Proceedings of the International MultiConference of Engineers and Computer Scientists 2010 Vol I,
IMECS 2010, March 17 - 19, 2010, Hong Kong

A Novel Re-ranking System Using a Radar Chart Interface

Jianwei Zhang * Yukiko Kawai † Tadahiko Kumamoto ‡

Abstract— **The users of search engines generally input less than 2 query keywords and only browse the top ranked pages from the ordered search results. However, credible information which users want to obtain may be hidden in the pages with low ranks. When the results of an initial input query are not satisfactory, they re-retrieve pages by adding or revising keywords until they find their desired information. We propose a novel re-ranking system based on a radar chart interface which considers the importance relationship among multiple sub-keywords. The proposed system can automatically extract some sub-keywords which complement the meaning of the query keywords entered by a user, and form a radar chart interface by using the extracted sub-keywords as its chart items. Users can easily adjust the importance degrees of the extracted sub-keywords or replace them with new words by using the radar chart. Our system makes the re-ranking more effective and flexible.**

Keywords: Web search, Re-ranking, Sub-keywords, Radar chart

1 Introduction

Recently, search engines such as Google and Yahoo! are widely used to retrieve information from a large quantity of pages on the Web. These websites have the benefit of being able to search and return several thousands to several tens of thousands Web pages that contain the query keywords in a split second. Generally, users only input several query keywords and browse the search results with high ranks. Our investigation into a total of 200 individuals shows that 90% of people input less than 2 query keywords when retrieving information, and that about 80% of them check no more than the top 30 pages returned from a search engine. When the top ranked pages are not user-desired, users have to try to add or revise the query keywords one or more times until the targeted results are achieved.

Recently, research and development [2, 5, 1, 3, 4] that make the re-ranking more convenient have attracted a number of research interests. These researches can reduce users' cost of considering and inputting sub-keywords. However, the prior work only provides some interfaces by which sub-keywords are dealt with in a binary way: selected or abandoned. Sometimes, a user is not entirely explicit when conducting a retrieval. A binary usage of sub-keywords is not enough to reflect such query intent. To solve this problem, we consider the importance relationship among multiple sub-keywords, and propose a flexible re-ranking system using an interface called "radar chart" that automatically extracts sub-keywords to complement the meaning of the query keywords and allows users to easily adjust the importance of multiple sub-keywords, rather than selecting or abandoning the extracted sub-keywords.

More specifically, the search results corresponding to the initial query keywords entered by the user are first obtained using the search engine. The top fifteen keywords with the highest averages of the tf·idf values are extracted as sub-keywords. Among them, the top five sub-keywords become the items of the radar chart. Furthermore, the tf·idf average value of each sub-keyword is used as the chart value of each item. By freely changing chart items or their chart values of this radar chart, users can balance the importance of multiple sub-keywords, and consequently find pages which they want to browse by re-ranking. A query vector for the re-ranking is generated using the changed tf·idf values of sub-keywords as its elements. Each search result is also considered as a 5-dimension vector, each element of which is the tf·idf value of each sub-keyword in the page. The pages returned from the search engine are re-ranked in the descending order of the similarity between the query vector and each page's vector.

Figure 1 is a radar chart's example. When a tourist enters a query keyword "Kyoto"[1], our system extracts the words, "sightseeing", "university", "policy", "culture", and "gourmet", as the sub-keywords. Assume that their tf·idf average values are 0.3, 0.2, 0.2, 0.1, 0.1, respectively. If the tourist is especially interested in "sightseeing" and "gourmet" and is somewhat concerned about "culture" (not feeling no interest in "culture"), the tourist can set the highest scale (e.g., 10) to "sightseeing" and "gourmet" and the relatively high scale (e.g., 4) to "cul-

*Kyoto Sangyo University, zjw@cc.kyoto-su.ac.jp
†Kyoto Sangyo University, kawai@cc.kyoto-su.ac.jp
‡Chiba Institute of Technology, kumamoto@net.it-chiba.ac.jp

[1]Kyoto is a tourist city, formerly the imperial capital of Japan.

ISBN: 978-988-17012-8-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

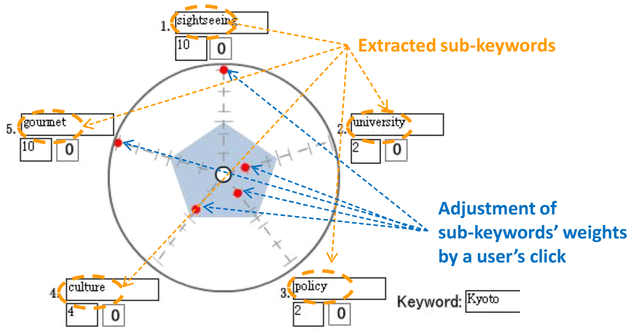

Figure 1: Example of a radar chart

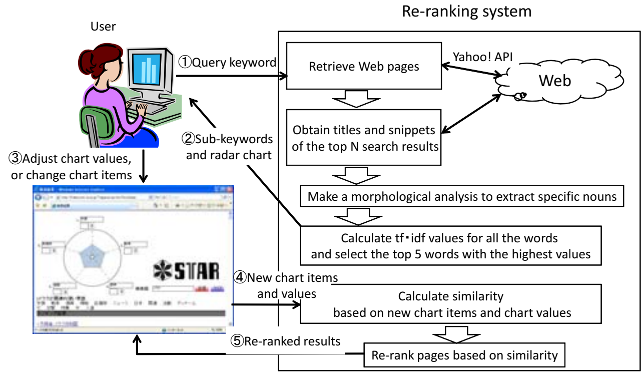

Figure 2: Overview of the re-ranking system

ture" by operating on the radar chart. The system receives the user adjustment of importance degrees of sub-keywords and changes the chart values of "sightseeing", "culture" and "gourmet" to higher values, e.g., 0.9, 0.4 and 0.9. A new query vector $(0.9, 0.2, 0.2, 0.4, 0.9)$ is generated (the chart values of "university" and "policy" are not modified and remain to 0.2), and the search results for the query "Kyoto" are re-ranked based on the similarity of the query vector, which considers the importance order: "sightseeing" = "gourmet" > "culture" > "university" = "policy".

The rest of this paper is structured as follows. In Section 2, the overview of the proposed system is presented. Section 3 describes how to extract radar chart items and how to calculate radar chart values. Section 4 shows how to adjust radar chart values. Section 5 describes the re-ranking method based on the radar chart. Section 6 shows the experimental results and their evaluations. Section 7 reviews the related work. Finally, we conclude this paper and discuss the future work in Section 8.

2 System overview

The proposed system can automatically extract multiple sub-keywords from search results and allow users to freely change these sub-keywords and their importance, so that user-desired pages can be highly re-ranked.

Figure 2 shows the overview of the proposed system. First, when a user enters one or more query keywords, the system retrieves the search results using the Yahoo! API. Then, the titles and snippets of the search results are acquired, and morphologically analyzed to extract specific nouns. Furthermore, the tf·idf value of each word for each page in which the word appears is calculated, and the words are sorted in the descending order of the average of tf·idf values for the pages in which the words appear. The top fifteen words with the highest tf·idf average values are extracted as sub-keywords and presented to the user. Among them, the top five words and the tf·idf average values are used respectively as radar chart items and radar chart values. Next, the user can adjust the value of each item in ten levels (1-10) by operating on the radar chart. The user can also replace any of the five items by selecting appropriate ones from the presented fifteen sub-keywords or using any new keywords which they can recall. After receiving new chart values or new chart items from the user input, the system generates a query vector consisting of the chart values of the five items and calculates the similarity between the query vector and page vectors. Finally, the search results are re-ranked based on the similarity.

3 Extraction of radar chart items and calculation of radar chart values

After a user enters one or more query keywords, the proposed system retrieves the top N search results using the Yahoo! API. Considering the real-time processing, our current system only collects the titles and snippets from the retrieved pages. Furthermore, these collected titles and snippets are analyzed morphologically to extract specific nouns. The tf·idf value of an extracted noun w appearing in a title or snippet's text t_i is calculated using the following formula:

$$tf \cdot idf(w, t_i) = \frac{C(w, t_i)}{N(t_i)} \times log \frac{N}{N(w)} \tag{1}$$

where $C(w, t_i)$ is the number of times that word w appears in text t_i, $N(t_i)$ is the number of words extracted from t_i, N is the number of all the retrieved pages, and $N(w)$ is the number of pages in which word w appears.

The tf·idf value values of each word w are averaged for the pages in which w appears. The top fifteen words with the highest tf·idf average values are extracted as the sub-keywords. Excluding the query keywords from them, the top five ones are used as the radar chart items. The initial radar chart value of each item is the tf·idf average value of the word.

ISBN: 978-988-17012-8-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Proceedings of the International MultiConference of Engineers and Computer Scientists 2010 Vol I,
IMECS 2010, March 17 - 19, 2010, Hong Kong

4 Adjustment of radar chart values

After the radar chart is presented, the user can set the scale of each item by selecting a value of 1-10 through the interface. The new radar chart value $y(w)$ of sub-keyword w is calculated as follows:

$$y(w) = \frac{max(tf \cdot idf(w,t_i)) - min(tf \cdot idf(w,t_i))}{9} \times (x-1) + min(tf \cdot idf(w,t_i)) \qquad (2)$$

where t_i is a title or snippet's text, and x is the selected scale (an integer value ranging from 1 to 10).

For example, a keyword w has a maximum tf·idf value 0.9 for a page and a minimum value 0.3 for another page, and a user sets the scale to 8. The conversion formula is

$$y(w) = \frac{0.9 - 0.3}{9} \times (8-1) + 0.3 = 0.77 \qquad (3)$$

The value $y(w)$ is used as the new chart value of the radar chart. The user can also replace any of the five chart item by selecting appropriate ones from the presented fifteen sub-keywords or using another new keyword if there are no desired keywords among the fifteen sub-keywords.

5 Re-ranking of initial search results

After receiving the five sub-keywords and their chart values from the user, the system generates a query vector $V_q = (v_{q1}, v_{q2}, v_{q3}, v_{q4}, v_{q5})$ using the chart value of each chart item as its element. For each retrieved page in the search results, a vector $V_p = (v_{p1}, v_{p2}, v_{p3}, v_{p4}, v_{p5})$ (p = 1,...,N) is determined, each element of which is the tf·idf value of each sub-keyword. The similarity between vector V_q and vector V_p is calculated using the following formula:

$$sim(V_q, V_p) = \frac{v_{q1}v_{p1} + \cdots + v_{q5}v_{p5}}{\sqrt{{v_{q1}}^2 + \cdots + {v_{q5}}^2} \times \sqrt{{v_{p1}}^2 + \cdots + {v_{p5}}^2}} \qquad (4)$$

The search results are re-ranked in a descending order of the similarity and presented to the user. When the radar chart value of a sub-keyword is set to 10, the pages with the highest tf·idf values for that sub-keyword tend to be ranked to the top place. When the radar chart value of a sub-keyword is set to 1, the pages with the highest tf·idf values for that sub-keyword tend to be given a lower ranking.

6 Experimental evaluation

We implemented a prototype of the proposed system [6]. The titles and snippets of the top N=150 pages returned from the search engine were used. Figure 3 is a snapshot of the initial search result with the corresponding radar chart for the query "influenza". "provision", "avian", "information", "prevention" and "cold" are extracted as the initial five radar chart items. Figure 4 shows the re-ranked results after changing the sub-keywords' importance. Using the radar chart, users need not reconsider the sub-keywords by themselves, and they can freely and easily change relative importance of the extracted sub-keywords.

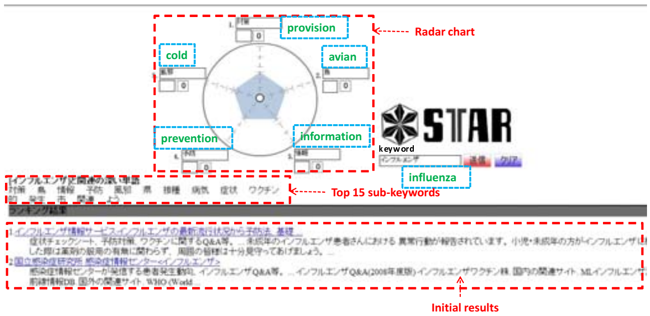

Figure 3: Initial search result

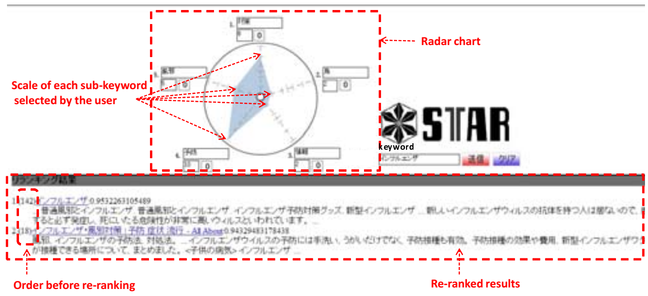

Figure 4: Re-ranked search result

We next describe the evaluation on the extracted sub-keywords' appropriateness and the re-ranking effect of search results based on the author's judgment (Section 6.1) and the questionnaires answered by 100 individuals (Section 6.2).

6.1 Evaluation by the author

6.1.1 Chart item appropriateness

The author evaluated whether the extracted sub-keywords for 10 selected query keywords were appropriate as their complement. The evaluation was one of three levels: "appropriate", "neither appropriate nor inappropriate", and "inappropriate". The evaluation result is shown in Figure 5. For most of the query keywords, two or three of the top five sub-keywords were appropriate, and at least one appropriate item was included in the sub-keywords for all the ten query words. This indicates our method can extract useful words for the re-ranking.

ISBN: 978-988-17012-8-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Also, since the chart items and chart values can be easily changed using the interface of the radar chart, the inappropriate items can be assigned smaller chart values, or be replaced by words that the user wants to use.

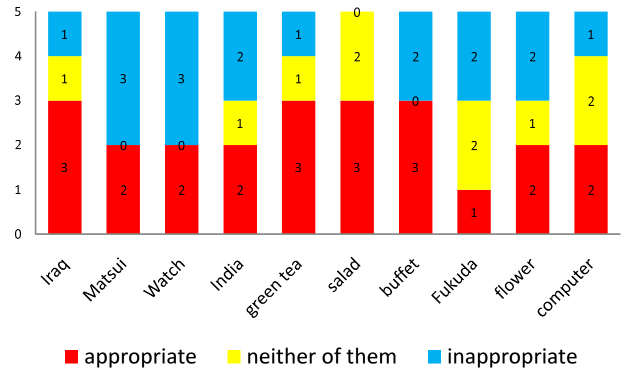

Figure 5: Evaluation of extracted sub-keywords by the author

6.1.2 Re-ranking effect

We also compared the precision and recall of the initial retrieval results and re-ranked results for four query keywords. For each query keyword, the top 50 pages were evaluated and assigned one of three levels: "Y(es)", "N(o)" , "B(order)". The pages regarded as desired ones were marked as Y, the undesired pages were marked as N, and the pages which were difficult to be categorized into Y or N were marked as B. Furthermore, precision and recall were calculated focusing on the top 10 pages as follows:

$$\begin{aligned} precision &= \frac{number\ of\ pages\ marked\ as\ Y\ in\ 10\ pages}{10} \\ recall &= \frac{number\ of\ pages\ marked\ as\ Y\ in\ 10\ pages}{number\ of\ pages\ marked\ as\ Y\ in\ 50\ pages} \end{aligned}$$

In the experiments, the scales of the sub-keywords which were evaluated as appropriate in Section 6.1.1 were set to the largest values (10), and the scales of the inappropriate sub-keywords were set to the smallest values (1). Figure 6 and Figure 7 show that both precision and recall are remarkably improved for all the four query keywords after search results are re-ranked. The detailed evaluation of the top 10 pages before and after re-ranking are shown in Table 1. As we can see, the appropriate pages are ranked to the higher place.

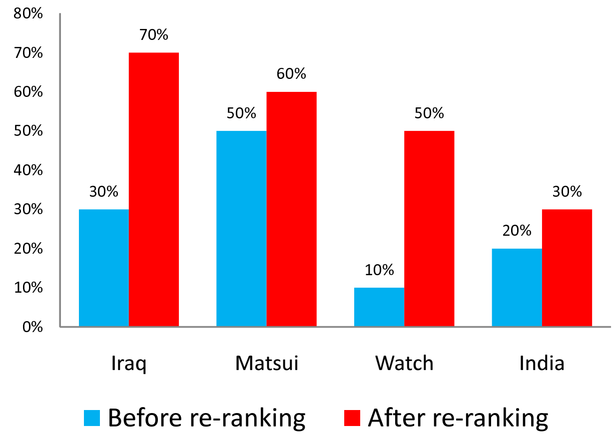

Figure 6: Precision

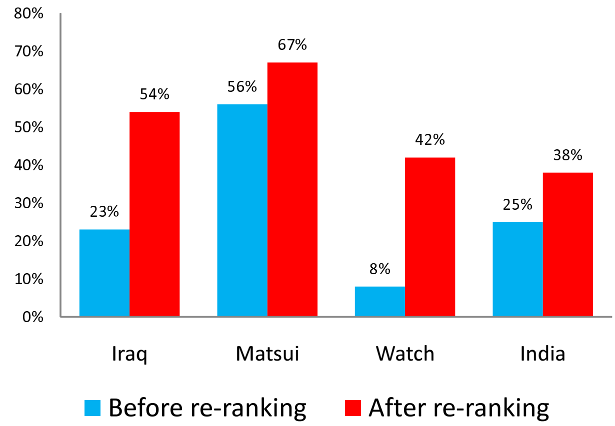

Figure 7: Recall

6.2 Evaluation by 100 individuals' questionnaires

6.2.1 Chart item appropriateness

To evaluate the overall appropriateness of the sub-keywords, questionnaires were filled out by 100 individuals. These individuals inputted their query keywords five times and gave their overall evaluation of the returned sub-keywords from four levels: "appropriate", "somewhat appropriate", "somewhat inappropriate", "inappropriate". Figure 8 shows the appropriateness evaluation of the five radar chart items. The percentage of individuals who gave the evaluation "appropriate" (7%) and "somewhat appropriate" (53%) was 60%. Figure 9 is the evaluation result of the top 15 sub-keyword appropriateness. The percentage of "appropriate" (7%) and "somewhat appropriate" (63%) increased to 70%. This indicates that the extracted sub-keywords are useful for complementing the query keywords.

ISBN: 978-988-17012-8-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Table 1: Top 10 search results and their evaluation

Query keywords		Ranking									
		1	2	3	4	5	6	7	8	9	10
Iraq	before	B	Y	B	N	Y	N	N	B	Y	B
	after	Y	Y	B	B	Y	B	Y	Y	Y	Y
Matsui	before	N	Y	Y	N	N	Y	Y	N	Y	B
	after	Y	Y	Y	Y	Y	N	B	B	Y	B
Watch	before	N	N	B	B	B	B	Y	N	B	B
	after	Y	Y	Y	Y	Y	B	N	B	B	N
India	before	Y	B	Y	B	B	B	N	B	B	N
	after	Y	B	B	B	B	Y	B	B	Y	B

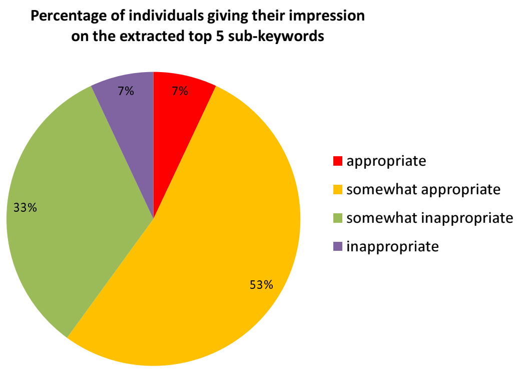

Figure 8: Evaluation of the extracted top 5 sub-keywords by 100 individuals

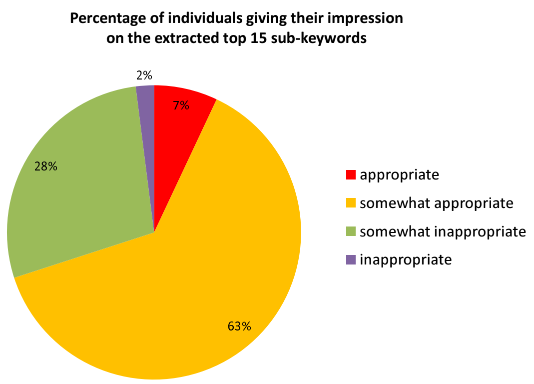

Figure 9: Evaluation of the extracted 15 sub-keywords by 100 individuals

6.2.2 Re-ranking effect

We randomly selected six query keywords and specified a page as the desired page for each query keyword. A total of 100 individuals evaluated the time that they found the specified page before re-ranking and after re-ranking. They were asked to select the level of the time for catching the desirable pages from five options: "found immediately", "found somewhat immediately", "found by spending time", "found by spending a long time", "not found".

Figure 10 shows the percentages of individuals who gave evaluation at each level before and after re-ranking for the query keyword "Iraq". The percentage of individuals who thought the specified pages was "found immediately" increased from 11% to 34%. The percentage of "found somewhat immediately" increased from 27% to 33%. Also, the percentage of "not found" decreased from 15% to 1%.

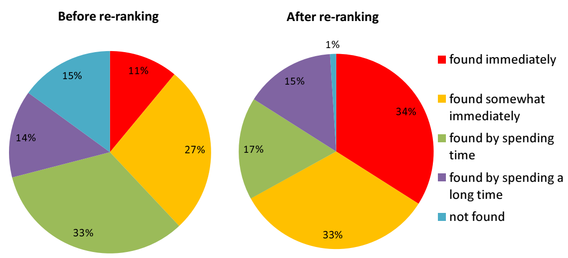

Figure 10: Re-ranking effect for the query keyword "Iraq"

Focusing on the total number of individuals who gave the evaluation levels "found immediately" and "found somewhat immediately", the improvement of percentage of individuals for all the six selected query keywords after re-ranking is shown in Figure 11. The total numbers of these two high evaluation levels after re-ranking were about 2-3 times than those before re-ranking. This indicates that our re-ranking method is very effective.

ISBN: 978-988-17012-8-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

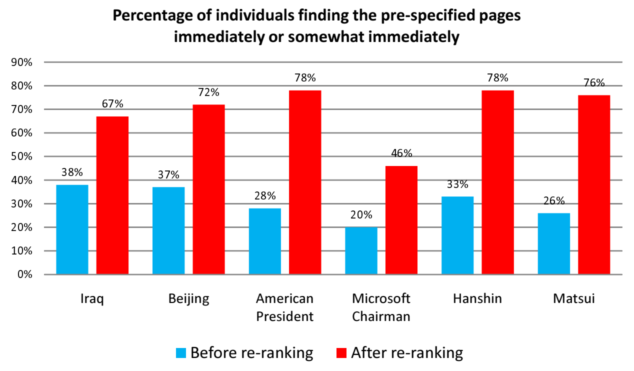

Figure 11: Re-ranking effect for the six query keywords

7 Related work

There have been a number of studies on interfaces which help users perform effective and efficient re-ranking. Xu et al. [2] developed a technique of query expansion by analyzing word relationships and documents retrieved by the initial query. Zeng et al. [5] organized Web search results into clusters. Specially, their method extracted and ranked salient phrases as candidate cluster names and assigned documents to relevant salient phrases to form candidate clusters. Seki et al. [1] proposed a multiple viewed search engine. Considering two aspects of search results, a matrix of the distribution of the clustering of them was generated. The characteristic words of each cluster were displayed in the matrix which supported narrowing the search. Yamamoto et al. [3] developed a re-ranking system which depends on users' deletion and emphasis operations. These operations are supported by using Tag-Clouds. Yoshida et al. [4] developed an interface by which users could perform a complex AND, OR, NOT search. The important terms were extracted from the search results and displayed on a two-dimensional graph. Our work also intends to help users easily perform a re-ranking of search results. However, the radar chart interface is different from the above methods because the importance relationship among multiple sub-keywords is considered in our system.

8 Conclusions and future work

In this paper, we proposed a retrieval system using a radar chart interface which automatically extracted multiple sub-keywords from search results based on tf·idf values and allowed users to easily combine and set the importance of the sub-keywords. The radar chart was generated by using the extracted sub-keywords as the chart items and using the tf·idf values as the chart values. The experimental evaluation indicated that most of extracted sub-keywords were appropriate and the re-ranking effect was good.

In the current system, real-time processing was taken into account, and thus only the titles and snippets of the retrieved pages were used. In the future, all the text of the retrieved pages will be analyzed. The tf·idf values were used for the extraction of sub-keywords. Co-occurrence frequency is also a reasonable factor which will be considered in our future work. Comparison between radar chart and other visualization styles is an interesting work which will be done. Also, we consider expanding the radar chart system to evaluate the credibility of Web information.

Acknowledgments

This work was supported in part by the National Institute of Information and Communications Technology, Japan, and by the MEXT Grant-in-Aid for Young Scientists (B) (#21700120, Representative: Yukiko Kawai).

References

[1] T. Seki, T. Wada, Y. Yamada, N. Ytow, and S. Hirokawa. Multiple viewed search engine for an e-journal - a case study on zoological science. In *HCI (4)*, pages 989–998, 2007.

[2] J. Xu and W. B. Croft. Query expansion using local and global document analysis. In *SIGIR*, pages 4–11, 1996.

[3] T. Yamamoto, S. Nakamura, and K. Tanaka. Rerank-by-example: Efficient browsing of web search results. In *DEXA*, pages 801–810, 2007.

[4] T. Yoshida, S. Nakamura, S. Oyama, and K. Tanaka. Query transformation by visualizing and utilizing information about what users are or are not searching. In *ICADL*, pages 124–133, 2008.

[5] H.-J. Zeng, Q.-C. He, Z. Chen, W.-Y. Ma, and J. Ma. Learning to cluster web search results. In *SIGIR*, pages 210–217, 2004.

[6] Radar Chart.
`http://klab.kyoto-su.ac.jp/~zjw/cgi-bin/radar/`

ISBN: 978-988-17012-8-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)