Web Search Disambiguation by Collaborative Tagging

Ching-man Au Yeung, Nicholas Gibbins, and Nigel Shadbolt

Intelligence, Agents, Multimedia Group,
School of Electronics and Computer Science,
University of Southampton,
Southampton SO17 1BJ, UK
{cmay06r,nmg,nrs}@ecs.soton.ac.uk

Abstract. Existing Web search engines such as Google mostly adopt a keyword-based approach, which matches the keywords in a query submitted by a user with the keywords characterising the indexed Web documents, and is quite successful in general in helping users locate useful documents. However, when the keyword submitted by the user is ambiguous, the search result usually consists of documents related to various meanings of the keyword, in which probably only one of them is interesting to the user. In this paper we attempt to provide a solution to this problem by using the semantics extracted from collaborative tagging in the social bookmarking site del.icio.us. For an ambiguous word, we extract sets of tags which are related to it in different contexts by performing a community-discovery algorithm on folksonomy networks. The sets of tags are then used to disambiguate search results returned by del.icio.us and Google. Experimental results show that our method is able to disambiguate the documents returned by the two systems with high precision.

1 Introduction

The amount of information on the World Wide Web is huge and keeps increasing as people from all over the world continue to contribute to this information network. It was estimated that there were already over 11.5 billion Web pages on the Web as of the end of January 2005 [4]. Such rapid growth has made the retrieval of information that is relevant to the needs of a user very difficult. While search engines help ease the problem by indexing the Web and returning search results based on ranking algorithms such as the PageRank [2] algorithm, in many situations the results returned are not as useful as the users have expected.

An obvious example of such situations is when a keyword with multiple meanings is used to query the search engines. Very often the retrieved documents are relevant to multiple meanings of the keyword, but the user is probably only interested in one of the meanings or one of the contexts in which the keyword is used. For example, when a user queries Google with the keyword *bridge*, he might be presented with Web pages about bridge as a kind of card game, as a design pattern in programming, or as a physical structure built across a river. The user who is only interested in bridge as a kind of card game will have to scan through the list of returned documents and single out those which are really relevant.

In recent years, collaborative tagging systems such as del.icio.us and Flickr have become very popular among Web users as a means of organising their favourite Web resources.[1] In these systems, users are allowed to choose any words they like as tags to describe Web resources, resulting in a user-generated classification scheme now commonly known as a *folksonomy* [20]. Not only does a folksonomy provide metadata of Web resources in the form of tags, it also provides a lot of information on the relations between different tags when they are used together. We have shown [1] that by performing clustering on documents tagged in a folksonomy, it is possible to extract the sets of tags related to the different contexts in which an ambiguous tag is used.

In this paper, we discuss how such implicit semantics extracted from a folksonomy can be utilised to enhance Web search. We propose a method to disambiguate Web search results by classifying returned documents into different contexts in which an ambiguous keyword is used. Evaluation is performed by applying the method on search results returned by del.icio.us and Google.

The remaining of this paper is structured as follows. The next section presents an example which motivates this research. In Section 3, we describe our method for automatically extracting the different meanings of an ambiguous tag from a folksonomy. In Section 4, we describe how we apply the results of tag meaning disambiguation on Web search disambiguation. Section 5 presents the experimental results. Finally, we mention some related work in Section 6 and give conclusions and future research directions in Section 7.

2 Motivating Example

It is very common for a user of a Web search engine to find that the search results are not as useful as expected. This is particularly true when the keywords used in the query represent different concepts when used in different contexts. In such cases the users have to go through the list of returned documents and single out those documents that are relevant to their needs.

Consider the following example of searching for information about the type of card game called Bridge. Table 1 lists the top ten pages returned by Google UK when *bridge* is used as the query string. While the first and the third pages returned are about the card game, the search results actually consist of pages about other meanings of the word *bridge*. For example, the second item is a page from Wikipedia describing bridges as architectural structures, and the sixth item is a page which contains travel information about the Golden Gate Bridge. There are also pages (e.g. 7th and 10th) which involve organisations or projects with the name 'Bridge' but are by no means related to any commonly used meanings of the word.

[1] http://del.icio.us/, http://www.flickr.com

1	Contract bridge - Wikipedia, the free encyclopedia http://en.wikipedia.org/wiki/Contract_bridge
2	Bridge - Wikipedia, the free encyclopedia http://en.wikipedia.org/wiki/Bridge
3	Play bridge card game online http://www.bridgeclublive.com/
4	Bridge Travel http://direct.bridge-travel.co.uk/
5	River Kwai Bridge Travel http://www.riverkwaibridge.com/
6	Golden Gate Bridge Guide — Attraction Travel Guide http://www.worldtouristattractions.travel-guides.com/attraction/170/attraction_guide/North-America/Golden-Gate-Bridge.html
7	Bridge - Mainstreaming Gender Equality http://www.bridge.ids.ac.uk/
8	Bridge to Reuters http://www.bridge.com/
9	The Bridge SE1 - London venue for parties, gigs, films, conference http://www.thebridgese1.co.uk/
10	BRIDGE (Building Radio Frequency IDentification for the Global Environment) http:// www.bridge-project.eu/

Table 1. The top ten pages returned by Google UK when '*bridge*' is used as a query string.

Two major problems can be observed in this example. Firstly, extra effort is required for the user to go through the list and select those results which are useful. Secondly, the presence of pages which are irrelevant to the user's need reduces the number of relevant pages that can be presented to the user at one time, especially when users tend to only inspect the first set of items returned [8, 18]. Although in some search engines terms which are commonly used together with the keyword are suggested to the users for refining the search results, it will be much more efficient if the search engine is able to classify the pages into different categories which correspond to different meanings of the keyword before presenting the results to the user. To tackle these problems, we propose a method for Web search disambiguation by using the semantics extracted from a folksonomy.

3 Tag Meaning Disambiguation

Our first step to Web search disambiguation involves obtaining the different meanings of an ambiguous word from a folksonomy.[2] We have shown that, for

[2] Since a word is referred to as a tag, a keyword or a term depending on the context in which it is being mentioned, we will use these terms interchangeably in the rest of this paper.

an ambiguous tag in a folksonomy, documents which are relevant to the same meaning of the tag tend to be grouped together [1]. This suggests that clustering algorithms can be applied to extract groups of documents which correspond to different meanings of the tag. Our target is to extract sets of tags which constitute different contexts in which an ambiguous tag is used. The proposed process for tag meaning disambiguation is described as follows.

A folksonomy is generally considered to consist of at least three sets of elements [13, 21], namely users, tags, and documents. Formally, we define a folksonomy as a tuple $\mathbf{F} = (U, T, D, A)$, where U is a set of users, T is a set of tags, D is a set of Web documents, and $A \subseteq U \times T \times D$ is a set of annotations. When we want to understand the different meanings of an ambiguous tag t, only a subset of the folksonomy involving the tag is required. This can be obtained by extracting the bipartite graph UD_t by restricting $\mathbf{F}$ to t:

$$UD_t = \langle U \cup D, E_{ud} \rangle, E_{ud} = \{(u, d) | (u, t, d) \in A\}$$

This graph can be represented in matrix form, which we denote as $\mathbf{Y} = \{y_{ij}\}$, $y_{ij} = 1$ if there is an edge connecting user u_i and document d_j, and $y_{ij} = 0$ otherwise. We further fold this bipartite graph into a one-mode network of documents by performing matrix multiplication, obtaining $\mathbf{C} = \mathbf{Y}'\mathbf{Y}$. In this one mode network, an edge is weighed by the number of users who have assigned tag t to the documents represented by the vertices on the two ends of the edge.

From this network of documents, we can extract groups of documents where each group corresponds to a single meaning of the tag t. This can be done by applying clustering algorithms to the network represented by $\mathbf{C}$. We adopt the fast greedy algorithm for community discovery in networks proposed in [15], which optimises modularity [16] by connecting the two vertices at each step which result in the largest increase (or smallest decrease) of modularity. If D_t is the set of documents which are assigned the tag t, the result of the clustering process is a set of sets of documents: $\mathbf{X}_t = \{X_{t,1}, X_{t,2}, ..., X_{t,m}\}$ where $X_{t,1} \cup X_{t,2} \cup \cdots \cup X_{t,m} = D_t$. Finally, for each set $X_{t,i}$, we obtain a set $T_{t,i}$ of the top 10 tags which are used most frequently by the users on the documents in the set.

While each of these sets of tags is likely to be related to a single meaning of the ambiguous tag t, it is possible that two or more of these sets are related to the same meaning. To eliminate the redundancy in the result we combine two sets of tags if there is significant overlap between the two with the help of the following function:

$$overlap(T_{t,i}, T_{t,j}) = \frac{|T_{t,i} \cap T_{t,j}|}{|T_{t,i} \cup T_{t,j}|} \tag{1}$$

We introduce a threshold α, and merge the two sets of documents $X_{t,i}$ and $X_{t,j}$ when $overlap(T_{t,i}, T_{t,j}) \geq \alpha$. The top 10 tags with the highest frequencies are extracted to form a new set. Hence, the final result of this tag meaning disambiguation process is a set of sets of tags: $\mathbf{T}_t = \{T_{t,1}, T_{t,2}, ..., T_{t,n}\}$, where $n \leq m$. The whole process is summarised in Algorithm 1.

Algorithm 1: Tag meaning disambiguation

```
Input: Adjacency matrix C of the network of documents
Output: A set T of sets of tags
begin
    // Document clustering;
    X ← FastGreedyCommunityDiscovery(C);
    T ← {};
    // Extract top 10 tags;
    for X_i ∈ X do
        T_i ← Top10Tags(X_i);
        T ← T ∪ {T_i};
    end
    // Merge similar sets of tags;
    merged ← 1;
    while merged = 1 do
        merged ← 0;
        for T_i, T_j ∈ T and i ≠ j do
            if overlap(T_i, T_j) ≥ α then
                X_new ← X_i ∪ X_j;
                T_new ← Top10Tags(X_new);
                T ← T − {T_i, T_j};
                T ← T ∪ {T_new};
                merged ← 1;
            end
        end
    end
    return T;
end
```

4 Web Search Disambiguation

The result of the tag meaning disambiguation obtained from the method described in the previous section can be used to disambiguate Web search results. This is done by comparing the tags corresponding to the different meanings of an ambiguous tag with the keywords characterising a document in the search results. The steps are described in detail as follows.

Given a set D_t of documents returned by a search engine when queried with an ambiguous keyword t, our target is to classify the documents into different categories, each corresponding to a different meaning of t, yielding a set of sets of documents $\mathbf{D}_t = \{D_{t,0}, D_{t,2}, ..., D_{t,n}\}$ where $D_{t,0} \cup D_{t,1} \cup \cdots \cup D_{t,n} = D_t$. We assume that each document is only related to one meaning of t. Each set of documents $D_{t,i}$ corresponds to the meaning represented by T_i, except that $D_{t,0}$ is the set of documents which cannot be classified into any of these categories represented by $T_{t,1}, ..., T_{t,n}$. We further assume that each document $d_j \in D_t$ is characterised by a set $K_{t,j}$ of keywords, which could be the keywords used to

Algorithm 2: Web search disambiguation

```
Input: A set T of sets of tags, a set D_s of documents
Output: A set D of sets of classified documents
begin
    // Initialisation;
    D ← {};
    for i ← 0 to |T| do
        D_i ← {};
        D ← D ∪ {D_i};
    end
    // Classify documents;
    for d ∈ D_t do
        x ← Cat_A(d);
        D_x ← D_x ∪ {d};
    end
    return D;
end
```

index the document by the search engine, or the tags assigned to the document by users in a collaborative tagging system.

Firstly, we define the function *match* which calculates the extent to which the set $K_{t,j}$ of keywords of document d_j matches the set $T_{t,i}$ of tags of a particular meaning of the term t.

$$match(K_{t,j}, T_{t,i}) = \frac{|K_{t,j} \cap T_{t,j}|}{|T_{t,i}|} \tag{2}$$

By comparing the different values returned by the *match* function when different sets of tags are used, a document d_j is assigned to a particular category as follows.

$$Cat_A(d_j, t) = \begin{cases} \operatorname*{argmax}_i \, match(K_{t,j}, T_{t,i}), & \text{if } \max_i match(K_{t,j}, T_{t,i}) \geq \beta \\ 0, & \text{if } \max_i match(K_{t,j}, T_{t,i}) < \beta \end{cases} \tag{3}$$

The function Cat_A (the subscript A stands for automatic) assigns d_j a category which corresponds to the meaning of t represented by the set $T_{t,i}$ of tags which match the best with the keywords of d_j. However, if the keywords of d_j match poorly with any of the sets of tags, the document is assigned the category of 0. The threshold β is a value in the range of 0 to 1. The whole process is summarised in Algorithm 2.

5 Evaluation

In order to evaluate our proposed method of Web search disambiguation, we apply the method to Web search results obtained by querying del.icio.us and

Tag	Number of Documents	Number of Users
sf	426	446
tube	476	427
bridge	915	338
wine	421	896

Table 2. Statistics of the dataset collected from del.icio.us.

Google UK using four ambiguous terms, namely *sf*, *tube*, *bridge* and *wine*. These four terms are selected because it is observed that they are used to represent multiple concepts in del.icio.us, and that search results returned by Google when using these terms in the query also consist of documents related to a rather diverse topic. By applying the method on documents returned by del.icio.us, we can test whether our tag meaning disambiguation is able to identify all of the meanings of the ambiguous tags used in the system. On the other hand, by applying the method on documents returned by Google, we are able to study its performance on a traditional search engine.

5.1 Data Preparation

To generate the sets of tags representing the different meanings of the ambiguous terms, we collect data involving the four tags from del.icio.us by using a crawler program. The dataset includes documents which have been assigned the tags and users who have used the tags on the documents. In the data collection process, we skip documents which are only tagged by one user. Table 2 summarises the statistics of the dataset.

For Google UK, we submit queries using each of the four terms and obtain the top 50 pages returned. We denote the set of documents retrieved for the term t by GD_t. Del.icio.us, although primarily a collaborative tagging system, also provides search service on its data. However, search results returned by del.icio.us are ranked by how recent an item is tagged by a user instead of how relevant an item is to the keyword in the query. Hence, for each of the terms t we extract the top 50 items which are tagged by the greatest number of users with the tag in question as the search result and denote it by DD_t.

Finally, we construct a set of keywords for each document which are used to characterise the document. For documents returned by del.icio.us, the aggregated set of tags contributed by the users are used to form the set of keywords. For documents returned by Google, we first process the texts in the documents and extract keywords by filtering out stop words and non-text symbols, and then enrich the set by querying del.icio.us for the tags, if any, which are assigned to the documents.

Tag	Context	Tags Extracted
sf	San Francisco	sf, sanfrancisco, bayarea, san, francisco, california, travel, events, art, san_francisco
	Science fiction	sf, scifi, fiction, books, sci-fi, literature, writing, sciencefiction, science, fantasy
tube	YouTube videos	tube, youtube, video, funny, videos, fun, cool, music, feel.good, flash
	Vacuum tubes	tube, audio, electronics, diy, amplifier, amp, tubes, music, elect, guitar
	London underground	tube, london, underground, travel, transport, maps, map, uk, subway, reference
bridge	Design pattern	bridge, programming, development, library, code, ruby, tools, software, adobe, dev
	Card game	bridge, games, cards, game, imported, howto, conventions, card, bidding, online
	Computer networking	bridge, networking, linux, network, howto, software, sysadmin, firewall, virtualization, security
	Architecture	bridge, bridges, structures, engineering, science, physics, school, education, building, reference
wine	Software application	wine, linux, ubuntu, howto, windows, software, tutorial, emulation, reference, games
	Beverage	wine, food, shopping, drink, reference, vino, cooking, alcohol, blog, news

Table 3. Meanings of tags discovered and related tags extracted for each meaning.

5.2 Experiments

We first attempt to discover the different contexts in which the ambiguous tags are used by applying our proposed tag disambiguation algorithm on the del.icio.us dataset with $\alpha = 0.2$. By setting $\alpha = 0.2$, we effectively require two sets of tags to have more than three tags in common before we will combine them. This is based on the observation that very often the first three or four most frequently used tags in a set are sufficient for one to decide the meaning to which it corresponds.

The tags extracted for each of the ambiguous tags are shown in Table 3. We can see that the proposed algorithm performs well in revealing the multiple meanings of the tags. For example, four different meanings of the tag *bridge* are discovered, in which the tags extracted are closely related to the contexts in which *bridge* is used.

Next, we apply our proposed Web search disambiguation method, with $\beta = 0.3$, to the search results obtained from del.icio.us and Google. β is chosen based on a similar reason of the choice of α. We first manually classify the returned documents into the categories discovered in the tag meaning disambiguation phrase by inspecting their content. Our classification can be represented by a mapping $Cat_M(d, t)$ which assigns each document d a category x, where $x \in$

Tag	Case	Total	Classified	Unclassified	Classifiable	Correct	Precision	Recall	Coverage
sf	D	50	50	50	50	50	1.00	1.00	1.00
	G	50	38	12	38	37	0.97	0.97	0.74
tube	D	50	50	50	50	50	1.00	1.00	1.00
	G	50	34	16	33	31	0.91	0.94	0.62
bridge	D	50	43	7	49	42	0.98	0.86	0.86
	G	50	16	34	24	13	0.81	0.54	0.26
wine	D	50	50	50	50	50	1.00	1.00	1.00
	G	50	27	23	50	27	1.00	0.54	0.54

Table 4. Results of web search disambiguation. D stands for an experiment on del.icio.us-returned pages, while G stands for one on Google-returned pages.

$\{0, 1, 2, ..., n\}$. Category x corresponds to the meaning of the term t represented by the set $T_{t,x}$ of tags, and the category 0 is reserved for unclassified documents.

We evaluate the performance of the method by using three different measures, namely *precision*, *recall* and *coverage*. ***Precision*** measures the extent to which the documents are classified correctly. It is calculated by dividing the number of correctly classified documents by the total number of classified documents. ***Recall*** measures the fraction of classifiable documents which the method is able to classify. By classifiable documents we refer to documents which should fall into any one of the contexts discovered in the tag meaning disambiguation phase. Finally, ***coverage*** measures how many documents can be classified given the total number of documents returned. Let R_t be the set of retrieved documents, where $R_t = DD_t$ or $R_t = GD_t$ depending on the dataset on which we apply our algorithm. The three measures are defined as follows.

$$\text{Precision} = \frac{|\{d \in R_t | Cat_M(d,t) = Cat_A(d,t) \wedge Cat_M(d,t) \neq 0\}|}{|\{d \in R_t | Cat_A(d,t) \neq 0\}|} \tag{4}$$

$$\text{Recall} = \frac{|\{d \in R_t | Cat_M(d,t) = Cat_A(d,t) \wedge Cat_M(d,t) \neq 0\}|}{|\{d \in R_t | Cat_M(d,t) \neq 0\}|} \tag{5}$$

$$\text{Coverage} = \frac{|\{d \in R_t | Cat_M(d,t) = Cat_A(d,t) \wedge Cat_M(d,t) \neq 0\}|}{|R_t|} \tag{6}$$

The experimental results are shown in Table 4 and Figure 1.

5.3 Discussions

The experimental result shows that documents are classified to the correct categories the majority of the time, with precision ranging from 81% to 100%. This suggests that the tags extracted in the tag meaning disambiguation phase can be used to identify precisely the different contexts in which the ambiguous terms are used. Precision in the cases of del.icio.us was always higher than or equal to those in the cases of Google, probably because the documents in del.icio.us

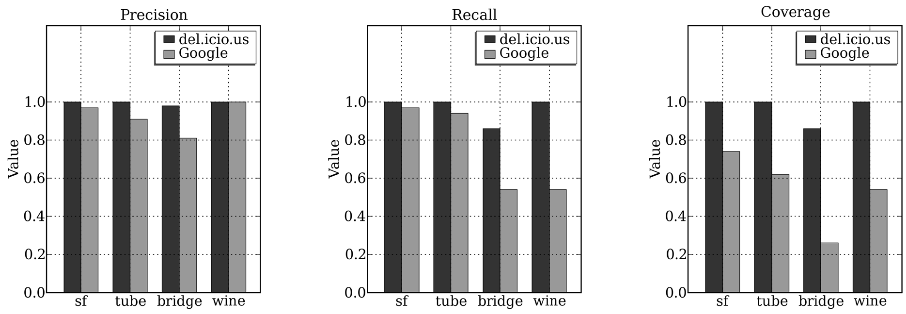

Fig. 1. Precision, recall and coverage of web search disambiguation.

feature keywords which are more similar to the tags used for disambiguation. After all, they are all contributed by users of del.icio.us.

Applying our technique to search results from del.icio.us generally results in higher recall (86% to 100%) than when we apply it to those from Google (54% to 97%). Low recall means that the algorithm is unable to classify documents which are actually related to one of the meanings discovered in the tag meaning disambiguation phase, probably due to poor matches between the tags extracted from del.icio.us and the keywords characterising the documents. This suggests that the tags extracted from del.icio.us may not be comprehensive enough to reconstruct the contexts in which those ambiguous words are used. Recall is particularly low when classifying documents returned by Google for *bridge* and *wine*. We find that quite a number of pages about bridges as architectural structures cannot be identified. These pages are characterised by keywords like *river*, *stream* and *architecture*, which are not present in the set of tags extracted from del.icio.us. Similarly, some pages about wine as a kind of beverage are not identified because they contain keywords like *red*, *white* and *bottle* which are absent from the set of tags for disambiguation. This problem is much less serious when classifying documents from del.icio.us, because the tags extracted are also the tags used frequently on these pages.

Performance in terms of coverage of our method on documents returned by del.icio.us is very satisfactory, suggesting that the tag meaning disambiguation method is able to identify all or most of the multiple meanings of the ambiguous tags used in del.icio.us. However, relatively low coverage (26% to 74%) can be observed in all the cases of classifying documents returned by Google. While low coverage is partly predicted by the low recall in these cases, this result also suggest that the tag meaning disambiguation process is not able to return the different meanings of an ambiguous tag used in a more general situation. For example, the common usage of *tube* to refer to a hollow and circular structure is not identified, which makes the Web search disambiguation algorithm unable to identify documents related to this meaning. On the other hand, among the documents returned by Google, there are in fact a certain number of documents which are not related to any commonly known meanings of the query terms. For example, the coverage in the case of *bridge* is particularly low because some of the documents are only about places or organisations which are named *Bridge*. From this observation, we believe that a low coverage is not as undesirable as it seems, because the algorithm actually helps to filter out documents which are not semantically related to the query term.

In summary, our proposed method for Web search disambiguation is able to classify documents with high precision based on the implicit semantics extracted from collaborative tagging, though in some cases it is not able to identify all relevant documents for the categories. A major issue which requires further investigation is how to increase the comprehensiveness of the tags extracted from folksonomies in order to increase recall and coverage.

6 Related Work

To the best of our knowledge, this is the first study of the use of user-contributed annotations in collaborative tagging systems to disambiguate Web search results. Different methods have been used to discriminate word meanings or senses in the literature. These include the use of manual-constructed rules [9] and the use of dictionaries or thesauri [11, 12]. Our work is similar in part to studies which employ lexical co-occurrence to discover the different senses of an ambiguous word. For example, Schütze and Pedersen [17] derive a term vector for each word which represents word similarity derived from lexical co-occurrence. The vectors are then combined to form context vectors which are clustered to represent different senses of ambiguous words.

In addition, our work is also similar in principle to studies which apply document clustering techniques on Web search results. This is a problem quite extensively studied in the literature [3, 5, 19, 23] and is also addressed by commercial systems such as Vivisimo [10].[3] Existing document clustering techniques in general extract keywords from documents and calculate their similarity based on the keywords to obtain a set of clusters. Our approach differs from these techniques in that instead of performing clustering based on the vocabulary found in the documents returned by the search engine, we obtain a set of categories from analysis of collaborative tagging systems to aid classification of the documents. We believe our proposed method is better than existing approaches, as it is more focused in terms of the meanings of the keywords, while existing document clustering techniques might result in clusters which are not necessarily meaningful to the users.

On the other hand, while there have been no studies which directly address the problem of tag ambiguity, tag meaning disambiguation can be observed as a by-product in some research work which focuses on tag clustering. For example, in the work of Wu et al. [21] latent semantic analysis is applied to study the co-occurrence of tags, and ambiguous tags are found to score highly in multiple pre-

[3] The public version of Vivismo's Web search engine, Clusty, can be found at http://clusty.com/.

defined dimensions. Zhou et al. [24] also report that, in building a tag hierarchy by using deterministic annealing to perform tag clustering, tags with multiple meanings are found to appear in different branches of the resulting hierarchy. In addition, collaborative tagging is also used to improve Web search in general, such as by providing a better ranking of the search results [6, 22]. In contrast to these prior studies, our work directly addresses the problem of tag ambiguity, proposes a feasible solution and studies how the extracted semantics of tags can be applied to novel applications.

7 Conclusions and Future Work

In this paper, we propose a method for automatic Web search disambiguation which uses the implicit semantics extracted from folksonomies. Our preliminary evaluation shows that the tags extracted from tag meaning disambiguation can be used to classify search results returned by Web search engines with high precision. This suggests that tags contributed by users in collaborative tagging systems can be used to enhance the performance of Web search engines. Also, we note a distinct advantage of using tags extracted from collaborative tagging systems for Web search disambiguation. Our proposed method for tag meaning disambiguation is able to discover some unconventional meanings of the ambiguous words, such as *tube* for the video-sharing site YouTube, or *bridge* for the design pattern used in programming. These meanings are rather new and are of specific domains that they may not be available in dictionaries or thesauruses such as Wordnet [14], which are commonly used for word sense disambiguation in the literature [7].

At the same time, we are aware of several problems in the proposed method. In particular, the levels of recall and coverage are significantly lower than that of precision, meaning that some relevant documents cannot be identified with the tags we extract from a folksonomy. Based on the results reported in this paper, we plan to extend our research work in several directions. Firstly, we will study how the comprehensiveness of the set of tags which represents a particular meaning of an ambiguous term can be increased, such as by expanding it with tags which co-occur frequently with the set in order to increasing the chance of matching the keywords which characterise the documents. Secondly, we will investigate how we can identify more meanings of an ambiguous word to increase recall and coverage in Web search disambiguation, such as by complementing the contexts discovered in tag meaning disambiguation by information obtained from dictionaries. In addition, as our method for tag meaning disambiguation requires a post-processing step of combining clusters corresponding to the same meaning of a tag, we will also investigate how this process can be incorporated into the clustering process such as by considering other clustering algorithms. Finally, we will perform further evaluations which involve larger dataset and more ambiguous tags, in order to understand the performance of our proposed method in more general cases.

References

1. C. M. Au Yeung, N. Gibbins, and N. Shadbolt. Understanding the semantics of ambiguous tags in folksonomies. In Peter Haase, Andreas Hotho, Luke Chen, Ernie Ong, and Philippe Cudre Mauroux, editors, *Proceedings of the International Workshop on Emergent Semantics and Ontology Evolution (ESOE2007) at ISWC/ASWC2007, Busan, South Korea*, pages 108–120, November 2007.
2. Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual web search engine. In *Proceedings of the seventh international conference on World Wide Web (WWW7)*, pages 107–117, Amsterdam, The Netherlands, The Netherlands, 1998. Elsevier Science Publishers B. V.
3. Douglass R. Cutting, David R. Karger, Jan O. Pedersen, and John W. Tukey. Scatter/gather: a cluster-based approach to browsing large document collections. In *Proceedings of the 15th annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR '92)*, pages 318–329, New York, NY, USA, 1992. ACM.
4. A. Gulli and A. Signorini. The indexable web is more than 11.5 billion pages. In *Special interest tracks and posters of the 14th international conference on World Wide Web (WWW2005)*, pages 902–903, New York, NY, USA, 2005. ACM.
5. "Peter Hannappel, Reinhold Klapsing, and Gustaf Neumann. MSEEC - a multi search engine with multiple clustering. In *Proceedings of the '99 Information Resources Management Association International Conference*, Hershey, Pennsylvania, May 1999.
6. Andreas Hotho, Robert Jäschke, Christoph Schmitz, and Gerd Stumme. Information retrieval in folksonomies: Search and ranking. In *Proceedings of the 3rd European Semantic Web Conference (ESWC 2006)*, pages 411–426, 2006.
7. N. Ide and J. Véronis. Introduction to the special issue on word sense disambiguation: the state of the art. *Computational Linguistics*, 24(1):1–40, 1998.
8. iProspect. Search engine user behaviour study, April 2006.
9. E. Kelly and P. Stone. *Computer Recognition of English Word Senses*. North-Holland, Amsterdam, The Netherlands, 1975.
10. Sherry Koshman, Amanda Spink, and Bernard J. Jansen. Web searching on the vivisimo search engine. *Journal of the American Society for Information Science and Technology*, 57(14):1875–1887, 2006.
11. R. Krovetz and W. B. Croft. Word sense disambiguation using machine-readable dictionaries. *SIGIR Forum*, 23(SI):127–136, 1989.
12. Michael Lesk. Automatic sense disambiguation using machine readable dictionaries: how to tell a pine cone from an ice cream cone. In *Proceedings of the 5th annual international conference on Systems documentation (SIGDOC '86)*, pages 24–26, New York, NY, USA, 1986. ACM.
13. Peter Mika. Ontologies are us: A unified model of social networks and semantics. *Journal of Web Semantics*, 5(1):5–15, 2007.
14. George A. Miller. Wordnet: a lexical database for english. *Commun. ACM*, 38(11):39–41, 1995.
15. M. E. J. Newman. Fast algorithm for detecting community structure in networks. *Physical Review E*, 69:066133, 2004.
16. M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Physical Review E*, 69:026113, 2004.
17. H. Schütze and J. Pedersen. Information retrieval based on word senses. In *Proceedings of the 4th Annual Symposium on Document Analysis and Information Retrieval*, pages 161–175, 1995.

18. Craig Silverstein, Hannes Marais, Monika Henzinger, and Michael Moricz. Analysis of a very large web search engine query log. *SIGIR Forum*, 33(1):6–12, 1999.
19. Jerzy Stefanowski and Dawid Weiss. Carrot^2 and language properties in web search results clustering. In *Proceedings of the International Atlantic Web Intelligence Conference (AWIC 2003)*, pages 240–249, 2003.
20. Thomas Vander Wal. Folksonomy definition and wikipedia. *http://www.vanderwal.net/random/entrysel.php?blog=1750*, November 2, 2005. Accessed 13 Feb 2008.
21. Xian Wu, Lei Zhang, and Yong Yu. Exploring social annotations for the semantic web. In *WWW '06: Proceedings of the 15th international conference on World Wide Web*, pages 417–426, New York, NY, USA, 2006. ACM Press.
22. Yusuke Yanbe, Adam Jatowt, Satoshi Nakamura, and Katsumi Tanaka. Can social bookmarking enhance search in the web? In *JCDL '07: Proceedings of the 2007 conference on Digital libraries*, pages 107–116, New York, NY, USA, 2007. ACM.
23. Oren Zamir and Oren Etzioni. Grouper: a dynamic clustering interface to web search results. In *Proceeding of the eighth international conference on World Wide Web (WWW99)*, pages 1361–1374, New York, NY, USA, 1999. Elsevier North-Holland, Inc.
24. Mianwei Zhou, Shenghua Bao, Xian Wu, and Yong Yu. An unsupervised model for exploring hierarchical semantics from social annotations. In Karl Aberer, Key-Sun Choi, Natasha Fridman Noy, Dean Allemang, Kyung-Il Lee, Lyndon J. B. Nixon, Jennifer Golbeck, Peter Mika, Diana Maynard, Riichiro Mizoguchi, Guus Schreiber, and Philippe Cudré-Mauroux, editors, *The Semantic Web, 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference, ISWC 2007 + ASWC 2007*, volume 4825 of *Lecture Notes in Computer Science*, pages 680–693. Springer, 2007.