2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology

Exploring Feedback Models in Interactive Tagging

Robert Graham James Caverlee
Department of Computer Science
Texas A&M University
College Station, TX 77843-3112 USA
{rbgraham,caverlee}@tamu.edu

Abstract

One of the cornerstones of the Social Web is informal user-generated metadata (or tags) for annotating web objects like pages, images, and videos. However, many real-world domains are currently left out of the social tagging phenomenon due to the lack of a wide-scale tagging-savvy audience – domains like the personal desktop, enterprise intranets, and digital libraries. Hence in this paper, we propose a lightweight interactive tagging framework for providing high-quality tag suggestions for the vast majority of untagged content. One of the salient features of the proposed framework is its incorporation of user feedback for iteratively refining tag suggestions. Concretely, we describe and evaluate three feedback models – Tag-Based, Term-Based, and Tag Co-location. Through extensive user evaluation and testing, we find that feedback can significantly improve tag quality with minimal user involvement.

1. Introduction

Tags – words or phrases that serve as informal metadata for objects like Web pages, images, and videos – have grown in popularity and purpose in the last few years. Social tagging as a phenomenon corresponds with a Web 2.0 mentality that users can create not only content but a richer, more adaptive and responsive way to navigate and search both existing and new media. In practice, tagging has gained traction among weblogs, social bookmarking sites like del.icio.us, photo sites like Flickr, as well as more traditional media companies like The New York Times and Amazon.com, among many others.

Widespread social tagging promises better and more intuitive information access through tag-based faceted browsing (e.g., [1]), tag-based search (e.g., [8]), and new applications centered around the emergent semantics inherent in the aggregation of the tagging habits of millions of users (e.g., [2]). In contrast to traditional metadata annotation by experts, tagging can overcome less precision in individual tags (e.g., through misspellings, spam tags, and off-topic tags) through the sheer volume of tags that can be generated for an object, especially by a Web-scale audience.

However, many high-value real-world domains are currently left out of the social tagging phenomenon due to the lack of a wide-scale tagging-savvy audience. For example, users manage thousands of local documents on their desktop computers. Few, if any, of these documents are exposed to a Web-scale audience for tagging, and users are typically resistant to go back through their archives to manually apply tags. Privacy concerns also limit the potential effectiveness of social tagging on the personal desktop. Similar factors have slowed the adoption of social tagging in government and industry enterprises, digital libraries, and other specialized information services where archival content and large email and document sharing networks are prime candidates to take advantage of the social web phenomenon.

To bring the power of tag browsing, tag search, and emerging tag-based information access approaches over these untagged domains, we believe there could be some benefit to an *interactive tagging system* for intelligently guiding users. We envision that such a system should be:

- Effective: It should recommend high-quality tags for untagged objects.
- Adaptable: Since different users have different perspectives, it should adapt to each user's needs and interests.
- Lightweight: It should require the user to expend little effort to support the goals of high-quality tags and ease-of-adaptation, plus give novice users some guidance.
- Social: It should take advantage of the collective intelligence of existing socially tagged domains to guide tagging (e.g., to recommend tags the user may want but had not thought of.)

As a first step, we propose in this paper an interactive tagging framework called Plurality. Plurality supports lightweight interactive tagging through a service-based system architecture. Starting with baseline tag suggestions de-

978-0-7695-3496-1/08 $25.00 © 2008 IEEE
DOI 10.1109/WIIAT.2008.419

rived through content-based analysis, Plurality iteratively refines the tag suggestions by incorporating user feedback. Concretely, we describe and evaluate three feedback models inspired by relevance feedback in traditional information retrieval – Term-Based, Tag-Based, and Tag Co-location. Each model represents a different hypothesis about what is the best way to capture user feedback and explore the space of possible tags. Through extensive user evaluation and testing, we find that feedback can significantly improve tag quality with minimal user involvement.

The rest of the paper is organized as follows. In Section 2, we highlight some related work in tagging. Section 3 describes the overall system framework and architecture, including a detailed discussion of the feedback models in Section 3.3. We evaluate the proposed system in Section 4 and conclude in Section 5 with some final thoughts.

2. Background and Related Work

Social tagging (or social annotations) have received growing research attention, particularly as a form of bottom-up user-generated semantics for the Web. At its core, social tagging is a community-based activity, in which users assign tags (typically a word or short phrase) to objects. On the del.ic.ous social tagging service, a user could tag the Web resource `www.espn.com` with tags like "sports", "my-favorites", and "scores". While a user's tags can be used in isolation to provide personalized metadata over Web objects (e.g., as a form of bookmarking), most social tagging systems support community-based aggregation for smarter tag-based browsing [1], search [8], and information access (e.g., through tag-based clustering [2]).

Even in the presence of large and heterogeneous user communities, the tags applied to objects tend to display clear structural properties (which might be counter-intuitive, especially in contrast to the tightly controlled metadata produced by domain experts). For example, Halpin, Robu, and Shepherd [6] studied the complex dynamics of the large-scale social tagging system del.ic.ous and found a power-law distribution for the tags applied to Websites – meaning that in the aggregate, distinct users independently described site using a common tagging vocabulary. These results echo Golder and Huberman [5], who also found a number of clear structural patterns in del.ic.ous, including the stabilization of tags over time. This stabilization suggest an imitation of others or shared knowledge in tagging communities; both features suggest that our proposed interactive tagging framework can leverage existing tagging systems to extract relevant structure for application to new data.

Several papers have studied why users tag and what incentives are in place for continued tagging. Sen et al. [14] enumerate several goals of tagging (self-expression, organization, learning, finding, and decision support) and three tag types (personal, subjective, and factual). Veres [16] finds the same stability as other tagging papers even in low frequency tags that do not get recommended by the user interface. Veres [16] also finds that tags are goal derived and not taxonomic in nature, further differentiating tags from professionally-assigned metadata. Marlow et al. [9] analyze tag systems incentives and benefits. Since our goal is to bring tagging to traditionally untagged domains, we make note of a corporate effort to leverage tagging by Farrell et al. [3]; they show some promise in that people tend to update their employee profiles more often if others are allowed to assign tags to it. We also note that even for the most advanced social tagging domain – the Web – only a fraction of all Web content has actually been tagged (on the order of $\frac{1}{1000}$ of the Web) [7]. And for Web content that has already been tagged, a particular user's personalized view over the Web content may not be reflected in the tags applied by others, again stressing the need for new approaches to extend the reach of traditional social tagging.

Several researchers have examined the problem of automatically generating tags for an object [15], typically through an analysis of the textual content of the object [2], [4], [10], [11]. For example, Mishne [10] proposes a collaborative filtering approach to identifying relevant tags in other closely related documents. Paul et al. [11] use textual analysis of a user's personal documents to apply tags to unseen Web documents. These tag suggestion approaches are complementary to the framework studied here. We promote an *interactive* tagging framework in which users are intentionally kept in-the-loop for guiding the tagging process.

3. The Plurality Model and System Description

In this section, we begin by describing the high-level abstract framework for interactive tagging (Section 3.1). Next, we present the concrete implementation of interactive tagging in the Plurality system (Section 3.2). Finally, we propose and describe three feedback models that are at the heart of the interactive tagging framework (Section 3.3) for iteratively refining the suggested tags.

To motivate the need for an interactive tagging system, consider a user who wishes to leverage social tagging advances over her desktop content (for example, in email and archival documents). Manually tagging each document may be too burdensome and exposing the content to a Web-scale audience for tagging may be infeasible (particularly for privacy reasons). Relying on purely text-based analysis of the user's desktop content (e.g., extracting keywords from documents to be used as tags) is closely aligned with traditional text-based search and may miss out on the collective intelligence that arises in the tagging habits of thousands of users (e.g., that a document mentioning "Kobe Bryant" may also be relevant to the Beijing Olympics, even though no keyword in the document cites “Beijing” or “Olympics”). How can our user take advantage of the “social” aspects of social tagging and the correlated benefits of smarter tag-based browsing and search?

3.1. Interactive Tagging Framework

In response to the challenges outlined above, we propose an interactive tagging framework that supports high-quality personalized tag suggestions over untagged content. The interactive tagging framework is designed to bridge the gap between the solitary user (or enterprise) who can slowly tag each document and the large-scale social intelligence embedded in existing social tagging services. The idea of iterative tagging is for the user to provide guidance to the system in the selection of tags.

In our framework, our goal is to identify a high-quality tag set $T = \{tag_1, tag_2, ..., tag_n\}$ for an untagged object o. We assume the user has access to an existing tag database D consisting of $(object, user, tag, time)$ tuples, for example a tag database extracted from an existing social tagging service like Flickr or del.ic.ous. Each tuple defines the relationship between each object (e.g., a Web resource, an image), the users who have tagged the resource, the tags applied to the resource, and the time at which each tag was applied. The iterative process for identifying high-quality tags for the untagged object o proceeds as follows:

1. Identify the top-k most relevant objects in database D with respect to the input object o.
2. Extract and rank the tags from the top-k matching objects.
3. The user marks some of the extracted tags as relevant or nonrelevant.
4. The system revises its object representation o^* based on the user feedback, and identifies the top-k most relevant objects in database D to o^*.
5. Goto Step 2, until a stopping condition is met.
6. The user selects a final tag set T to be applied to the object o. Additionally, the user may opt to add in additional tags that were not suggested by the system.

A successful interactive tagging framework requires the selection of an appropriate tag database to bootstrap the initial tag suggestions. An ill-matched tag database may provide poor tag coverage for the set of untagged objects to be tagged, e.g., using a sports-related tag database to aid the tagging of finance-related documents. Similarly, the definition of “relevance” is important for identifying related objects in the tag database D. Both tag database selection and

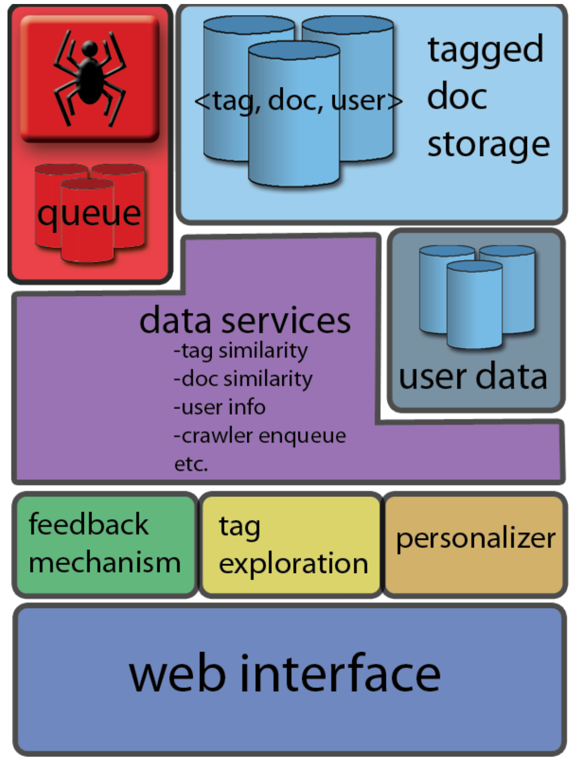

Figure 1. Plurality: System Architecture

relevance definition are important research questions, and to some degree several related studies have considered the problem of identifying relevant tags to untagged content, e.g., [2, 4, 10, 11].

In this paper, we study how to intelligently support feedback within the interactive tagging framework. By incorporating user feedback to guide the tagging process, we can support a more personalized, hybrid tagging approach that takes the best of the overarching semantics inherent in the tagging habits of millions of users, plus the specific goals of a particular tagger.

3.2. Plurality System Description

We have implemented the interactive tagging framework in a service architecture-based system called Plurality (see Figure 1). In its current version, Plurality supports interactive tagging over text and HTML-based documents using a variation of the standard vector space model used in information retrieval. Concretely, Plurality utilizes the Solr web services stack over the Apache Lucene search engine. Solr is an adaptable service layer that supports HTTP GET requests, meaning the system can be easily adapted to support additional document types (e.g., PDF, Word).

Given an untagged candidate document c, Plurality makes its initial tag suggestions (recall Step 1 in the iterative tagging process described above) based on a nearest-neighbor search over the tag database D. We simplify the tag database to consider tuples of the form $(document, user, tag)$ (dropping the time that the tag was applied and considering only objects that are documents).

Each document $d \in D$ is scored with respect to the candidate document c according to the popular cosine similarity used in information retrieval:

$$docScore(c, d) = \frac{\vec{c} \cdot \vec{d}}{|\vec{c}||\vec{d}|} = \frac{\sum_{i=1}^{n} w_{i,c} w_{i,d}}{\sqrt{\sum_{i=1}^{n} w_{i,c}^2} \sqrt{\sum_{1=1}^{n} w_{i,d}^2}}$$

The weights of each term $w_{i,c}$ and $w_{i,d}$ are determined by the application of a fairly standard TF-IDF variant common in information retrieval:

$$w_{i,d} = \sqrt{freq(w, d)} * \left(1 + \log \frac{N}{df(w) + 1}\right)$$

where $freq(w, d)$ is the frequency of word w in document d and where $df(w)$ is the number of documents in D containing word w.

Plurality takes the top-n best matching documents, extracts their tags from the tag database D, and ranks the tags according to the tag scoring function:

$$tagScore(c, t) = \sum_{i=1}^{n} docScore(c, d_i)^4 * freq(t, d_i)$$

where $freq(t, d)$ is the number of times that tag t has been applied to document d in the tag database D. Weighting the tag scores by the pairwise document similarity $docScore(c, d)$ favors the tags in highly relevant documents. The hypothesis is that the tags applied to a very similar document would most likely be applied to the new untagged document. Weighting the tag scores by the frequency of the tag's occurrence on the document $freq(t, d)$ favors more popular tags over rarer, probably less important tags (recall the discussion of tag structure in the Section 2).

Finally, the Plurality system takes the top-k highest-ranked tags and recommends them to the user as initial tag suggestion for the candidate document c. An example screenshot of tag recommendations for a user-selected document is shown in Figure 2.

Figure 2. Tag Suggestions Screenshot

3.3. Exploring Feedback Models

Given the overall framework and system description, we now turn our attention to exploring several feedback models for guiding the quality of tag suggestions. In this section we describe three feedback models at the heart of the iterative process for revising tag suggestions. This feedback is similar in spirit to relevance feedback in information retrieval, which has received considerable research attention [12, 13] and shown promising results.

Our goal is to let the user guide the system's tag recommendations, so that the final tags suggested are satisfactory to the user. Of critical importance is that the system be lightweight in the sense that users can easily revise and improve the tag suggestions without much effort (e.g., many additional clicks, reading more text, etc.).

We propose three models for incorporating user feedback into the tag suggestion process: (1) **Tag Feedback**; (2) **Term Feedback**; and (3) **Tag Co-location**. In the rest of this section, we describe each of these models in detail.

Tag Feedback

In the first model, users provide feedback on the highest-scoring *tags* currently suggested by the system. The goal of the first feedback model is to inherent tags from the most similar documents in the tag database. The hypothesis motivating this model is that the tags assigned to the most similar documents are most likely to be the best tags for the candidate document.

After finding the initial batch of best-k matching documents (recall Step 1 in the interactive tagging framework), the system selects the top-t highest-scoring tags (by $tagScore(c, t)$) from these matching documents to present to the user. For each of these top-t tags, the user can indicate if the tag is relevant to the candidate document or irrelevant to the candidate document. Based on this feedback, a new candidate document vector representation $\vec{c}_{new}$ is constructed based on the original candidate document term vector $\vec{c} = (w_{1,c}, ..., w_{n,c})$ and the user's expressed interests. This new candidate vector is used to find better matching documents (and then tags) following the basic procedure outlined in Section 3.2.

In our current implementation, positive feedback is considered only, and it has the following impact. For each relevant tag tag_{rel}, the system takes each document from the top-k matching documents that has been tagged with tag_{rel} and adds the top-j terms (as scored by TF-IDF) from each of those documents to the augmented candidate document representation $\vec{c}_{new}$. In practice, the choice of k and j can be tuned to widen the range of tags recommended and to optimize the performance of the system; experimentally, we consider $k = 10$ and $j = 5$.

Term Feedback

Instead of providing feedback on tags, in the second model, users provide feedback on *terms* extracted from the content of the most similar documents. Like the first model, the goal of the second feedback model is to inherent tags from the most similar documents in the tag database. The difference here is that the user provides preferences over terms instead of tags. The intuition is that by providing term-based feedback, the user can more directly push the system to identity more related documents (since the underlying document match is based on the content of the documents).

The procedure begins in a fashion similar to the first approach: the system first identifies the best-k matching documents to the candidate document using the text-based cosine similarity described above. But instead of presenting tags to the user, the system selects the top-t highest-scoring *terms* (by TF-IDF weight) from these matching documents to present to the user. For each of these top-t terms, the user can indicate if the term is relevant to the candidate document or irrelevant to the candidate document. Based on this feedback, a new candidate document vector representation $\vec{c}_{new}$ is constructed based on the original candidate document term vector $\vec{c} = (w_{1,c}, ..., w_{n,c})$ and the user's expressed interests in the presented terms.

In our current implementation, positive feedback has two impacts on the augmented document representation $\vec{c}_{new}$. First, the selected term is added to the augmented document representation $\vec{c}_{new}$ unconditionally. Second, terms that co-occur with the selected term are also added to $\vec{c}_{new}$; the top-j terms from each document in the top-k matching document set that contain the term are added to $\vec{c}_{new}$.

Tag Co-location

The first two models use feedback to identify the closest matching documents to the candidate document; the tags from these documents are assumed to be good tags for the candidate document. In contrast, the Tag Co-location model takes advantage of the collective intelligence of all users and their tag decisions to find related tags for the candidate document. The hypothesis is that tags and the tagging habits of real users are more important than terms and documents for finding relevant tags.

As in the previous two models, the system begins by finding the initial batch of best-k matching documents and selecting the top-t highest-scoring tags (by $tagScore(c, t)$) from these matching documents to present to the user. But now, when the user provides feedback over the top-t tags, the tags judged to be relevant are then examined for *co-location* with other tags and the most commonly co-located tags are presented to the user. There is no more refinement of the document representation for finding closely matching documents.

For each tag pair, we pre-compute a co-location score for every tag pairing in the database. If there are M documents in the tag database, tag_i is represented by a document-based vector $\vec{tag_i}$ where each entry is a count of the number of times tag_i was applied to document d:

$$\vec{tag_i} = (d_{i,1}, d_{i,2}, ..., d_{i,M})$$

and the co-location similarity score is a straightforward adaptation of the cosine similarity measure:

$$coloc(tag_i, tag_j) = \frac{\vec{tag_i} \cdot \vec{tag_j}}{|\vec{tag_i}||\vec{tag_j}|}$$

For the list of all tags marked as relevant by the user T_r, the system can compute an aggregate score for all tags based on the scoring function:

$$score(tag_j) = \sum_{tag_i \in T_r} coloc(tag_i, tag_j)$$

In this way, tags that are co-located with several relevant tags (as indicated by the user in her feedback) can be given additional weight.

4. Experiments

In this section, we evaluate the iterative tagging framework and consider the relative benefits of the three feedback models described in Section 3.3. In summary, we find that feedback can significantly improve the quality of tag recommendations with and that the best performing technique is the Tag Co-location model.

4.1. Experimental Setup

We conducted a user study over the course of one week with 15 volunteer participants. Our goal was to understand how well users interact with the system through qualitative and quantitative data about the quality of Plurality's tag suggestions. The participants were mostly graduate students in Computer Science and ranged in age from 22 and 51. We recorded over 200 tagging sessions. In each session, a user was given a sequence of documents to tag; each

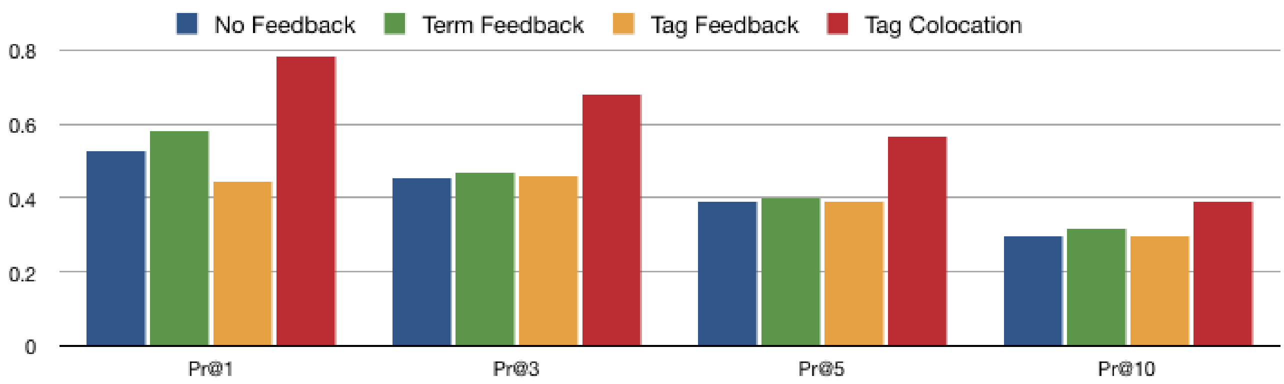

Figure 3. Pr@k Comparison Among Feedback Methods

document was randomly selected from a pool of 100 randomly selected documents tagged on `del.icio.us` in June 2008. The tag database was taken to be a collection of 20,000 `del.icio.us` documents we crawled in the Spring of 2008; note that the test set is non-overlapping with the tag database. For each document to be tagged, one of the three feedback models (Tag-Based, Term-Based, and Tag Co-location) was used without alerting the user to the technique being used. In each case, up to 10 tags were presented to the user.

We measured the quality of the final ranked list of tags recommended by Plurality to the user with the Precision@k measure. Precision@k measures how many of the top-k tags are actually relevant to the document being tagged. Relevance judgments for tags were recorded by considering whether the user checked the tag during the session. The only points at which we calculated the values for these metrics is at the initial screen of tags we recommend to the user – which is the same across all feedback methods for one document – and the final screen that the user selects the tags to be applied to the document with the option to add their own tags. We measured the statistical significance of the result using the Wilcoxon Signed-Rank Test. The comparison for improvements shown by the feedback methods is a comparison between the relevance of the first set of tags they see and the last. In the case of terms for the Term Feedback method we simply scored it as relevant terms to start vs. tags at the end.

4.2. Results

First, we compare in Figure 3 the Precision@k for the baseline tag suggestions (without user feedback) and for the three feedback models. Note that with no feedback, the system generates slightly less than 3 relevant tags on average. The Pr@1 without feedback was over 50% and the Pr@3 was around 45%. This is the baseline against which all the

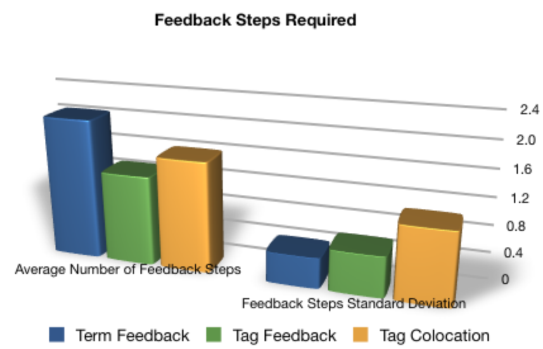

Figure 4. Number of steps required to tag a document in each method.

feedback methods were compared. We see that each feedback method produces some improvement over the original (no feedback) suggestions, but only the Tag Co-location method produces a statistically significant improvement (at the 95% confidence level) in every category.

Users observed that Tag Co-location had the most tag turnover of any feedback method. When a user selected tags using this method the most commonly co-located tags appeared with them on the next list which users said changed more of the list more often than the other two. This makes sense if you consider an example. If the user selects one tag using Tag Co-location then that tag and the nine most commonly co-located tags for that tag will be displayed. This list could be different for nine of ten spots. In contrast, the Tag Feedback and Term Feedback models are fundamentally a refinement process that is unlikely to change the majority of the matching documents and hence the resulting tag list.

Figure 4 reports how many interface screens the user

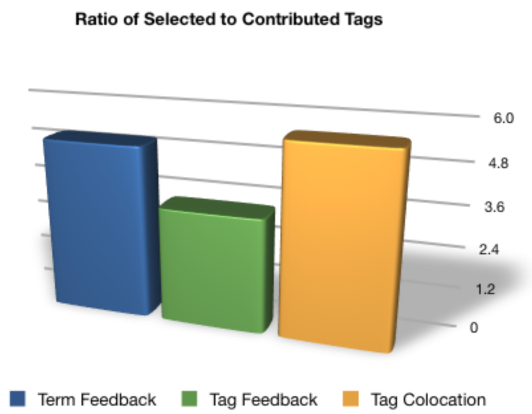

Figure 5. Ratio of Selected Tags to Contributed Tags

went through before finishing the tagging process. We see that on average users needed around 2 iterations with Term Feedback (i.e., two rounds of feedback) versus around 1.5 for Tag Co-location, meaning that the Tag Co-location model required less effort on the part of users. Note however the higher standard deviation for Tag Co-location. Since Tag Co-location explores a wider tag space than the other techniques (there is higher tag turnover per round), this model sometimes requires additional effort.

Finally, we report in Figure 5 the ratio of tags that the user checked from the suggestions to the number they added explicitly. Recall that the system design allows users to manually add tags beyond what the system recommends. We see that the ratio of selected tags (those tags recommended by the system) to user submitted tags is highest for the Tag Co-location model, indicating that this model is the best at capturing good tags and reducing the amount of user work (by manually adding tags).

5. Conclusions and Future Work

We have proposed a lightweight interactive tagging framework for providing high-quality tag suggestions and explored three feedback models. We introduced and evaluated three feedback models – Tag-Based, Term-Based, and Tag Co-location. Our initial results are encouraging, especially with respect to the Tag Co-location feedback model. In our ongoing work, we are revisiting two important factors. First, what are the most important qualities of an external tag database with respect to the quality of tags recommended? We are interested in developing tag database selection algorithms that can dynamically select the best tag database for a particular user, for a particular context, or for a particular corpus of untagged documents. Second, how can we improve the initial tag recommendations made in the first place? And how does the choice of initial tags impact the quality of revised tags through user feedback?

References

[1] S. Bao, G. Xue, X. Wu, Y. Yu, B. Fei, and Z. Su. Optimizing web search using social annotations. In *International World Wide Web Conference (WWW)*, 2007.

[2] C. H. Brooks and N. Montanez. Improved annotation of the blogosphere via autotagging and hierarchical clustering. In *International World Wide Web Conference (WWW)*, 2006.

[3] S. Farrell, T. Lau, S. Nusser, E. Wilcox, and M. Muller. Socially augmenting employee profiles with people-tagging. In *ACM Symposium on User Interface Software and Technology (UIST)*, 2007.

[4] S. Fujimura, K. Fujimura, and H. Okuda. Blogosonomy: Autotagging any text using bloggers' knowledge. In *IEEE/WIC/ACM International Conference on Web Intelligence*, 2007.

[5] S. Golder and B. A. Huberman. Usage patterns of collaborative tagging systems. *Journal of Information Science*, 32(2):198–208, 2006.

[6] H. Halpin, V. Robu, and H. Shepherd. The complex dynamics of collaborative tagging. In *International World Wide Web Conference (WWW)*, 2007.

[7] P. Heymann, G. Koutrika, and H. Garcia-Molina. Can social bookmarking improve web search? In *International Conference on Web Search and Web Data Mining (WSDM)*, 2008.

[8] R. Li, S. Bao, Y. Yu, B. Fei, and Z. Su. Towards effective browsing of large scale social annotations. In *International World Wide Web Conference (WWW)*, 2007.

[9] C. Marlow, M. Naaman, D. Boyd, and M. Davis. Ht06, tagging paper, taxonomy, flickr, academic article, to read. In *International Conference on Hypertext and Hypermedia (HT)*, 2006.

[10] G. Mishne. Autotag: a collaborative approach to automated tag assignment for weblog posts. In *International World Wide Web Conference (WWW)*, 2006.

[11] Paul, S. Costache, W. Nejdl, and S. Handschuh. P-TAG: large scale automatic generation of personalized annotation tags for the web. In *International World Wide Web Conference (WWW)*, 2007.

[12] J. Rocchio. Relevance feedback in information retrieval. In *The SMART Retrieval System*, pages 313–323, 1971.

[13] G. Salton and C. Buckley. Improving retrieval performance by relevance feedback. *Journal of the American Society for Information Science*, 41(4):288–297, 1990.

[14] S. Sen, S. K. Lam, A. M. Rashid, D. Cosley, D. Frankowski, J. Osterhouse, M. F. Harper, and J. Riedl. Tagging, communities, vocabulary, evolution. In *International Conference on Computer Supported Cooperative Work (CSCW)*, 2006.

[15] Y. Song, Z. Zhuang, H. Li, Q. Zhao, J. Li, W. Lee, and C. Giles. Real-time automatic tag recommendation. In *SIGIR*, 2008.

[16] C. Veres. The language of folksonomies: What tags reveal about user classification. *Natural Language Processing and Information Systems*, pages 58–69, 2006.