリスト中のブックマークに関連するタグのナビゲーション。

表 1 は各項目に対するクリックスルーとコンバージョンレートを示している。クリックスルー(**#Click Through**)は各ナビゲーションのクリック回数であり, コンバージョン(**#Conversion**)は各ナビゲーションで表示される Web ページにおいてコピー行動が起こった回数である。コンバージョンレート(**%Conversion Rate**)は, クリックスルーに対するコンバージョンの割合である。関連ブックマークの推薦効果は 31%程度であり効果の高い他のナビゲーションにはやや劣るということを示している。ただし, 特殊なケースをのぞいた場合は括弧の部分の 44%程度になる。特殊なケースとは, 関連ブックマークのナビゲーションは提供されたが, ナビゲーションのリンクのクリック後に関連ブックマークが一つも表示されないケースである。このケースをのぞいた場合は, 効果の高い他のナビゲーションと同程度の効果を示している。

#### 5.4 考察

関連ブックマークは, 情報発見の支援に対して効果の高いナビゲーションと同程度の効果があるとみなせる。なぜならば, ユーザインタフェースを改善することで 44%の推薦効果が見込めるからである。したがって, ナビゲーションとしては十分に機能しており, 情報発見の支援に対して役立つという見込みが確認された。また, 今回関連ブックマークのコンバージョンレートの低かった原因は, Connotea のオリジナルデータを絞り込んで削減したデータを使ったために, ブックマークの登録の共起があまりおこらず, 登録共起関係を用いた推薦アルゴリズムでは関連度の弱いブックマークしか提供できなかったことが挙げられる。これらの問題については, ユーザインタフェースの設計と推薦アルゴリズムについて

検討する必要がある。

また, 関連ブックマークの評価は, コンバージョンレートの客観的なデータによる評価だけでなく, ユーザによる評価も必要である。なぜならば関連ブックマークでは推薦するブックマークをランキングして提供したが, かならずしも上位にくるものがユーザにとって有用な情報であったとは限らないからである。そのため, 推薦されたブックマークのランキングに対してユーザに評価を行ってもらうなどユーザによる評価が必要である。

#### 6. おわりに

ソーシャルブックマークにおける登録共起関係を用いたブックマークの推薦機能を実現した。この推薦機能の評価実験の結果は, 情報発見の支援に対して役立つという見込みを示した。今後の課題は, ユーザインタフェースの設計, 推薦アルゴリズムの検討, およびランキングして推薦したブックマークに対するユーザの評価実験などのユーザによる評価を行うことである。

#### 参考文献

- [1] Agrawal, R., Imclinski, T., Swami, A. (1994). Fast Algorithms for Mining Association Rules. In *Proceedings of the 20<sup>th</sup> VLDB Conference*. pp.487-499
- [2] Amazon.com. <http://www.amazon.com>
- [3] B.M. Sarwam et al., "Analysis of Recommendation Algorithms for E-Commerce," *ACM Conf. Electronic Commerce*, ACM Press, 2000, pp.158-167.
- [4] Connotea. <http://www.connotea.org>
- [5] Connotea Code. <http://www.connotea.org/code>
- [6] Greg Linden, Brent Smith, and Jeremy York. "Amazon.com recommendations: Item-to-item collaborative filtering." *IEEE Internet Computing, January 2003*.
- [7] Ishikawa, T, Klaisubun, P, Honma, M. "Navigation Efficiency of Social Bookmarking Service". In *Proceedings of IPSJ-GN<sup>2</sup> 2007, Japan, November 2007*.
- [8] ReMarkables2. <http://remarkables.nit.ac.jp/remarkables2>
- [9] 神寫敏弘. 推薦システムのアロリズム(1). 人工知能学会誌, Vol. 22, No. 6, 2007.
- [10] 本間 正大, 銭正宇, Piyanuch Klaisubun, 石川 孝: 研究支援のためのソーシャルブックマッキングサービスの開発, *IPSJ-GN63*, pp.73-78 (2007)