Title	Personalized retrieval in social bookmarking
Authors(s)	Bateman, Scott, Muller, Michael J., Freyne, Jill
Publication date	2009
Publication information	Bateman, Scott, Michael J. Muller, and Jill Freyne. "Personalized Retrieval in Social Bookmarking." Association for Computing Machinery, 2009.
Conference details	Presented at the 2009 ACM SIGCHI international conference on supporting group work, May 10–13 2009, Sanibel Island, Florida
Publisher	Association for Computing Machinery
Item record/more information	http://hdl.handle.net/10197/1249
Publisher's version (DOI)	10.1145/1531674.1531688

Personalized Retrieval in Social Bookmarking

Scott Bateman
Department of Computer Science
University of Saskatchewan
110 Science Place, Room 176
Saskatoon, SK, Canada
+1 306 966 4886

scott.bateman@usask.ca

Michael J. Muller
IBM Research / Collaborative User
Experience
One Rogers Street
Cambridge, MA, USA 02142
+1 617 693 4235

michael_muller@us.ibm.com

Jill Freyne
CLARITY: Centre for Sensor Web
Technologies
University College Dublin,
Belfield, Dublin 4, Ireland
+353 1 7165353

jill.freyne@ucd.ie

ABSTRACT

Users of social bookmarking systems take advantage of pivot browsing, an interaction technique allowing them to easily refine lists of bookmarks through the selection of filter terms. However, social bookmarking systems use one-size-fits-all ranking metrics to order refined lists. These generic rankings ignore past user interactions that may be useful in determining the relevance of bookmarks. In this work we describe a personalized ordering algorithm that leverages the fact that refinding, rather than discovery (finding a bookmark for the first time), makes up the majority of bookmark accesses. The algorithm examines user-access histories and promotes bookmarks that a user has previously visited. We investigate the potential of our algorithm using interaction logs from an enterprise social bookmarking system, the results show that our personalized algorithm would lead to improved bookmark rankings.

Categories and Subject Descriptors

H3.3. Information storage and retrieval: Information filtering; H5.4: Hypertext/Hypermedia. – User issues.

General Terms

Algorithms, Human Factors.

Keywords

social bookmarking, personalization, information retrieval

1. INTRODUCTION

Social bookmarking systems have become popular applications for people to store, retrieve and share personal bookmarks. Bookmarks are uploaded to a bookmarking service, annotated with keywords (called *tags*), and shared for all users to see. While users of bookmarking systems have varied goals for their contributions, which can include community building and publishing [11], the primary reason for a user to contribute content to a bookmarking system is to facilitate refinding at a later time, using the tags as hooks for recalling bookmarks [9]. Studies

of web browsing behaviours have reported that between 50% [2] and 80% [10] of webpage visits are to pages previously visited by a user. Here we present, for the first time, results that show a conservative estimate of refinding in the social bookmarking domain is 59%.

Pivot browsing characterizes the lightweight interactions that are common for exploration and discovery in social bookmarking systems. It allows users to "reorient" their view of bookmark lists, by clicking on different tag artifacts [9]. Pivot browsing in effect is equivalent to providing a text-based filter or query refinement, but allows users to more quickly explore the information-space by alleviating the need to recall and correctly type a tag or username. Users can see not only their own tags but tags entered by other users, so pivot browsing also supports discovery by following the tag links created by other users. However, this technique encourages users to enter queries one piece at time (by clicking on a single tag link), resulting in several refinement steps in order to complete a single query. Further, the result lists returned after each refinement step are ordered in generic ways that do not take into account the possible relevance of bookmarks for the user.

Our research investigates an area that has received little attention in the past: improving ordering algorithms to assist users in finding content in social bookmarking systems. We present a technique that exploits the viewing actions of a searcher as a means of judging relevance; thus the probability the searcher will select a bookmark. Our ordering algorithm coupled with the pivot browsing model can reduce the effort required by users to refind bookmarks. We demonstrate this using the logs of an enterprise social bookmarking system, Dogear [9], and report on the savings in terms of reduced browsing steps and improved result list positioning.

2. SOCIAL BOOKMARKING

Social bookmarking systems allow users to save URLs, and associated metadata, they wish to revisit or share with others to a central repository. A bookmark has 4 components, a **URL**, a **user**, a set of **tags** and a **timestamp.** Each of these components is often used in retrieval; tags and user profiles as filtering criteria, and URLs and timestamps in ordering metrics. Systems such as delicious (http://delicious.com), furl (http://furl.net) and Dogear (https://www.ibm.com/dogear) share similar interfaces for displaying bookmark lists, by using simple ordered lists. By default, bookmark lists are ranked by recency or by popularity (the number of times a URL has been bookmarked), or by some other hybrid method [3].

Filters can be achieved in two ways. First, users can type queries into search boxes, which match tags or usernames. Secondly, users

may filter their current view by pivot browsing. This filter is achieved by the selection of a single refinement tag or username from a list or cloud. We refer to the selection of each filtering criterion as a *refinement step*, which produces a sublist of results (or *result list*) based on the current criteria.

For example, in Figure 1, a filter query has been performed for the tag 'collaboration', which matches 5,545 bookmarks (only the top two results are shown). For each bookmark displayed, "hooks" for pivot browsing are provided. The query could be further refined by clicking the username 'Ling Shin', resulting in all bookmarks that 'Ling Shin' has created with the tag 'collaboration' being displayed. The pivot browsing style of interface makes query refinement simple; however, it requires multiple steps to reduce result lists to a manageable size.

Figure 1. The first two results for the tag 'collaboration'.

Even with the opportunity for simple and expedient query refinements, social bookmarking systems suffer with the problem of information overload. As the size of social bookmark collections grow, so too do the number of possible matches for a given query and thus the number of single-step refinements required to retrieve a manageable set of possible bookmarks. This is further compounded by the convergence of vocabulary used in tagging systems [1,6,9], where limited vocabulary is applied to a large corpus making relevance decisions more difficult and options for refining filters more limited. Our analysis of a large set of user logs (detailed in Section 4) showed that 70% of successful retrieval sessions involved multiple refinement steps.

Social bookmarking systems typically present results in order of the most recently added URLs (date-based) or the most frequently bookmarked URLs (popularity). We believe that the order in which bookmarks are presented in lists to users is of great importance in helping users effectively find and refind bookmarks of interest. Studies confirm that users typically make selections from the top of result lists, as shown by Keane, et al. [7]. Keane found that users selected the top ranked result over 70% of the time in search engine result lists, and that the second result was selected only 10% of the time. They showed that a bias existed even when search results were presented with the least relevant results at the top of the list. This work shows the importance of presenting results in an ordering appropriate to the user; as users seem, in part, to select results based on this ordering.

While Web search engines are able to make relevancy judgments based on web-scale measures, current approaches in social bookmarking systems have been limited to data contained within the system itself. Bookmarking systems generally do not incorporate measures of relevance – other than what might be new or popular at a system-wide level. Therefore, finding 'relevant' bookmarks within bookmarking systems is often difficult.

Recent work in the area of enhancing web search has concentrated on harnessing previous interaction data in order to improve the relevance of search results by providing a level of personalization. The collaborative Web search technology monitors user communities as they search [4]. It notes each query and resulting

page selection to model community preferences, and it makes predictions about the probability of a page being selected. In contrast, Liu et al. map queries to predefined groups [8]. ASSIST [5] monitors search and browsing patterns in order to recommend relevant pages and paths through a repository of information.

We initially examined Dogear logs for usage patterns that could be used to inform new social bookmark relevancy algorithms. We discovered that the majority of bookmark selections were revisits – bookmarks that a user had visited before or were already in the user's personal collection. Our analysis of the interactions of 10,755 users showed that within this bookmarking system 59% of bookmark selections were revisits, which suggests that re-finding behaviour is a typical usage pattern in social bookmarking.

3. PERSONALIZED RANKING

Social bookmarking systems aim to use the contributions of all users for the benefit of each individual user – the wisdom of the crowd. As such, we initially concentrated on examining the interaction history of the user population as a whole to improve result list orderings; for example, by ranking bookmarks according to accesses over all users. However, we did not see any improvement in result list orderings. We also attempted to exploit smaller groups of similar users. Using cosine similarity metrics to determine cliques, we used clique activities and accesses to determine bookmark relevancy, but the results were not encouraging. For reasons of space we have not included these results here. Following these results we refocused on individual personalization, only considering an individual searcher's previous actions as an alternative to community-based reordering metrics.

The proposed technique exploits an individual's previous selection history to make relevancy judgments on bookmarks. Consider a bookmark, $bkmk_j$, and user, u_i. The relevance of $bkmk_j$ to u_i is the relative number of times that $bkmk_j$ has been selected (viewed) by user u_i in the past (denoted as $selected_{ij}$), as a percentage of all the selections the user has made previously (see Eq 1).

$$relevance(user_i, bkmk_j) = \frac{selected_{ij}}{\sum_j selected_{ij}} \qquad (1)$$

With our personalized ranking metric, all bookmarks that have been previously selected by a user are identified as having a higher probability of being relevant to the user's current need. Our personalized ranking metric moves previously selected bookmarks toward the top of the list in order of decreasing relevance, with bookmarks that have not been selected appearing below, in reverse chronological order. Bookmarks with the same relevance score are ranked using a reverse chronological order.

Consider a situation where a user remembers encountering a tutorial on the programming language, Java, in a social bookmarking system. They may have added the bookmark themselves or found it by browsing other people's entries in the system's collection. To refind the tutorial bookmark, the user must enter some keyword(s), possibly through pivot browsing, and navigate to the bookmark. Assume the user selects the tag 'java' as their initial query. Many thousands of results would be returned. The user must either go through many result pages, or refine their query; in either case the user must exert more effort. Our personalized algorithm would automatically identify from the pool of bookmarks returned by the system those bookmarks that a user

has previously viewed, and bring those to the top of the list, thus eliminating the need for further query refinements.

4. EVALUATION

To evaluate our technique we used a dataset from the Dogear Social Bookmarking System [9]. This system has been in use since July 2005, it has over 10,000 users and 425,000 bookmarks. Access to Dogear is available to IBM employees through an intranet user interface, and to several internal systems via an intranet API. In this work we concentrate our analysis on events using the web-based user interface only.

The aim of the experiment was to retrospectively determine whether our personalized ranking algorithm would have made it easier for a user to locate a target bookmark. We hypothesized that a ranking metric more tuned to the use case of the searcher would produce result lists with target bookmarks closer to the top. If the lists presented during refinement were more relevant for the searcher, we could reduce the number of refinements required to locate a relevant bookmark, thus reducing user effort.

4.1 Dataset

Dogear activity logs contain all user actions with timestamps, for the history of the system's deployment. We isolated a set of 6 months worth of search activity from late 2007 as a set of 22,271 sessions for evaluation (we will refer to these as *query sessions)*. These query sessions are composed of a series of queries and refinements in which users sought and found a target bookmark, and made up approximately 70% of all sessions. The other 30% were query sessions where users restarted queries with all new query terms or quit searching. We assumed that all of the query sessions in the logs were viewed by the system's default ordering (reverse chronological). The log information detailed the number of results seen per page and the number of pages viewed by the searcher in each session.

For each query session we generated two result lists. The first list is a recreation of the actual result list seen by the user (date-based ordering). The second list is the result list that the user would have seen if our algorithm had been in use, using the selection history of the user (personalized ordering). We compared both the number of refinement steps and the actual rank of the bookmark in the two lists. 89% of sessions produced a result list that contained at least one bookmark that the searcher had previously accessed, and thus had the potential for personalization through our algorithm. Our experiments incorporated all sessions regardless of the possibility for personalization. We did this to get accurate results reflecting the effectiveness of our algorithm had it been originally used.

5. Results

Our first metric for determining the effectiveness of our algorithm relates to the position of the target bookmark (the selected bookmark) in each of the result lists. Overall (see Figure 2a), the default date-based ordering had an average rank for the user-selected bookmark of 37.8 and a median rank of 5. We were surprised to find that in many sessions (72%) the target bookmark was presented to the user in the initial result list (after the initial filter), however, the user continued to refine their query through pivot browsing. We hypothesize that the result was not ranked highly enough in the lists of bookmarks to grab the user's attention (lists were of length of 10, 25 or 50, based on a user-set preference). Our personalized ordering improved the average position of a target selection by 16.2 positions, bringing the target bookmark to a higher average position of 21.6 and median of 3.

Overall, the personalized ordering metric decreased the median position for selected results by 2 positions over the default date-based ordering. Reducing the mean by 2 might seem like a small change, however, when we consider this in terms of selection bias in search engine results the improvement is sizeable. By increasing the ranking of a target bookmark we increase the likelihood that the bookmark is seen and selected by a user. The figures discussed in the remainder of this section will concentrate on the medians, because the sample set is highly skewed.

Our second metric for determining success was based on how many times the target bookmark appeared in a result list during a query session. Figure 2b shows the percentage of times that the target bookmark was presented to the user, over all refinement steps in a query session. Recall that the date-based orderings presented the target bookmark 72% of the time (18,965 results lists), whereas our personalized ranking metric made a 9.8% relative increase in presenting the target bookmark (20,818 lists).
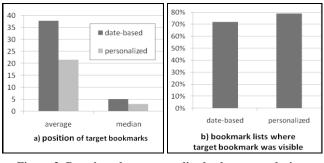

Figure 2. Date-based vs. personalized relevance ordering.

In order to get a better idea of the circumstances under which our technique best performed, we decomposed our results by type of filter event. Table 1 presents the results by each type of query filter used, for each of the ordering metrics. We see that our algorithm outperforms or equals the date-based ordering in all cases. In general the size of improvement seems to be proportional to the size of the result lists. As previously mentioned social tagging systems suffer from limited and converging vocabularies, resulting in large numbers of bookmarks being associated with individual tags. Therefore, when a user initiates a query using only a single tag they are often presented with a large result list. We see that for query refinements using a single tag, the median rank of the selected bookmark is 29. Meaning most users would need to navigate from the first page of results to find a target bookmark (the default list size is 25 results per page). The personalized ranking algorithm reports a much lower median of 19 (a 33% reduction) reducing the median by 10 and placing the result on the first page. We see that when users use more than one tag in their queries our algorithm is not as effective, returning identical medians. This is due to the smaller result list sizes, and the fact that users have already completed at least one refinement step. When users view their personal bookmark collections we see another large ranking improvement, with the median rank dropping from 10 to 6. This is logical, as a user will go to their own profile when they are refinding a bookmark.

In the last two cases we see an increase in rank of only one position, which when compared to the other larger position increases seems to be disappointing. However, when the position of the increase is examined we see an increase in rank from 3 to 2 when people are searching over other people's contributed content and from 2 to the, all important, top result when people are searching others content with a tag in their refinement. This

highlights an encouraging detail in that the personalized algorithm; it performs well when users view the profiles of other users. This also provides evidence that users are refinding bookmarks within the collections of others – whether the bookmark is in the collection of the searcher or not. This is an interesting finding, given that there are links allowing users to very quickly add a new bookmark to their collection or discover if the bookmark already exists in their collection.

filter type	date-based		personalized	
	avg.	med.	avg.	med.
tag - single tag	182.7	29	138.2	19
tag - multiple tags	8.9	4	8.5	4
user - view own profile	44.9	10	22.7	6
user - view other profile	19.7	3	6.7	2
user and tag	3.9	2	2.9	1

Table 1. Target bookmark position by query filter type.

6. CONCLUSION AND FUTURE WORK

We have presented new evidence that helps to characterize social bookmarking system usage. We have incorporated these results into a simple, yet effective, personalized ordering algorithm and our initial evaluation clearly shows its strong potential. Overall, there is a substantial improvement in bookmark ordering, leading to a target bookmark being shown both higher in a list, and being presented after fewer refinement steps. These improvements occur despite the conservative approach of the algorithm, which incorporates all query types and user purposes, and does not consider other possible factors such as contributor and time frame. Our algorithm has potential to improve the user experience using the popular pivot browsing mechanism by improving bookmark orderings, thus reducing the number of query term refinements needed to find a bookmark of interest. Further analysis of the actual implications of our technique would require a live user trial.

A major advantage of our algorithm is its simplicity: It leverages data that can be easily captured and applied in social bookmarking systems. We believe that this provides designers of social bookmarking systems several new options in presenting personalized bookmark lists: by providing a new option for ordering any bookmark list in addition to date and popularity based orderings; and, perhaps, in a new support widget that provides users with a list of their most frequently visited sites.

We will investigate the effects of including additional features to improve the current algorithm performance. For example, the relevancy of a bookmark could also incorporate whether or not the searcher has it in their collection. A bias toward older bookmarks exists in the current algorithm, as older bookmarks may have an increased chance to accrue selections. This bias could be alleviated through the use of a simple decay model. Our analysis shows that refinding is common in social bookmark information accesses with at least 59% of selections being a repeated selection. However, this also means that 41% of selections are to new content. Our technique works best when users are refinding. If we could identify in the initial steps of query refinement a user's intention (whether a user is refinding or looking to discover new bookmarks) we could more intelligently decide when it is appropriate to use our technique and when another metric may be more appropriate.

Finally, despite the lack of success we have had incorporating the influence of group interactions we believe there are still potential insights to be gained from groups of users. An algorithm incorporating group influence would help when a user is looking to discover new information; by identifying a pool of similar users the searcher could be guided by their group's access patterns. Further, it could help overcome the cold start problem, when new users have no history profile for use by relevance algorithms.

8. REFERENCES

1. Chi, E. H. and Mytkowicz, T. Understanding the efficiency of social tagging systems using information theory. *Proc HT 2008*. 81-88, 2008.

2. Cockburn, A. and B. McKenzie. What do Web users do? An empirical analysis of Web use. *Int. J. of Human-Computer Studies*, 54(6): 903-922, 2001.

3. Dover, D. Reddit, Stubleupon, Del.icio.us and Hacker News Algorithms Exposed! Online article, published on SEOmoz July 2nd, 2008, accessed January 2nd, 2009. http://www.seomoz.org/blog /reddit-stumbleupon-delicious-and-hacker-news-algorithms-exposed

4. Freyne, J., Smyth, B., M. Coyle, Balfe, E., Briggs, P. Further Experiments on Collaborative Ranking in Community-Based Web Search. *Art. Intel. Rev.*, *21*, 3-4, (2004), 229-252.

5. Farzan, R., Coyle, M., Freyne, J., Brusilovsky, P., Smyth, B. Adaptive Social Support for Information Space Traversal. In *Proc HT 2007*, (2007), 199-208.

6. Heymann, P., Koutrika, G., and Garcia-Molina, H. Can social bookmarking improve web search?. *Proc WSDM '08*, (2008), ACM, 195-206.

7. Keane, M. T., O'Brien, M., and Smyth, B. Are people biased in their use of search engines? *Commun. ACM. 51*,2,(2008), 49-52.

8. Liu, F., Yu, C., and Meng, W. Personalized web search by mapping user queries to categories. *Proc. CIKM 2002, (2002)*, 558-565.

9. Millen, D.R., Feinberg, J., & Kerr, B. Dogear: Social bookmarking in the enterprise. *Proc.CHI'06,* (2006), 111-120.

10. Tauscher, L. and S. Greenberg. How people revisit Web pages: Empirical findings and implications for the design of history systems. *Int. J. of Human-Computer Studies*, *47*, 1, (1997), 97-137.

11. Thom-Santelli, J., Muller, M., & Millen, D. Social Tagging Roles: Evangelists, Publishers, Leaders. *Proc CHI 2008*, (2008), 1041-1044.