Journal of Universal Computer Science, vol. 15, no. 4 (2009), 926-940
submitted: 30/6/08, accepted: 28/2/09, appeared: 28/2/09 © *J.UCS*

A Personalized URL Re-ranking Method using Psychological User Browsing Characteristics[1]

Shohel Ahmed
(Inha University, Korea
shohel@inhaian.net)

Sungjoon Park
(Kongju Communication and Arts College, Korea
sjpark@kcac.ac.kr)

Jason J. Jung
(Yeungnam University, Korea
j2jung@intelligent.pe.kr)

Sanggil Kang[2]
(Inha University, Korea
sgkang@inha.ac.kr)

Abstract: This paper proposes a personalized URL re-ranking method based on psychological characteristics of users browsing. The characteristics are classified into three groups, which are "common-mind," "uncommon-mind," and "extremely uncommon-mind." Our personalization method constructs an index of the anchor text retrieved from the web pages that the user has clicked during his/her past searches. Our method provides different weights to the anchor text according to the psychological characteristics for re-ranking URLs. In the experimental section, we show that our method can provide better performance than Google and another web personalization method in terms of the average rank.

Key Words: Personalization; User browsing; Search engine; Re-ranking.
Category: H.1.1, H.3.5, I.2.11

1 Introduction

Searching activities taken by users are dependent on distinct features of general search engines (e.g., domains, user interface, and query expression rules). However, given a user query, search results from the search engines are massive in amount, so that users usually spend much time to determine whether the results are relevant or not. We believe that search result provided by a search engine is collected from the web without any consideration of user intentions. Most of the user queries to the search engines are short and ambiguous, and also, other users

[1] This paper is significantly revised from an earlier version presented at The 2nd KES International Symposium on Agent and Multi-Agent Systems: Technologies and Applications (KES-AMSTA 2008) in March 2008.
[2] Corresponding author

may have completely different information needs and goals for the same query. For example, search results on "java" return a wide variety of topics, which includes Java Island, java programming, java coffee, and so on. One of the best ways to solve this problem is to personalize the search results by observing the user browsing behavior. Google, Yahoo, and MSN have already proposed their personalization methods [Sun et al. 2005]. However, they have three drawbacks;

- they need explicit input of user profile/interest,
- they do not consider that a users interest changes over time, and
- interests of a user group might not match that of individual user in all cases.

In order to solve the problem, it is needed to automatically capture user's interest and provide personalized results based on the interests.

In this paper, we propose a novel methodology to capture user browsing behavior and aim to overcome limitations of the problems. The proposed method dynamically constructs an index of the anchor texts retrieved from the web pages on his/her browsing paths. We assign unique 'weight' to the anchor texts extracted from the web pages the user has clicked. After collecting sufficient anchor texts, the proposed method uses the weight of the anchor text to re-rank the search results from a search engine. Also, we want to consider psychological characteristics obtained from user browsing behavior for elaborating this re-ranking process, because the user psychological characteristics can play a critical role on identifying user search intention. In this paper there are three types of psychological characteristics such as "common-mind," "uncommon-mind," and "extremely uncommon-mind."

In experimental section, we show the analysis and performance comparison of our method with Google and another web personalization method which is mostly relevant to the proposed method. The experimental results using the proposed re-ranking method shows that the average rank reduces significantly and test of significance is positive.

The rest of the paper is organized as follows. Sect. 2 discusses about the related work. Sect. 3 introduces our system. Experimental results and conclusions are presented in Sect. 4 and 5, respectively.

2 Related Works

This section covers various research literatures related to user profiling and ranking techniques related to personalize the web search results. The research literatures show that user profile can be a great asset to identify the user intent for future search. Generally, the user profile is created either explicitly or implicitly, and search engines recommend search results based on the discovered profile. A number of research literatures have highlighted user profiling [Liu et al. 2004, Pitkow et al. 2002, Sugiyama et al. 2004, Gauch et al. 2007, Kim et al. 2008, Kundu et al. 2008] to personalize search results. These literatures combine several techniques to learn user profile explicitly or implicitly from users' browsing histories. In explicit profiling, personalization is done by explicitly asking user profile. The inherent limitation of this approach is that the user profile may change over time. Also the studies have shown that users are quite reluctant to provide explicit input of their profile or any explicit feedback on search results.

To overcome this kind of limitations many researchers introduced various implicit profiling methods which predict user preference from users' interaction. Kelly and Teevan [Kelly and Teevan 2003] reviewed several possible approaches for inferring user preferences. They categorized user behavior across many dimensions, e.g., examine, retain, reference, and so on. Agichtein et al. [Agichtein et al. 2006] have organized the user interests as a set of features. Shen et al. [Shen et al. 2005] collects user interests by using clicked document summaries, title, and browsing history accumulated over a session. They define one session as a period consisting of all interaction for the same information need. Teevan et al. [Teevan et al. 2005] and Chirita et al. [Chirita et al. 2006] have used user's desktop to estimate their interests and construct their profiles. Also, user's web browser cache is used to construct user profiles by classifying web pages into appropriate concepts in the reference ontology [Gauch et al. 2003, Jung 2007]. The proposed method in this work is different from all these approaches in the sense that we use searchers psychological characteristics and accumulate anchor text to find out user interest.

Practically, it has become impossible for any search engine to provide search results considering individual intension as there are millions of available options. This problem is minimized by the researchers with the introduction of collaborative filtering (CF). In CF, user intention of other people is considered to provide personalized result for any individuals. Das et al. [Das et al. 2007] and Chidlovski et al. [Chidlovskii et al. 2000] have presented a personalized search result in which profiles are created using community browsing history. However, experiments are not well discussed in their paper. Also, in [Smyth et al. 2005, Balfe and Smyth 2005, Smyth et al. 2003], CF technique has been used for web personalization. They take advantage of submitted query and the result clicked by a community of users to provide personalized result for similar queries in future [Freyne et al. 2004]. The underlying assumption of CF approach is that those who agreed in past tend to agree again in future [Freyne and Smyth 2005]. The drawback of this approach is that the group interest might not match the individual interest. Besides, to provide personalized result, the considered browsing history or users' community should be very large.

Search engines usually attempt to present results more or less relevant to the user query [Haveliwala 2002]. A large community of researchers has become interested to re-rank search results provided by the search engines for a personalized web search [Jeh and Widom 2003, Haveliwala et al. 2003, Chirita et al. 2004].

Haveliwala [Haveliwala 2002] introduced an offline computation of topic-oriented page rank using ODP[3]. Jeh and Widom [Jeh and Widom 2003] constructed an incorporate personalized page rank vector using personalized hub pages. An automatic hub selection process for personalized ranking is discussed in [Chirita et al. 2004]. Also, collaborative filtering is used to re-rank web pages for personalization in [Rohini and Varma 2007, Uppuluri and Ambati 2006].

Agichtein et al. [Agichtein et al. 2006] have used their own measure and supervised machine learning technique for re-ranking search results. Dou used search engine logs for constructing user profiles [Dou et al. 2007]. They re-rank the search results by computing a personalized score for each URL returned for a query. [Boydell and Smyth 2006] They introduced four formulas for re-ranking: two methods closely relate to collaborative filtering and the other two relate to personal level.

In this paper, we introduce a novel approach to re-rank URLs by using an index comprising of anchor text collected from the URLs visited by the users. Aktas et al. [Aktas et al. 2004] firstly introduce hyperlink feature for personalization where a personalized page rank vector is constructed using some predefined profiles. However, they did not consider users' psychological characteristics which can be a critical factor for providing re-ranked URLs. We consider not only weighted anchor text extracted from the web pages clicked by user but also user's psychological characteristics for re-ranking URLs.

3 The Framework of Proposed System

The proposed system provides personalized re-ranking of URLs. An outline of the proposed system is shown in Fig. 1.

On the client side, a user issues a query and chooses a search engine from the available four options (Google, Yahoo, MSN, and Naver). The returned search results (set of URLs) are logged along with the query and user ID. If the user clicks a URL, the system logs the selected URL along with the query and user ID. The anchor text extraction module extracts anchor text from the documents clicked by the user. The extracted anchor texts are logged and used for re-ranking search results in the re-ranking web page module. The detailed description is presented in the following sections.

[3] Open Directory Project(ODP). http://dmoz.org

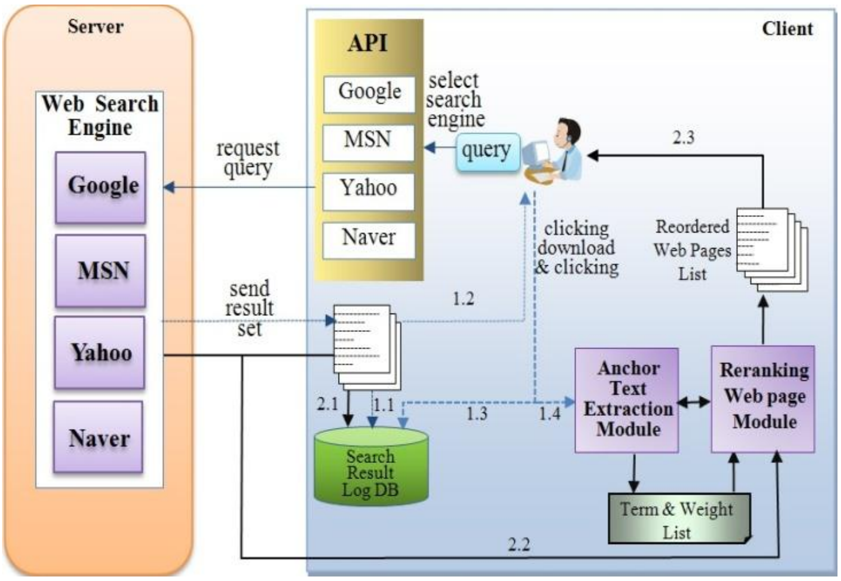

Figure 1: Overview of the proposed system

3.1 Computation of Anchor Text Weight

A search engine returns a list of URLs (denoted as set U) which are relevant with user query. The URLs clicked by the user are denoted as a set V ($V \subseteq U$). URLs clicked by any user are related to his/her personal interests. Clicking behavior of users is used as a measure of the user's interests. The clicked URLs contain anchor text. It has been reported by [Eiron and McCurley 2003] that there is a similarity between search queries and anchor text. They also showed that anchor text is a succinct description of a web page. Therefore, we extract and create an index file of anchor text from the URLs clicked by the user.

The order of web pages in a search result list indicates importance or relevance of the web pages for the query. For instance, a high ranked URL is highly relevant with the query. In this context, anchor text in a high ranked URL can also be considered as highly relevant to the query. Taking the ranking of URLs into consideration, we assign the weight to each anchor text, denoted as a_i, according to the rank of URLs containing the anchor text and store it as an index. Anchor text weight, denoted as w_i, is computed by using the rank of the URL (clicked by the user) that contained the anchor text. If an anchor text already exists in the indexed file and the same anchor text appears in another web page, the weight w_i of that anchor text is calculated by

$$w_i = \sum_j \frac{N - R_{i,j} + 1}{N} \tag{1}$$

where N is the cardinality of set U, and $R_{i,j}$ is a rank of the j-th URL containing anchor text a_i. The accumulated anchor text over a period of time represents user interests. These are further used for re-ranking search results for a new query. The value of w_i increases exponentially as the number of high ranked web pages increases. To avoid the problem, we normalize w_i to $\widetilde{w}_i$ with the log sigmoid function $\widetilde{w}_i = \frac{1}{1+exp(-s \times w_i)}$ where s the slope of the function. In general, an anchor text can be extracted from multiple web pages. Thus, the value of w_i increases, as the number of high ranked URLs containing anchor text a_i increases. As an example, assume that for a query Q a search engine returns 10 results, $U = \{U_i | i \in [1, 10]\}$. The user clicks on U_1, U_3, U_{10}, $V = \{U_1, U_3, U_{10}\}$. If the clicked URLs contain anchor text a_1, a_2, and a_3 and anchor text a_1 appears in URLs corresponding to U_1, U_3 and U_{10}, then the weight of anchor text a_1 is computed by

$$w_1 = \frac{(10 - 1 + 1)}{10} + \frac{(10 - 3 + 1)}{10} + \frac{(10 - 10 + 1)}{10} = 1.9 \tag{2}$$

where first component evaluates to 1 and the last component evaluates to 0.1.

3.2 Re-ranking of search results

To re-rank search results, we will consider user browsing characteristics. In general, those who are interested in the high ranked URL provided by a search engine can be considered "common-mind" people. If a user is interested in low ranked URL then the user can be considered as "uncommon-mind" people. It means that the characteristics of user clicking activities can influence on determining how the search results should be personalized. Also, we have to keep in mind the original rank of URLs returned by the search engine for re-ranking the search results. It can be done by taking the linear combination of i) the original rank, and ii) the index file of anchor texts along with their weights. The adjusted rank of j-th URL $\widetilde{R}_j$ can be computed by

$$\widetilde{R}_j = (1 - \beta) R_j + \beta \sum_{a_i \in A_j} \widetilde{w}_i \tag{3}$$

where A_j is a set of anchor texts of web page U_j Variable $\beta \in [0, 1]$ identifies a degree of personality (i.e., psychological characteristics) during re-ranking. It means that β can control how much $\widetilde{w}_i$ and the original search engine rank can be contributed to re-rank the search engine results. If $\beta = 0$ (i.e., he is a "common-mind" user), the new rank is same as the original one.

Rank correlation coefficient β determines the degree of personalization for user's browsing behavior. To estimate the value of β, we have collected user clicking patterns of URLs during a certain period, and analyzed the most recent clicking information of the user because we assume that recent clicking patterns have much potential to describe the user present browsing behavior.

3.2.1 Reordering by statistical testing

The clicking patterns of a user can be represented as a frequency-based distribution. We reorder the distribution according to the clicking frequency for the calculation of rank correlation coefficient. To do that, a threshold value is needed. The threshold is a tolerance for reordering URLs having close number of clicking frequencies. Using the confidence interval method [Robinson 1975], it can be obtained. The confidence interval can be estimated from the clicking histories of M users.

Average clicking frequency for the URL U_k by the users is given by

$$f_{avg,k} = \frac{1}{M} \sum_{p=1}^{M} f_{p,k} \tag{4}$$

where $f_{p,k}$ represents the clicking frequency of URL U_k by the p-th user. If the average clicking frequency for U_i and U_j is almost same, it is hard to say which one is more preferable between U_i and U_j to the users. In this case, to estimate the threshold, we assume that clicking frequency difference collected from users is normally distributed. Hence, confidence interval of clicking frequency difference can be established. The average clicking frequency difference $\Delta f_{(k,k+1)}$ is computed by

$$\Delta f_{(k,k+1)} = f_{avg,k} - f_{avg,k+1} \tag{5}$$

To find the confidence interval of clicking frequency difference, the clicking frequency difference's standard error SE is calculated by

$$SE = \frac{s}{\sqrt{n}} \tag{6}$$

where s is the standard deviation of the clicking frequency difference and n is the number of clicking differences. It can be applied when the data of clicking information is huge. In real situation, the data is sparse, so it is needed to estimate the confidence interval using a statistical distribution method.

As shown in Eq. 7, we use t-distribution to calculate the confidence interval in this paper, because t-distribution is closely related to the normal distribution and t-distribution with an infinite number of degrees of freedom is same as normal distribution.

$$C_{\Delta f} = \mu_{\Delta f} \pm |t_{(n-1),\frac{\alpha}{2}} \times SE| \tag{7}$$

$C_{\Delta f}$ is the confidence interval of clicking frequency difference, $\mu_{\Delta f}$ is the mean of average clicking frequency difference, and t is the confidence coefficient with $n-1$ degree of freedom. The value of α depends on level of confidence (LOC) such as 95% and 99%. We calculate the value of α using

$$LOC = (1-\alpha) \times 100 \tag{8}$$

According to the value of α, the value of t to be used in the calculation of confidence interval can be looked up in the standard chart of the t-distribution.

By using the confidence interval, we generate the reordered distribution by the following reordering algorithm.

Input: Confidence interval ($C_{\Delta f}$), Number of URLs (N), Clicking frequency of the URL at position q in the list ($f[q]$), Original rank of the URL at position q in the list ($R[q]$)
Output: Reordered distribution of URLs
$q = 1$;
while $q \leq N$ **do**
  if $f[q] - f[q+1] \leq C_{\Delta f}$ *and* $R[q] > R[q+1]$ **then**
    exchange $(f[q], f[q+1])$;
  end
  else $q = q + 1$;
end

Algorithm 1: Reordering algorithm

The algorithm can return the corrected positions the URLs, according to their ranks provided by the search engine if their clicking frequency difference is less than or equal to the confidence interval. It means that we respect the original rank provided by the search engine in reordering the URLs.

3.2.2 Calculation of β

Next the rank correlation coefficient (β) is calculated using the rank difference between the original rank provided by a search engine and the reordered rank. The rank difference d_i is calculated by

$$d_i = |R_j - \widehat{R_j}| \tag{9}$$

where R_j is the rank of j-th URL provided by search engine, and $\widehat{R_j}$ is the reordered rank of that URL. For example, the rank difference has minimum value zero (or maximum value), when the original and reordered rank list is same (or completely opposite). The maximum value is $\frac{n^2}{2}$ in which n is the number of URLs in the list. The normalized β is given by

$$\beta = \frac{2\sum d_i}{n^2} \tag{10}$$

4 Experiment

We collected 200 volunteers' browsing behavior over a period of three weeks and tested our method for two weeks. They are the students and faculties of INHA University in Korea. Out of them, 150 and 50 volunteers are used as training and test data, respectively. We have used Google search API[4] for sending/receiving user query/search results from servers of Google. The following section compares the proposed personalization method with Google and the other web personalization method (HFPWS) [Aktas et al. 2004] using evaluation metric *Average Rank* [Teevan et al. 2005], because HFPWS is the most related work to our method. We have tested all our results for test of significance (t-Test). The following section shows the analysis of user browsing behavior and determines the confidence interval of clicking frequency difference ($C_{\Delta f}$) from the training data. Also, we compare performance accuracy of the proposed method with the non-personalized Google search and another personalized web search method which is the most relevant one out of the personalized methods mentioned in Sect. 2.

4.1 Evaluation metric

The metric Average Rank [Gauch et al. 2003] is used for measuring the quality of personalized search. The average rank (AR) of a query q is defined by

$$AR_q = \frac{1}{|V|} \sum_{p \in V} R(p) \tag{11}$$

where $R(p)$ is the rank of URL p. The final AR over all the queries for a user is computed by

$$AR = \frac{1}{|Q|} \sum_{p \in Q} AR_q. \tag{12}$$

Smaller value of AR indicates better placement of results.

4.2 Performance Comparison

Fig. 2 is the average clicking frequency of each URL from training data with 30 URLs. It was obtained by considering only three pages of search result list per a query, because there were only a few users who have visited beyond the third page for each query in this experiment. Indeed, by Eq. 7, we obtained $C_{\Delta f} = 3.8$ for $\alpha = 95\%$.

Fig. 3 shows the performance comparison of the proposed method, Google, and HFPWS, with respect to AR The page depths mean the page length of search list considered for computing AR. Also, the comparison was done for each user

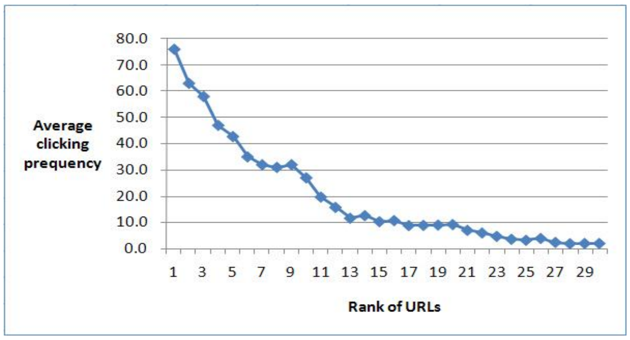

Figure 2: Average clicking frequency of each URL from training data

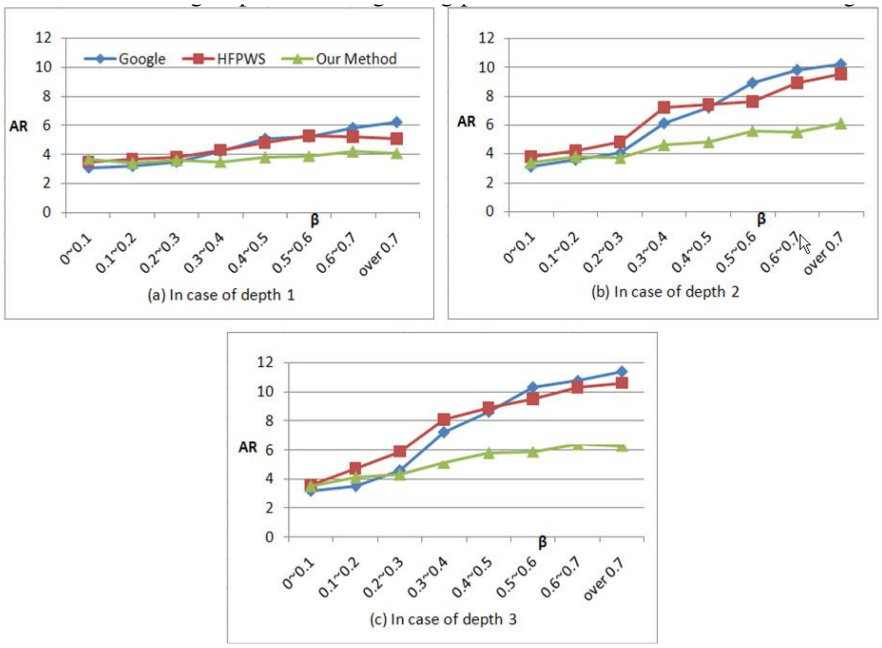

Figure 3: Performance comparison of Google, HFPWS, and the proposed method for different user groups

group organized by the range of β. For a common-mind group (i.e., $\beta = [0, 0.3]$), there is no significant difference on AR of all three methods regardless of the depth. For depth 1, ARs of Google, HFPWS, and the proposed method are 3.3, 3.7, and 3.6, respectively. Common-mind user usually prefers high-ranked URLs provided by Google search engine. The performance analyses of depth 2 and depth 3 are similar to that of depth 1.

For uncommon-mind group (i.e., $\beta = [0.3, 0.7]$), the average AR of Google, HFPWS, and the proposed method for depth 1 is about 5.0, 4.9, and 3.8, respectively. Also, in case of depth 2, the average ARs are 8.0, 7.7, and 5.1, respectively. In case of depth 3, the average ARs are 9.2, 9.2, and 5.8, respectively. For uncommon-mind user, the proposed method outperforms Google and HFPWS, because the proposed method considers the quantity of user's psychological characteristics for uncommon-minded users.

For extremely uncommon-mind group (i.e., $\beta = [0.7, 1.0]$), the average ARs of Google, HFPWS, and the proposed method for depth 1 are about 6.2, 5.1, and 4.1, respectively. Also, in case of depth 2, the average ARs are 10.2, 9.5, and 6.1, respectively. In case of depth 3, the average ARs are 11.4, 10.6, and 6.3, respectively. We can conclude that the proposed method can provide a good personalized search result for uncommon-mind group without degrading performance for the common-mind group.

To check whether there is a significant difference between means of ARs for Google/the proposed method and HFPWSA/the proposed method, we did a paired sample t-test as seen in Table 1.

Table 1: Paired t-Test summary statistics for Google and the proposed method

	N	Mean (μ)	SD	$\|\mu_1 - \mu_2\|$	t-value	$Pr > \|t\|$
Google	24	6.20 (μ_1)	2.80	1.65	4.43	0.000096**
The proposed method	24	4.54 (μ_2)	1.02			

$p^* < 0.05$, $p^{**} < 0.01$

The result of paired t-Test indicates difference between the means AR of Google and the proposed method. The difference of means is 1.65 with t-value = 4.43 and P-value = 0.000096. This indicates that the difference between the means of ARs for Google and the proposed method is significantly different at the 0.01 significance level.

Table 2 shows the result of paired t-Test for difference between the means AR of HFPWS and the proposed method. The means AR difference of HFPWS

[4] Google search API. http://code.google.com/apis/ajaxsearch/

and the proposed method is 1.73 with t-value = 6.13 and P-value = 0.000001.

Table 2: Paired t-Test summary statistics for HFPWS and the proposed method

	N	Mean (μ)	SD	$\|\mu_1 - \mu_2\|$	t-value	$Pr > \|t\|$
HFPWS	24	6.27 (μ_1)	2.36	1.73	6.13	0.000001**
The proposed method	24	4.54 (μ_2)	1.02			

$p^* < 0.05$, $p^{**} < 0.01$

This indicates that the difference of the means AR of HFPWS and our method is significantly different at the 0.01 significance level.

5 Concluding remarks and future work

In this paper, we have proposed a personalized re-ranking of URLs retrieved from an existing search engine using users' psychological characteristics such as "common-mind," "uncommon-mind," and "extremely uncommon-mind." The psychological characteristics were obtained from user browsing behaviors collected during predetermined time period. In the experimental section, we showed that the proposed method outperforms Google and another web personalization method in terms of the Average Rank for uncommon-mind users without degrading performance for the common-mind users. We have conducted a paired sample t-Test, and verified that there are significant differences between means of ARs for Google/the proposed method and HFPWSA/the proposed method, respectively. We can draw a conclusion that the proposed method is practical for saving browsing time and effort to find users' preferred web pages.

However, the comparison of performances was done by using 200 sample users. For practical purpose, we need to do the experiment using more sample users as a further work. Furthermore, a various of relatedness from psychological characteristics will be designed for enabling people to link with each other and to build contextual communities [Jung et al. 2007].

Acknowledgement

This research was supported by the Ministry of Information and Communication, Korea, under the College Information Technology Research Center Support Program, grant number IITA-2006-C1090-0603-0031.

References

[Agichtein et al. 2006] Agichtein, E., Brill, E., and Dumais, S.: Improving web search ranking by incorporating user behavior information. In E. N. Efthimiadis, S. T. Dumais, D. Hawking, and K. Järvelin, editors, *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Seattle, Washington, USA*, pages 19–26. ACM, 2006.

[Agichtein et al. 2006] Agichtein, E., Brill, E., Dumais, S. T. and Ragno, R.: Learning user interaction models for predicting web search result preferences. In E. N. Efthimiadis, S. T. Dumais, D. Hawking, and K. Järvelin, editors, *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Seattle, Washington, USA*, pages 3–10. ACM, 2006.

[Aktas et al. 2004] Aktas, M. S., Nacar, M. A. and Menczer, F.: Using hyperlink features to personalize web search. In B. Mobasher, O. Nasraoui, B. Liu, and B. M. Masand, editors, *Proceedings of the 6th International Workshop on Knowledge Discovery on the Web (WebKDD 2004), Seattle, WA, USA*, volume 3932 of *Lecture Notes in Computer Science*, pages 104–115. Springer, 2004.

[Balfe and Smyth 2005] Balfe, E. and Smyth, B.: An analysis of query similarity in collaborative web search. In D. E. Losada and J. M. Fernández-Luna, editors, *Proceedings of the 27th European Conference on Information Retrieval (ECIR 2005), Santiago de Compostela, Spain*, volume 3408 of *Lecture Notes in Computer Science*, pages 330–344. Springer, 2005.

[Boydell and Smyth 2006] Boydell, O. and Smyth, B.: Title and snippet based result re-ranking in collaborative web search. In M. Lalmas, A. MacFarlane, S. M. Rüger, A. Tombros, T. Tsikrika, and A. Yavlinsky, editors, *Proceedings of the 28th European Conference on Information Retrieval (ECIR 2006), London, UK*, volume 3936 of *Lecture Notes in Computer Science*, pages 484–487. Springer, 2006.

[Chidlovskii et al. 2000] Chidlovskii, B., Glance, N. S. and Grasso, M. A.: Collaborative re-ranking of search results. In *Proceeding of The National Conference on Artificial Intelligence 2000 Workshop on AI for Web Search*, pages 18–23, 2000.

[Chirita et al. 2006] Chirita, P.-A., Firan, C. S. and Nejdl, W.: Summarizing local context to personalize global web search. In P. S. Yu, V. J. Tsotras, E. A. Fox, and B. Liu, editors, *Proceedings of the 2006 ACM CIKM International Conference on Information and Knowledge Management, Arlington, Virginia, USA*, pages 287–296. ACM, 2006.

[Chirita et al. 2004] Chirita, P.-A., Olmedilla, D. and Nejdl, W.: Pros: A personalized ranking platform for web search. In P. De Bra and W. Nejdl, editors, *Proceedings of the Third International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH 2004), Eindhoven, The Netherlands*, volume 3137 of *Lecture Notes in Computer Science*, pages 34–43. Springer, 2004.

[Das et al. 2007] Das, A., Datar, M., Garg, A. and Rajaram, S.: Google news personalization: scalable online collaborative filtering. In Carey L. Williamson, Mary Ellen Zurko, Peter F. Patel-Schneider, and Prashant J. Shenoy, editors, *Proceedings of the 16th International Conference on World Wide Web, WWW 2007, Banff, Alberta, Canada*, pages 271–280. ACM, 2007.

[Dou et al. 2007] Dou, Z., Song, R. and Wen, J.-R.: A large-scale evaluation and analysis of personalized search strategies. In C. L. Williamson, M. E. Zurko, P. F. Patel-Schneider, and P. J. Shenoy, editors, *Proceedings of the 16th International Conference on World Wide Web, WWW 2007, Banff, Alberta, Canada*, pages 581–590, 2007.

[Eiron and McCurley 2003] Eiron, N. and McCurley, K. S.: Analysis of anchor text for web search. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*, pages 459–460. ACM, 2003.

[Freyne and Smyth 2005] Freyne, J. and Smyth, B.: Communities, collaboration and cooperation in personalized web search. In *Proceedings of the 3rd Workshop on*

Intelligent Techniques for Web personalization (ITWP'05) in conjunction with the 19th International Joint Conference on Artificial Intelligence (IJCAI '05), pages 73–80, 2005.

[Freyne et al. 2004] Freyne, J., Smyth, B., Coyle, M., Balfe, E. and Briggs, P.: Further experiments on collaborative ranking in community-based web search. *Artificial Intelligence Review*, 21(3-4):229–252, 2004.

[Gauch et al. 2003] Gauch, S., Chaffee, J. and Pretschner, A.: Ontology-based personalized search and browsing. *Web Intelligence and Agent Systems*, 1(3-4):219–234, 2003.

[Gauch et al. 2007] Gauch, S., Speretta, M., Chandramouli, A. and Micarelli, A.: User profiles for personalized information access. In P. Brusilovsky, A. Kobsa, and W. Nejdl, editors, *The Adaptive Web, Methods and Strategies of Web Personalization*, volume 4321 of *Lecture Notes in Computer Science*, pages 54–89. Springer, 2007.

[Haveliwala et al. 2003] Haveliwala, T., Kamvar, A. and Jeh, G.: An analytical comparison of approaches to personalizing pagerank. Technical report, Stanford University, 2003.

[Haveliwala 2002] Haveliwala, T. H.: Topic-sensitive pagerank. In *Proceedings of the Eleventh International World Wide Web Conference, WWW2002, Honolulu, Hawaii, USA*, pages 517–526, 2002.

[Jeh and Widom 2003] Jeh, G. and Widom, J.: Scaling personalized web search. In *Proceedings of the Twelfth International World Wide Web Conference, WWW2003, Budapest, Hungary*, pages 271–279, 2003.

[Jung 2007] Jung, J. J.: Exploiting semantic annotation to supporting user browsing on the web. *Knowledge-Based Systems*, 20(4):373–381, 2007.

[Jung et al. 2007] Jung, J. J., Juszczyszyn, K. and and Nguyen, N. T.: Centrality measurement on semantically multiplex social networks: divide-and-conquer approach. *International Journal of Intelligent Information and Database Systems*, 1(3/4):277–292, 2007.

[Kelly and Teevan 2003] Kelly, D. and Teevan, J.: Implicit feedback for inferring user preference: a bibliography. *SIGIR Forum*, 37(2):18–28, 2003.

[Kim et al. 2008] Kim, C.-G., Jung, J.-Y., Zin, H.-C., and Jung, J.-J.: An Application of Meta Search Agent System Based on Semantized Tags for EnhancedWeb Searching. *Journal of Universal Computer Science*, 14(14):2400–2415, 2008.

[Kundu et al. 2008] Kundu, A., Dutta, R. and Mukhopadhyay, D.: Design of SMACA: synthesis and its analysis through rule vector graph for web based application. *International Journal of Intelligent Information and Database Systems*, 2(4):397–421, 2008.

[Liu et al. 2004] Liu, F., Yu, C. and Meng, W.: Personalized web search for improving retrieval effectiveness. *IEEE Transactions on Knowledge and Data Engineering*, 16(1):28–40, 2004.

[Pitkow et al. 2002] Pitkow, J. E., Schütze, H., Cass, T. A., Cooley, R., Turnbull, D., Edmonds, A., Adar, E. and Breuel, T. M.: Personalized search. *Communications of the ACM*, 45(9):50–55, 2002.

[Robinson 1975] Robinson, G. K.: Some counterexamples to the theory of confidence intervals. *Biometrika*, 62(1):155–161, 1975.

[Rohini and Varma 2007] Rohini, U. and Varma, V.: A novel approach for re-ranking of search results using collaborative filtering. In *Proceedings of the 2007 International Conference on Computing: Theory and Applications (ICCTA 2007), Kolkata, India*, pages 491–496. IEEE Computer Society, 2007.

[Shen et al. 2005] Shen, X., Tan, B. and Zhai, C.: Context-sensitive information retrieval using implicit feedback. In R. A. Baeza-Yates, N. Ziviani, G. Marchionini, A. Moffat, and J. Tait, editors, *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Salvador, Brazil*, pages 43–50. ACM, 2005.

[Smyth et al. 2005] Smyth, B., Balfe, E., Boydell, O., Bradley, K., Briggs, P., Coyle, M. and Freyne, J.: A live-user evaluation of collaborative web search. In L. P. Kaelbling and A. Saffiotti, editors, *Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence, Edinburgh, Scotland, UK*, pages 1419–1424. Professional Book Center, 2005.

[Smyth et al. 2003] Smyth, B., Balfe, E., Briggs, P., Coyle, M. and Freyne, J.: Collaborative web search. In G. Gottlob and T. Walsh, editors, *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence, Acapulco, Mexico*, pages 1417–1419. Morgan Kaufmann, 2003.

[Sugiyama et al. 2004] Sugiyama, K., Hatano, K. and Yoshikawa, M.: Adaptive web search based on user profile constructed without any effort from users. In S. I. Feldman, M. Uretsky, M. Najork, and C. E. Wills, editors, *Proceedings of the 13th international conference on World Wide Web, WWW 2004, New York, NY, USA*, pages 675–684. ACM, 2004.

[Sun et al. 2005] Sun, J.-T., Zeng, H.-J., Liu, H., Lu, Y. and Chen, Z.: Cubesvd: a novel approach to personalized web search. In *WWW '05: Proceedings of the 14th international conference on World Wide Web*, pages 382–390, New York, NY, USA, 2005. ACM.

[Teevan et al. 2005] Teevan, J., Dumais, S. T. and Horvitz, E.: Personalizing search via automated analysis of interests and activities. In R. A. Baeza-Yates, N. Ziviani, G. Marchionini, A. Moffat, and J. Tait, editors, *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Salvador, Brazil*, pages 449–456. ACM, 2005.

[Uppuluri and Ambati 2006] Uppuluri, R. and Ambati, V.: Improving re-ranking of search results using collaborative filtering. In H. T. Ng, M.-K. Leong, M.-Y. Kan, and D. Ji, editors, *Proceedings of the Third Asia Information Retrieval Symposium on Information Retrieval Technology (AIRS 2006), Singapore*, volume 4182 of *Lecture Notes in Computer Science*, pages 205–216. Springer, 2006.