Google+Facebook: A Social-Network-Optimized Web Search Approach

Yuqing Mao^{1,2}

Haifeng Shen²

Chengzheng Sun¹

¹ School of Computer Engineering, Nanyang Technological University, Singapore

² School of Computer Science, Engineering & Mathematics, Flinders University, Australia
{mao.yuqing, haifeng.shen}@flinders.edu.au, czsun@ntu.edu.sg

Social search and social network are both prevalent approaches for people seeking information on the Web. Existing work generally addresses issues of the two approaches in an isolated way. In this paper, we introduce a method for incorporating social network in social search processes. Our method integrates social networking into a novel epistemology-based social search framework, where users in a social search community contribute epistemologies that include all information derived from their search processes, and forms and utilizes social networks from their search activities to improve the social search experience. We designed and developed a prototype system in which the proposed approach is implemented. We also conducted experiments to validate this approach. The results show that the social-network-optimized social search system outperforms a conventional search engine.

Social network, social search, epistemology.

1. INTRODUCTION

Information overload has become a significant problem in web search. There is a huge amount of digital information created and replicated worldwide, and only a small fraction of the information encountered in an information seeking activity is actually relevant to the search goal. It is difficult to collect the most adaptive documents for people with vague information needs as it is difficult for people who are unfamiliar with the topical area to evaluate the retrieved results with uneven qualities. The ranking of the result pages by a search engine based on the PageRank algorithm may not well reflect the quality precedence in exploratory information seeking. In particular, user-generated contents (UGC), including those social media such as blogs or forums, and meta data such as tags or taxonomies, do not usually appear in the first search result page even if they are more relevant to the search goal.

As social activities on the Web are attracting growing interest, social search has been increasingly adopted to address these difficulties by utilizing the wisdom of crowds (Surowiecki, 2004) in recent years. Regarding that individual users' knowledge and skills are limited, social search has changed a search process from an individual activity to a social one.

There is no a widely accepted definition of social search. Generally speaking, searches carrying with human-labour are collectively called "social search", as opposed to algorithm-based searches. The core idea of social search is combining human intelligence with computer algorithms, and the central tenet of social search is "People + Algorithms > Algorithms". It is focused on how social groups can influence and potentially enhance the ability of algorithms to find meaningful information for end users, namely "better search through people" (Anderson, 2006).

On the other hand, a growing set of web-based systems are focusing on making connections and sharing common interests among geographically dispersed users. Today's popular social network services (SNS) are exemplified by *Facebook*, *Google+*, and *Twitter*. All these sites have experienced phenomenal growth in the last two years. Social networks definitely provide a valuable resource for many people to stay in touch with their friends and to meet new people they otherwise might not have the chance to meet.

However, the fusion of social search and social network has never been well addressed. Although search engines are incorporating social posts within search results and SNSs are offering the "people search engine" for searching people on one's social network, there is no unified solution that can bridge the gap between them, i.e.,

understanding the role of social network in social search. Actually, people can make friends with those who share similar information needs and then do further information seeking with them. For example, users would like to look for help from others if they are facing difficulties in their search processes. Users can find people with same hobbies or similar information requirements and thence build social networks with them. Further, effective strategies can be adopted to search for expertise in the social networks so that social search can benefit from the building of social networks, e.g., user may invite friends in their social network to join their search activities.

In this paper, we present a novel social-network-optimized social search approach called *SNOWS* (Social-Network-Optimized Web Search), which can seamlessly integrate social network into a social search system, i.e., utilize social networks to improve the quality of social search experience.

In the *SNOWS* approach, social search is supported by allowing multiple users to share, reuse, and refine the intimate search epistemologies (Mao et al. 2010) contributed by others in a social search community with same or similar search interests. Social networks are formed in a community based on the epistemologies, and social network analysis can be performed for information seekers to locate appropriate users in their social networks so that they can connect with experts in a given subject area when they need advice or help.

The *SNOWS* approach has been implemented in a prototype system *Baijia*. We have constructed the epistemology repository and social networks of the community based on automatically generated epistemologies. We also conducted a set of experiments to measure the system's performance and the results show that the social-network-optimized approach outperforms a conventional search engine in supporting Web search.

The rest of this paper is organized as follows. First we describe some work related to social search and social network. After that, we present the proposed approach utilizing social networks in epistemology-based social search, followed by the experiments with the prototype system. Finally we conclude the paper with a summary of major contributions and future work.

2. RELATED WORK

While conventional search engines are incompetent in judging the value of content in social media, studies on human factors on Web search are now in the ascendant. Collaborative

tagging systems exemplified by bookmark sharing (Delicious¹), photo sharing (Flickr²), and video sharing (YouTube³), allow users create various tags to make "shared" easy for retrieval. Such social media has been coined a name "folksonomy", which is a combination of "folk" and "taxonomy" (Smith 2004). Folksonomy has attracted increasing attention in Web document classification and search (Jaschke et al. 2007). For example, Yanbe (Yanbe et al. 2008) proposed a search model combining standard link-based ranking method with the one using data from social bookmarking. However, tags are generally different from keywords submitted to a search engine, and extra efforts are needed to facilitate search of tags (Heymann et al. 2008).

Recommender systems in the information retrieval field (Schein et al. 2002) are based on the collaborative filtering technique. For example, in Amazon.com (Linden et al. 2003), users will see "people who liked this product also like that product" when they browse a product's Web page. The community search assistant (Glance 2001) enables a community of searchers to search in a collaborative fashion by using query recommendation based on a graph where related queries are connected. I-SPY (Freyne et al. 2007) is a collaborative Web search system that can provide personalized search results for a community of users by capturing relationships between queries and results pages for particular users. Although research in these directions takes advantages of others' searches, it is not regarded as genuine social search because users are not explicitly engaged in reusing processes (Guo et al. 2009). These research areas mainly study reusing optimal queries by analyzing query logs and search results from search engines (Baeza-Yates and Tiberi 2007).

Most existing research works on social networks are focused on looking for people with specified names (Sharad et al. 2009, Monique et al. 2007). Although users can benefit from the social networks since they can connect with experts in a given subject area when they need advice or help, current social search systems (Horowitz and Kamvar, 2010) try to help users with their problems by discovering and enquiring experts only in their existing explicit social networks (e.g. social networks constructed in *Facebook*). However, new social networks can be formed during social search processes in our solution, as it is possible that people with same interest or similar information needs might network with each other. For example, some previous systems, such as Maze (Yang et al.

¹ <http://www.delicious.com>

² <http://www.flickr.com>

³ <http://www.youtube.com>

2004), allow users to make friends in the file sharing network. Moreover, based on the users' activities and epistemologies contributed, we can not only discover their interests and information needs but also identify their expertise.

Our solution emphasizes the role of the social network of users and their collaboration in a social search process. Moreover, a user's social network is established based on the epistemologies in the search community by clustering users with same or similar interest, and further analysis of the social network structure with artificial intelligence techniques such as machine learning can help discover providers who will or may generate information related to a user's information needs.

3. THE SNOWS APPROACH

3.1 The Epistemology-based Social Search Framework

The cornerstone of our approach is the epistemology-based social search framework for designing social search systems, where users can effectively collaborate on a search task/process by sharing their intimate search epistemologies. As depicted in Figure 1, the framework consists of the following major components:

Figure 1: The Epistemology-based Social Search Framework

Epistemology Search – this component is for users to reuse the shared search epistemologies. While a user types a query through the *Searching Interface*, the *Epistemology Search Engine* will first search the epistemology repository and return the relevant epistemologies. These epistemologies were contributed by other users with the same or relevant search interests or goals. If no relevant epistemology is found, the system will search the *Pages Index Base* through *APIs of existing Search Engine* (e.g., Google) and return relevant pages according to the keywords. Users can generate their own epistemologies from the result pages returned by the search engine through the *Generation Interface*.

Epistemology Generation - this component is for users to easily generate new epistemologies through the *Epistemology Generation Interface*, and store them into the *Epistemology Repository* through the *Epistemology Store Engine*.

Epistemology Repository – this component stores epistemologies contributed by a variety of user search communities. Social epistemologies are indexed by the combination of the search goal descriptions and relevant keywords for the sake of easy retrieval by the *Epistemology Search Engine*. Extra information is also stored in the *Epistemology Repository* for building social search communities, i.e., to build a social network of users based on their search interest based on the information about who have contributed related epistemologies.

Epistemology Editing & Refining – this component contains several sub-components that work together to support consumer-led interactive search by joint construction of the pre-structured epistemology.

Epistemology Services – this component has the following functions:

Social Networks Building helps users with the same or similar search goals build up online social networks to complete their search tasks together. In the next section we will describe how to utilize social networks to improve the quality of social search experience in this framework.

Incentive Mechanism provides some common services for making social epistemology-based search systems viable, reliable, and sustainable, for example, encouraging users to share their epistemologies in the social search community.

3.2 Building Social Networks from Social Search Activities

It is a common phenomenon that users would be likely to look for help from others while they are conducting search tasks, if they are unfamiliar with the subject domain of that task. If users are not sure about what they are looking for, seeking advices from experts in the right areas is always a good option. In our approach, users can find people with same interests or similar information needs from social epistemologies and thence build a social network with them. In other words, the social network is constructed from social search activities, and the constructed social network will improve social search activities in turn.

In a search process, the user usually needs to formulate a set of queries $sp = \{q_1, q_2 \dots q_n\}$, and the epistemology of this search $Epi(sp)$ is defined as:

$Epi(sp) = Epi(q_1) \cdots \oplus Epi(q_i) \cdots \oplus Epi(q_n) \oplus Epi(extra)$
where $Epi(q_i)$ is the epistemology for q_i ($1 \leq i \leq n$), and $Epi(extra)$ is the epistemology for related information that is not acquired through these queries, such as information from authoritative websites, and ' $\oplus$ ' is the operator to construct the epistemology for a search process out of those for constituent queries. For each $Epi(q_i)$, the definition is based on the user's interaction with the system. Such as pages selected by the user: $\{p_1, p_2 \dots p_m\}$, and the user's ranking and comments on the pages.

Therefore there are two spaces in the epistemology-based social search: epistemology space and user space. Each user might participate in several epistemologies and each epistemology might be contributed by several users. Figure 2 shows the two spaces in the epistemology-based social search.

Figure 2: Epistemology-based social search spaces

The connection between epistemologies can be derived from the content of each epistemology. The more two epistemologies are relevant, the shorter the distance between them is. The distance is defined based on the summation of all epistemic concepts, each of which is contributed by user i , and the match is based on the similarity between all elements of the concept in two epistemologies:

$$Dist(Epi_1, Epi_2) = \sum_i (w_1 \times D(q_i, q_k) + w_2 \times D(tags_i, tags_k) + w_3 \times D(comments_i, comments_k) + w_4 \times D(p_i, p_k)), q_i, p_i \in Epi_1, q_k, p_k \in Epi_2$$

where w_i is the weight assigned to an element of the epistemology according to its importance, e.g., an element such as a comment or a page with a higher user ranking will be assigned a heavier weight.

The similarity between two elements can be measured by various methods. In our solution, it is measured with the Kullback-Liebler divergence (KL-divergence) between two language models. For example, if a query q_1 and a query q_2 is generated by a generative model P_{q_1} of epistemology Epi_1 and P_{q_2} of epistemology Epi_2 respectively, their KL-divergence is defined as:

$$D(q_1 \| q_2) = \sum_w P_{q_1}(w) \log \frac{P_{q_1}(w)}{P_{q_2}(w)}$$

where $P_{q_1}(w)$ is the probability of generating word w by the language model for query (q_1), $P_{q_2}(w)$ is the

probability of generating word w by the language model for query (q_2).

Now we consider correlating users with epistemologies. The social network in the social search community is a weighted graph $G = (U, E)$, where each node represents a user and each edge $e = (u_1, u_2)$ is the correlation between users u_1 and u_2 . For each epistemology Epi contributed by one or more users from the community $U = \{u_1, u_2, \dots, u_m\}$, we have a set of keywords $Epi = \{k_1, k_2, \dots, k_n\}$.

We first model the contribution of a user u to an epistemology Epi using a monotone aggregate function g over the individual relevance for each keyword k in Epi :

$$Con(u | Epi) = g(rev(u | Epi, k_1), \dots, rev(u | Epi, k_n)),$$

where $rev(u | Epi, k)$ is the relevance of user u and epistemology Epi for a keyword k in Epi . The aggregation function g used in this paper is a

summation: $g = \sum_{k_i \in Epi} rev(u | Epi, k_i)$. We using a TF-IDF scoring function (Salton and Buckley, 1988) to measure the relevance, which amounts to a simplified form of BM25, as follows:

$$rev(u | Epi, k) = \frac{(p+1) \cdot freq(u | Epi, k)}{p + freq(u | Epi, k)} \times idf(k),$$

where p is an application dependent parameter, $freq(u | epi, k)$ is the overall term frequency of u given the epistemology epi and keyword k , i.e., the number of times k was quoted by user u , and $idf(k)$ is the inverse document frequency for keyword k , which is defined in fairly standard manner as follows:

$$idf(k) = \log \frac{|Epi| - |\{u | Quoted(u, Epi, k)\}| + 0.5}{|\{u | Quoted(u, Epi, k)\}| + 0.5}$$

Then the correlation between two users u_1 and u_2 can be obtained from the epistemologies they contributed $\{E1_1, E1_2, \dots, E1_m\}$ and $\{E2_1, E2_2, \dots, E2_n\}$ respectively:

$$Corr(u_1, u_2) = \sum_{i=1}^m \sum_{j=1}^n \frac{|Con(u_1, E1_i) - Con(u_2, E1_j)|}{Con(u_1, E1_i) + Con(u_2, E1_j)} Dist(E1_i, E2_j) \quad (1)$$

The rationale behind formula (1) is that if two users both made major contribution to some highly related epistemologies, they might have same or similar interest so that we can correlate them each other. For example, if a user has contributed a lot to an epistemology about "World Cup", and another user has deeply involved in an epistemology about "Messy", as these two epistemologies has many overlapped keywords, e.g., "goal", "champion", we can deduce that the two users are both soccer fans and there is a great opportunity that they can make friends online, because they can talk with and learn from each other when searching for common topics on the Web.

3.2 Exploring Social Networks for Social Search

We have built social networks of likeminded users who appear to have similar preferences in the social search community. The main purpose of exploring the social networks is to locate users that can be helpful for a user in her/his future search processes: predict potential information providers for her/him, or recommend to her/him the epistemologies contributed by Top- N trustworthy users and that she/he would like the most.

The connection between trust and user similarity has been established by Ziegler and Golbeck (Ziegler and Golbeck, 2007). They used experiments to demonstrate that there exists a significant correlation between the similarity of users and the trust expressed by them; the more similar two people are, the greater the trust between them.

In addition, user reputation is a powerful method of identifying high-quality providers over time and has been adopted in some social web applications, where reputation is based on feedback on items that user has created in the past, and serves as a signal of quality as well as an incentive to improve quality.

In our approach, epistemologies have been rated and commented by other users, and therefore the system will re-rank the epistemologies in the repository dynamically. The ranking of an epistemology is based on all received scores (one to five stars) for all pages in the epistemology, and the reputation (honest or fraudulent) and expertise (newcomer or skilled) levels of each contributor and commenter.

Based on epistemology-mediated social networks and the user reputation, our approach utilizes the user model to generate a cluster of the most similar and trustworthy users in the social network for a user and then to identify the Top- N users in the cluster that have gained highest reputation and can act as information providers or advisers to that user.

The k -means clustering algorithm is applied in the social network analysis, which partitions the users into k sets in a way that minimizes the variance within each group. For initializing k -means, the k "means" $\mathbf{m}_1^{(1)}, \dots, \mathbf{m}_k^{(1)}$ are initialized with users randomly selected from the social network. The algorithm k -means proceeds by alternating between two steps:

Assignment step: Each user is assigned to the cluster with the closest mean (i.e. partition the users according to the Voronoi diagram generated by the means).

$$S_i^{(t)} = \{\mathbf{x}_j : \|\mathbf{x}_j - \mathbf{m}_i^{(t)}\| \leq \|\mathbf{x}_j - \mathbf{m}_{i^*}^{(t)}\| \text{ for all } i^* = 1, \dots, k\}$$

where $\|\mathbf{x}_j - \mathbf{m}_i^{(t)}\|$ is the correlation between users calculated by formula (1).

Update step: Calculate the new means to be the centroid of the users in the cluster.

$$\mathbf{m}_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{\mathbf{x}_j \in S_i^{(t)}} \mathbf{x}_j$$

The algorithm is deemed to have converged when the assignments no longer change.

4. EXPERIMENTS

The main purpose of setting up experiments is to validate how much our prototype system *Baijia* can outperform a conventional search engine. The studies of human factors in social-network-optimized social search and usability of the system (including user interface evaluation) are currently on-going.

4.1 Dataset

We selected the AOL query logs (Pass et al. 2006) as the base of our experiments and initial epistemologies source. The AOL query logs consist of about 20 million search queries from about 650,000 users. Each query log is a set of {AnonID, Query, QueryTime, ItemRank, ClickURL}, where AnonID presents an anonymous user ID number, ClickURL is the URL the user clicked and ItemRank is the rank of the clicked item on the results list.

Although the dataset doesn't contain explicit users' feedback on search results, the URL clicking can be regarded as positive feedback because relative feedback signals generated from users' clicking behaviours have been proved to correspond well with explicit judgments (Joachims 2002). Therefore it is possible to backtrack users' search processes according to the query logs. Reposing on this technical foundation, we used intelligent agents to simulate users' interactions with epistemologies and search results (based on AnonID).

4.2 Procedure

In our experiments, the epistemology repository and social networks are constructed based on automatically generated epistemologies. Search epistemologies are contributed and shared through the following steps:

Step 1:

"Users" completed their searches through iterative interaction with the system and contributed their search epistemologies. To simulate the contribution from users, we extracted every user's search processes from their queries. Each search process

contains several queries that are contextually related. Cosine distance function is used to measure the contextual similarity between every two queries. We have totally extracted 1,201,497 search processes.

Step 2:

The system returned other users' search epistemologies that are relevant to the queries of the current user from its epistemology repository. To simulate the sharing of epistemologies, we retrieved the epistemology repository for relevant search epistemologies. An epistemology is relevant to a search process if its queries are similar to the search queries, and the selected pages of the epistemology completely/partially match the clicked URLs of the search.

Step 3:

If no relevant epistemology is found at step 2, the search process itself will be formulated as a search epistemology; otherwise, it will be integrated into existing relevant epistemologies. "Users" participated in the search activity by re-ranking the re-ranked results from other users or the ranked results from the search engine. To simulate the refinement of epistemologies, a computer-generated score following a Gaussian distribution is assigned to every clicked URL to represent the judgment from the current user.

Actually it is common that users may have different opinions on the same search result. As the motivation of social search is to utilize the wisdom of crowds, the result that is ranked highest by the majority is regarded as the best. For a social search system, a ranking mechanism that is based on the average scores of all participants' follows this rationale. This ranking mechanism is adopted by the *Baijia* system, which is adaptive to accumulative users' rankings, no matter whether they are computer-generated random scores or real human evaluation scores. Therefore, the score assignment following a Gaussian distribution serves our purpose for the experiments and we could envisage an even better performance gain if real human scores were used in the future experiments. In addition, URLs that are repeatedly clicked are given higher scores. The selected URLs of every epistemology are re-ranked according to the scores. We have finally built 480,254 records in the epistemology repository.

Following the above steps, we have built up the initial epistemology repository for *Baijia* by importing all search processes derived from the AOL query logs.

4.3 Results

We adopt some metrics that have been widely used to evaluate the performance of search engine, including Mean Average Precision (MAP), and Normalized Discounted Cumulative Gain (NDCG), to compare the performance of the following approaches:

The first approach is the AOL search engine, where the results are derived from the original data;

The second approach is *Baijia* without social network optimizing, where the search results are derived from all relevant epistemologies;

The third approach is *Baijia* with social network optimizing, where the search results are derived from epistemologies contributed by similar users clustered in the social network

Figure 3: MAP scores of the *Baijia* system and the AOL search engine

Figure 3 shows the MAP scores of the *Baijia* system as compared to those of the AOL search engine. As expected, the MAP score has been improved significantly in *Baijia* system and a growth in the search precision is shown when searching in *Baijia* system with the proposed SNOWS approach. This looks quite reasonable from the way our experiment was done, as while more searches are imported and more users are included in the social network, a user who involves in a search process will have a higher probability to get relevant search epistemologies and more users with similar interest whose epistemologies can be utilized to satisfy the user's information needs.

Figure 4: NDCG@10 of the Baijia system and the AOL search engine

Figure 4 shows the NDCG@10 of the *Baijia* system and the AOL search engine. We can observe that the ranking algorithm adopted in *Baijia* using the social-network-optimized approach outperforms the algorithm without social network optimizing and the algorithm adopted in AOL search engine.

5. CONCLUSION AND FUTURE WORK

Conventional search engines are incompetent in the situations where the users have difficulties in formulating proper keywords and must struggle to evaluate search results. In this paper, we propose a novel social-network-optimized web search approach to improving social search by incorporating the social network building and analysis in an epistemology-based social search system. Our work focuses on the integration of the substantial social network information into social search processes.

We have devised an epistemology-based social search framework for the design of a social search system. Epistemologies contributed by a mass of users include all information derived from the search processes. Social network services are in place to build social networks for each user in the community, and such social networks are utilized in future search process.

Furthermore, we have implemented the proposed approach in the *Baijia* prototype system, where epistemologies are formed from existing search processes, and users with the same or similar search interests are clustered in the social networks built based on the epistemologies. The system thereby can identify most similar and trustworthy information providers in the social network of a user. Through the experimental evaluation, we show that a social-network-optimized social search system outperforms a conventional search engine.

We have got some usage feedback after introducing *Baijia* on our intranet. Initial usability testing of the system has given positive feedback to the solution, which confirms the improvement of the search efficiency and quality in various social search situations. Further, a series of user studies are conducted to validate the proposed approach, including user interface evaluation and usability study.

6. REFERENCES

- Anderson, C. The long tail: Why the future of business is selling less of more. New York: Hyperion. 2006.
- Baeza-Yates, R., Tiberi, A. Extracting semantic relations from query logs. In *KDD'07, Proceedings of the 13th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, ACM Press, pages 76-85. 2007.
- Freyne, J., Farzan, R., Brusilovsky, P., Smyth, B., and Coyle, M. (2007). Collecting community wisdom: integrating social search & social navigation. In *Proceedings of the 12th international conference on Intelligent user interfaces*, Honolulu, USA, ACM Press, 52-61.
- Glance, N. S. Community search assistant. (2001). In *Proceedings of the 6th international conference on Intelligent user interfaces*, Santa Fe, USA, ACM Press, 91-96.
- Guo, F., Liu, C., Kannan, A., Minka, T., Taylor, M., Wang, Y., Faloutsos, C. (2009). Click chain model in web search. In *Proceedings of the 18th International World Wide Web Conference*, Madrid, Spain, ACM Press, 11-20.
- Heymann, P., Koutrika, G. and Garcia-Molina, H. (2008). Can social bookmarking improve web search? In *Proceedings of the 1st ACM International Conference on Web Search and Data Mining*, Stanford, USA, ACM Press, 195-205.
- Horowitz, D. and Kamvar, S. D. (2010) The anatomy of a large-scale social search engine. In *Proc. WWW '10*. ACM Press, 431-440.
- Jaschke, R., Marinho, L., Hotho, A., Schmidt-Thieme, L., and Stumme, G. Tag Recommendations in Folksonomies. *Lecture Notes in Artificial Intelligence*, 4702, pages 506-514. 2007.
- Joachims, T. Optimizing search engines using clickthrough data. In *KDD'02, Proceedings of the 8th ACM SIGKDD Conference on Knowledge*

Discovery and Data Mining. ACM Press, pages 133-142, 2002.

Linden, G., Smith, B., and York, J. (2003). Amazon.com recommendations. *IEEE Internet Computing* 7, 1, 76-80.

Mao, Y., Shen, H., and Sun, C. (2010) Supporting exploratory information seeking by epistemology-based social search. In *Proc. IUI '10*. ACM Press, 353-356.

Monique V. Vieira, Bruno M. Fonseca, Rodrigo Damazio, Paulo Braz Golgher, Davi de Castro Reis, Berthier Ribeiro-Neto: Efficient search ranking in social networks. In *CIKM'07, Proceedings of the 16th ACM conference on Conference on information and knowledge management*, ACM Press, pages 563-572. 2007.

Pass, G., Chowdhury, A. and Torgeson, C. A Picture of Search. In *Infoscale '06, Proceedings of the 1st International Conference on Scalable Information Systems*. 2006.

Schein A. I., Popescul A., Ungar L. H., Methods and Metrics for Cold-Start Recommendations, In *SIGIR'02, Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM Press, pages 253-260. 2002.

Sharad Goel, Roby Muhamad, Duncan J. Watts: Social search in "Small-World" experiments. In *WWW'09, Proceedings of the 18th World Wide Web Conference*, ACM Press, pages 701-710. 2009.

Smith. G. Folksonomy: social classification. Atomiq. http://atomiq.org/archives/2004/08/folksonomy_social_classification.html, August, 3, 2004.

Surowiecki. James (2004). *The Wisdom of Crowds: Why the many are smarter than the few and how collective wisdom shapes Business, Economies, Societies and Nations*. Little and Brown.

Yanbe, Y., Jatowt, A., Nakamura, S., and Tanaka, K. (2008). Can social bookmarking enhance search in the web? In *Proceedings of the Joint Conference on Digital Libraries*, Pittsburgh, USA, ACM Press, 107-116.

Yang, M., Chen, H., Zhao, B. Y., Dai, Y. and Zhang, Z. Deployment of a large-scale peer-to-peer social network. In *WORLDS'04, Proceedings of the 1st Workshop on Real, Large Distributed Systems*, December 2004.

Ziegler, C. and Golbeck, J. 2007. Investigating interactions of trust and interest similarity. *Decis. Support Syst.* 43, 2, 460-475.