Mining Navigation Behaviors for Term Suggestion of Search Engines

Hahn-Ming Lee[1], Chi-Chun Huang[2] and Wei-Tang Hung[3]
[1] Department of Computer Science and Information Engineering
National Taiwan University of Science and Technology
hmlee@mail.ntust.edu.tw
[2] Department of Information Management
National Kaohsiung Marine University
cchuang@mail.nkmu.edu.tw
[3] Department of Electronic Engineering
National Taiwan University of Science and Technology
bage@doya.url.com.tw

Abstract- *Query expansion is extensively applied in information retrieval systems, such as search engines. Most conventional approaches to query expansion have been developed based on textual analysis of documents. However, different issues such as segmentation and feature selection must be addressed, which might influence performance seriously. This work focuses mainly on avoiding the above problems of textual analysis and thus proposes a collaborative method of applying access logs in the search engines to term suggestion (i.e., query expansion). A co-clicked behavior-based term suggestion is presented to suggest user-oriented terms. Analyzing the co-clicked behaviors of users in the access logs for term suggestion eliminates the need to perform textual analysis and provides some positive characteristics that previous approaches neglected, such as content independent, adaptability, and extensibility. Furthermore, limitations of current search engines, including problems of word mismatch and partial match, can also be overcome. Here, a search engine prototype is also developed to demonstrate the results of term suggestion. Experimental results demonstrate that the precision rates of information retrieval systems can be improved by using the suggested terms of the proposal presented herein. The proposed term suggestion also performs better than another famous Chinese term suggestion system, namely OPENFIND.*

Keywords: query expansion, term suggestion, mining navigation behaviors, access logs, search engines.

1. Introduction

The Internet has recently grown in popularity [19] and now carries a diverse range of information [21]. However, the explosive growth of the new medium is actually making the process of finding useful information increasingly difficult. To overcome this problem, search engines [22,24,25] provide an interface for users to quickly find the information they require. Actually, search engines have been demonstrated to be one of the most popular applications on the Internet [1]. However, various search engine-related problems persist, such as short query term, word mismatch, and partial match [5,7,12]. To address these problems, query expansion (also called query suggestion or term suggestion) [11,12,15,17,20] is used to help users finding the information they require effortlessly.

In general, search engines return result lists for query terms entered by users based on matching similarities [22,24,25] and ranking algorithms [1,2,3,4,6]. Traditional query expansion methods analyze the relationships among user query term and documents to automatically or interactively generate some suggested terms for users. Reformulating the initial query term of a user by adding suggested terms may enhance the precision and recall rates of the retrieved documents. Nevertheless, the fact that suggested terms are generated from documents may create the following disadvantages [20]: 1) sensitivity to document quality, 2) ranking algorithm bias, 3) difficulty in comprehending suggested terms, and 4) poor capability to handle new associations of terms. These disadvantages may influence query expansion performance.

This work proposes a novel method of term suggestion by considering user navigation behavior. When users click on certain documents, it is assumed that they are interested in these documents [18]. Different query terms are thus assumed to be related if users click on the same document after entering them [18]. Based on this concept, the navigation behaviors of users derived from the access logs in the search engines are mined to suggest query terms. The suggested terms in the novel method are extracted from the access logs based on the popularities of terms and clicked hyperlinks. Since the access logs contain user information, such as

navigation behavior, the candidate terms are user-oriented.

Regarding the proposed term suggestion, the suggested terms are originally entered by users, and thus must be recognized and comprehended by users in advance, i.e., these suggested terms are implicitly verified by the users in advance. Consequently, these suggested terms must be considered reasonable for other users to comprehend. Therefore, the quality of the suggested terms in the novel method is expected to be reliable. Furthermore, the proposed method does not consider the relationships among query term and documents as conventional query expansions do. Consequently, the novel method is not influenced by document quality and ranking algorithm bias. Besides, log-based term extraction used herein is clearly adaptive. The proposed term suggestion can thus handle new associations of terms.

The rest of this paper is organized as follows. The concept used herein for term suggestion is described in Section 2. The system architecture of the proposed term suggestion is presented in Section 3. Some characteristics of the proposed approach are detailed in Section 4. In Section 5, a search engine prototype is developed to demonstrate the results of the proposed term suggestion. In addition, experimental results are given to demonstrate the performance of the proposal presented herein. Finally, conclusions are made in Section 6.

2.Co-clicked Behavior-based Term Suggestion

This section describes the concept used herein for term suggestion. For a query term, a search engine generally returns a lot of hyperlinks corresponding to different Web sites (or pages) and Web site (or page) information in the result list. The Web site (or page) information, as illustrated in Table 1, includes title, short description, and Web site (or page) category for each Web site (or page). Users use the above information to decide which hyperlinks in the result list to click on. After a hyperlink is clicked, the corresponding Web site (or page) will be displayed. This work thus assumes that the clicked hyperlinks (i.e., Web sites or pages), which have been confirmed by numerous users and contain the required information, must be highly related to the query term entered by these users. Obviously, this assumption is reasonable.

Table 1. An example of result list with query term "AI" in the search engine "Google" [22]

Title	Short Description	Web Site Category
AI Artificial Intelligence	AI Artificial Intelligence AI Artificial Intelligence DVD for ratings reasons	Arts > Movies > Titles > A > A.I.
AI on the Web	AI on the Web. This page links to 874 pages around the web with information on Artificial Intelligence.	Computers > Artificial Intelligence

Figure 1 gives an example of the concept employed herein for term suggestion. In this figure, numerous users click on the same hyperlink corresponding to Web site "911 Crime", via two different query terms, "Bin Laden" and "911", respectively. The Web site "911 Crime" can then be assumed to contain information required by these users. Meanwhile, the required information in the Web site "911 Crime" is highly related to the query terms "Bin Laden" and "911". Assuming that query terms "Bin Laden" and "911" are highly related in this case is thus reasonable. Here, the query terms "Bin Laden" and "911" can be viewed as synonym. Therefore, if other users enter the query term "911", then the term "Bin Laden" can also be suggested to them as an alternative search term. Restated, according to the co-clicked behaviors of users, highly related terms are grouped for term suggestion.

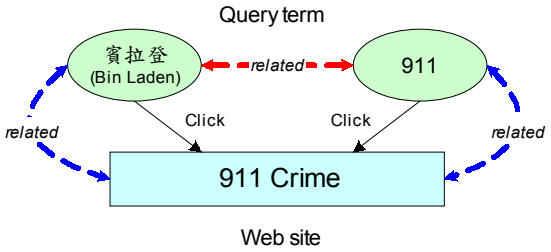

Figure 1. An example of the concept used herein for term suggestion.

3.System Architecture for Term Suggestion

This section describes the system architecture of the proposed term suggestion, shown in Figure 2. The Log Extractor extracts the access logs and stores the extracted results in the Log DB. The Term Refiner is then used to preprocess the extracted results, build the relations among various query terms and hyperlinks, and then weight these terms. Accordingly, the Suggestion Term DB stores the relations among query terms and hyperlinks. While the novel system receives a query term, the Suggestion Term DB will be searched to suggest query terms. The following subsections details each component of the proposed term suggestion.

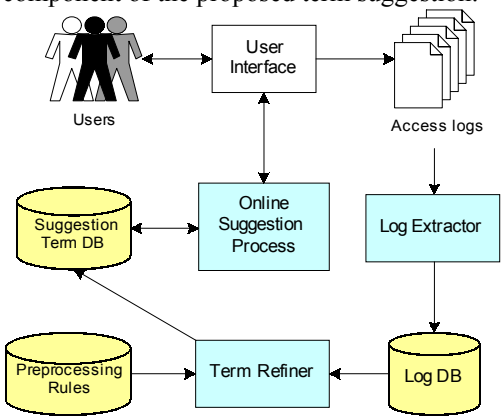

Figure 2. System architecture of the proposed term suggestion.

3.1 Log Extractor

The Log Extractor extracts the original access logs and stores the extracted results in the Log DB. For term suggestion, the Log Extractor only extracts the queries of users, along with their clicked hyperlinks (i.e., Web sites or pages), and other information such as date and IP. Noises are filtered using the Log Preprocessing component of the Term Refiner. The dataset used in this work is WWW server access logs (from Nov. 18 to Nov. 25, 2001) derived from the YAM portal site [25] (a portal site that collects numerous Chinese Web sites). A database Log DB with about three million records is created by extracting the original access logs.

3.2 Term Refiner

The Term Refiner contains two components: Log Preprocessing and Term Weighting. The Log Preprocessing component processes the Log DB and builds the relations among query terms and hyperlinks (i.e., Web sites or pages). The Term Weighting component then weights each term. Accordingly, these terms are stored in the Suggestion Term DB, which will be searched to suggest query terms. The above two components are detailed in Sections 3.2.1 and 3.2.2, respectively.

3.2.1 Log Preprocessing

The Log Preprocessing component uses the following two preprocessing rules to filter noises for the Term Refiner:

a. **Removal of special symbols**. Symbols such as "+", "&", and so on are removed from the database to improve the processing performance.

b. **Removal of infrequently-used terms**. In general, users often enter hot query terms (i.e., frequently-used query terms) to find the desired information. To reduce the complexity in the proposed method, infrequently-used query terms are removed, that is, only hot query terms are retained.

3.2.2 Term Weighting

This component calculates the weight of each search term based on two weighting factors, namely support and fitness. The weight is then used to rank the suggested terms. These two weighting factors are:

a. **Support.** Let S_{ij} be the number of times that users have clicked on hyperlink *j* after entering query term *i*. A high value of S_{ij} indicates that query term *i* is closely related to hyperlink *j*. This factor, namely *support*, is thus considered for term weighting. Notably, different search terms corresponding to the same clicked hyperlink have different values of S_{ij}. In the proposed term suggestion, S_{ij} will be normalized as follows.

$$NS_{ij} = \frac{S_{ij}}{\max\{S_{1j}, S_{2j}, \ldots, S_{nj}\}}, \quad (1)$$

where NS_{ij} is the *normalized support* of query term *i* corresponding to the clicked hyperlink *j*, and n is the total number of query terms corresponding to the clicked hyperlink *j*. In practice, the normalized support is employed herein for term weighting.

b. **Fitness**. Consider that users are provided with k hyperlinks in ranked order 1, 2, …, k, respectively after entering a query term *i*. Let R_j be the ranked order of hyperlink *j*. An average number, namely *fitness*, can be obtained as follows.

$$I_i = \frac{\sum_{j=1}^{k} P_j}{n}, \quad (2)$$

where n is the total number of clicked hyperlinks, and k is the total number of presented hyperlinks. If hyperlink *j* is clicked, then $P_j = R_j$; otherwise $P_j = 0$. In general, a query term *i* with I_i close to one means that users can quickly find their required information, i.e., web site or page with highest ranking, after entering term *i*. This term is thus appropriate to be a suggested term. Thus, the factor fitness can also be considered for term weighting.

By considering the above two factors, the normalized weight W_{ij} of each query term *i* corresponding to hyperlink *j* is calculated by using the following formula:

$$W_{ij} = \frac{NS_{ij} + \frac{1}{I_i}}{2}, \quad (3)$$

where NS_{ij} is the normalized support of query term *i* corresponding to clicked hyperlink *j*, and I_i is the fitness of query term *i*. Consequently, each pair of terms and clicked hyperlinks, which is stored in the Suggestion Term DB, is associated with its own weight. The normalized weight W_{ij} is then used in the Online Suggestion Process to rank suggested terms. That is, query terms with high normalized-weight will be suggested to users with high ranking. Following the generation of new access logs and the execution of the Log Preprocessing, the Term Weighting must be done immediately to ensure the consistency of the Suggestion Term DB. Consequently, the execution period of the Term Weighting must be set the same as that of the Log Preprocessing.

4. Characteristics of the Proposed Term Suggestion

From the concept and operations described in the previous sections, some characteristics of the novel term suggestion approach can be summarized as follows.

a. **User-oriented**. The suggested query terms, referring to terms submitted by users, are assumed to be able to be recognized and comprehended by users, otherwise they would not be submitted. Thus, the suggested terms must be user-oriented.

b. **No textual analysis is needed**. Since the novel approach is based on co-clicked behaviors of users, the Web site contents will be verified by numerous users who click the Web site. As a result, no textual analysis is performed.
c. **Adaptability**. Regarding new trends or events that create new associations of terms, the novel approach is able to handle these new associations. For example, the term "meteoric shower" is highly related to the term "astronomy". Thus, the suggestion results for "meteoric shower" would be related to those for "astronomy" in most search engines. However, the suggested terms for "meteoric shower" obtained by the novel approach also contain the term "F4" (see Table 2), a newly popular Taiwanese pop group that once released a well-known song entitled "meteoric shower". Here, the term 'F4' is suggested (regarding meteoric shower) because of the co-clicked behaviors of numerous users who click the same Web site. The fact that the novel approach includes the pop group in its suggestion results demonstrates that the novel approach is capable of handling new associations and catching current trends in the meanings of terms.
d. **Extensibility**. Since the navigation behaviors in access logs are considered for term suggestion, the novel approach can be easily applied to various types of applications, such as query systems for image databases.
e. **Good quality**. Owing to the log-based term extraction and the use of normalized support and fitness (i.e., two weighting factors in Section 3.2), unimportant or bad quality terms will eventually be filtered. The quality of the suggested terms in the novel approach is thus enhanced.

5. Experimental Results

In this section, a search engine prototype is developed to show the results of the proposed term suggestion. Accordingly, experimental results are given to demonstrate the performance of the proposal presented herein.

5.1 Search Engine Prototype

Figure 3 illustrates the architecture of the proposed search engine prototype, which is built by using a Windows 2000 platform, with an IIS 5.0 WWW server and a SQL 2000 database. Meanwhile, the programming language used to build the system is PHP 4.0. Since the prototype is built for experimental use, this work only implements basic search engine functions (such as keyword matching), that is, users can only submit their query terms to the system.

After a user entering a search term, the proposed prototype will provide two functions, normal search and direct search, to the user. Normal search is identical to the search method of traditional search engines, namely searching the matched Web sites by database scanning. With direct search, first the database is scanned to identify the hyperlink which has the highest clicked times corresponding to the search term. Accordingly, the user will be directly presented with a Web site corresponding to that hyperlink. Google also has a similar function, named "I'm Feeling Lucky".

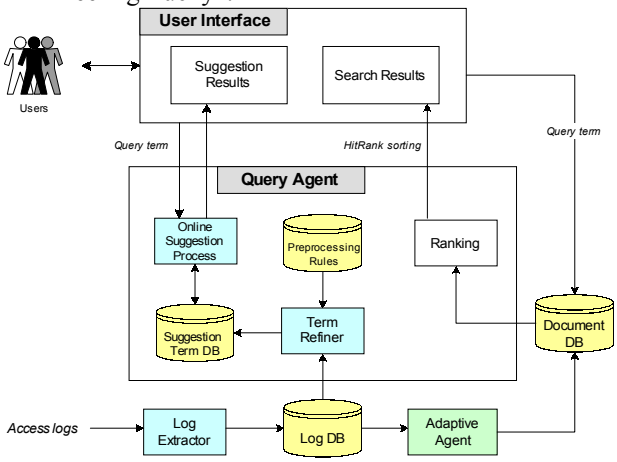

Figure 3. The proposed search engine prototype

The result list of the proposed prototype, as shown in Figure 4, contains two parts: suggested term list and matched Web site list. The suggested terms are ranked according to their weights, as stated in Section 3.2. Also the matched Web sites are ranked using the HitRank algorithm [14]. Once users use the suggested terms with the logic 'AND' or 'OR' operations to reformulate the initial query term, the reformulated query term will be sent to the query server again to return search results.

Query Term: CD-ROM
Web Matches: 66 Page Number: 1 of 7 *HitRank*
Suggested Query Terms:
⊙ *GISH* (1) ⊙ *AACOM* (1) ⊙ *ACER* (0.57) ⊙ *FUJITSU* (0.44) ⊙ *SCSI* (0.38) ⊙ *ELOHA* (0.33) ⊙ *CD-Recorder* (0.28) ⊙ *ACER Technology* (0.23) ⊙ *MO* (0.21) (a)

Web Matches:
- *ACER Technology* [1.42169] – Agent of High-performance desktop PCs, Notebooks, Tablet PC.
- *Chung Sheng Computer Ltd.* [1.19446] – Located in Kaohsiung, Retailer of computers and related products, including CPU, RAM, Motherboard, video accelerator, monitor, keyboard, CD-ROM, Speaker System, sound cards, modem, scanner, mouse, etc.
- *Advance Application Technology Inc.* [1.12365] – R&D designer, manufacturer and exporter of digital products, including LCD monitor, DVD-Player and Portable DVD-Player.
- *wangman.boys.com.tw* [1.04725] – introduces MP3 CD-ROM, Walkman and DVD-Player.
- *www.yuanyu.com.tw* [0.99792] – Agent of digital products, including camera, CD/RW and microscope.

First Page 1 2 3 4 5 6 7 *Last Page* (b)

Figure 4. An example of result list of the proposed search engine prototype. (a) Suggested Terms. (b) Result list with HitRank ranking

5.2 Precision and Recall Rates

The dataset used herein for term suggestion is WWW server access logs (from Nov. 18 to Nov. 25,

2001) derived from the YAM portal site [25] (a famous portal site that collects numerous Chinese Web sites). The access logs used for term suggestion contain 24130 records. Extracting the original access logs and storing the extracted results in the Log DB creates a database around three million records. In this database, most query terms are in Chinese.

To demonstrate the performance of the proposed term suggestion, the following two rates, namely recall rate and precision rate, were used.

$$\text{Precision} = \frac{\text{Number of Retrieval and Relevant Documents}}{\text{Number of Total Retrieval Documents}} \quad (4)$$

$$\text{Recall} = \frac{\text{Number of Retrieval and Relevant Documents}}{\text{Number of Total Relevant Documents}} \quad (5)$$

As stated earlier, after entering a query term, a user will be presented with a result list and suggested terms. Various presented results corresponding to different query terms generally have different recall rates. Thus, the precision rates corresponding to these presented results with recall rates varying from 0.1 to 1.0 can be calculated, respectively. In the experiments, two kinds of precision rates can be obtained when the original query term is reformulated by adding the suggested terms of the proposal used herein or not. Experimental results, as shown in Figure 5, demonstrate that the average precision rates can be enhanced (13.67%) when the proposed term suggestion was applied. That is, the proposal presented herein can help users quickly finding the information they require.

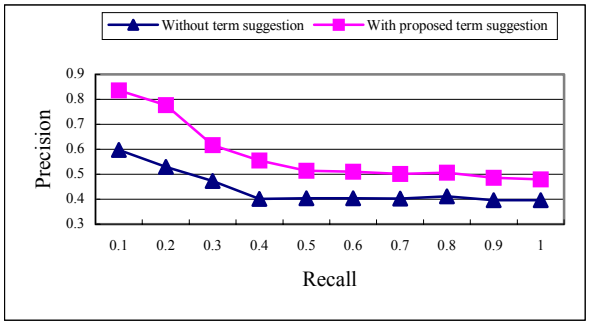

Figure 5. Precision rates corresponding to different recall rates when the proposed term suggestion was applied or not.

5.3 Compared with OPENFIND

In this subsection, the suggested terms of the proposed approach are compared with those of OPENFIND [23], a well-known Chinese search engines that support term suggestion. These suggestion results are listed in Table 2. As mentioned in Section 4, the characteristics of the proposed term suggestion are user-oriented, content independent, adaptability, extensibility, and good quality. Here, some characteristics can be further discussed and verified.

Table 2. A comparison of suggested terms in the proposed approach and OPENFIND

Original query term	Suggested terms of the proposed approach	Suggested terms of OPENFIND
Fossil	**Mineral**	Liquefied Petroleum Gas, High Pressure Gas, Sandstone, Store, Shale, Dinosaur Fossil, Living Fossil, Fossil Fuel
CDROM	GISH, AACOM, Acer, FUJITSU, SCSI, ELOHA, CD Recorder, Acer Tech., MO, ACER	Soundtrack, Compact Disc, Unreadable Disc, Extract Soundtrack, Virtual CDROM, CDROM, Disk, Hard Disk, Floppy Disk
LCD	LCD	Flash Lamp, Camera Shutter, Automatic focus, Diaphragm, Memory Card
A-Kuei	Download City of Nokia, Mashi Maro, YAM comic, Ringing of Cellular Phone, FLASH, Nokia, Gan Giau Long	Web Site of A-Kuei
911	**911 Event**, **America 911**, Porsche, America World Trade Center, Pentagon, ETTODAY, Car Picture, Bin Laden, Sport Car	N/A

In response to the problem of partial match, the novel system suggests a helpful term "Mineral" for the original query term "Fossil", which often cause the problem of partial match in Chinese. For example, after entering the query term "Fossil", an unrelated term "Liquefied Petroleum Gas" with problem of partial match in Chinese will be suggested in a famous Chinese term suggestion system, namely OPENFIND. In our experiences, if the goal of searching is to find information on "Fossil" in Chinese, the best query term is "Fossil" with "Mineral". Furthermore, regarding the problem of word mismatch in search engines, the novel system can provide synonyms and related terms to help users reformulate their original query terms. For example, for the query term "CD-ROM", terms suggested by the novel system include the manufacturers and models of CD-ROMs; for the query term "LCD", its English equivalent (LCD) will be suggested.

Generally, a query term may have different meanings. For example, the query term "A-kuei" in Table 2 has two meanings, a kind of flash animation and a kind of ringing of cellular phone. The proposed approach suggests terms including "Flash" and other flash animations like "Gan Giau Long" and "Mashi Maro" for the query term "A-kuei", which is viewed as a kind of flash animation here. Meanwhile,

the proposed approach also suggests terms including "Nokia", "Download City of Nokia" and "Ringing of Cellular phone" for the query term "A-kuei", which is viewed as a kind of ringing of cellular phone here. In this case, users can easily find the desired information related to the term "A-kuei", in different meanings.

Since the proposed term suggestion is based on user navigation behavior extracted from access logs, suggested terms can change rapidly. That is, the proposed approach can identify new associations of terms, and the old ones will be ranked in the bottom of the result list or even filtered out. For example, the access logs used herein are collected from Nov. 18 to Nov. 25, 2001. Regarding the query term "911", the suggested terms highly related to the terrorist attack on Sep. 11, 2001 are ranked in the top of the result list. The terms suggested by the proposed approach are thus clearly adaptive. By contrast, OPENFIND doesn't suggest any candidate terms regarding the original query term "911".

6. Conclusions

This paper proposed a novel term suggestion method for users to quickly find the desired information. A co-clicked behavior based term suggestion is presented to suggest search terms. Analyzing the co-clicked behaviors of users in the access logs for term suggestion eliminates the need to perform textual analysis. The proposed approach thus provides some positive characteristics that previous approaches neglected, including content independent, adaptability, and extensibility. In addition, limitations of current search engines such as problems of word mismatch and partial match can also be overcome. Here, a search engine prototype is also developed to show the results of term suggestion. Experimental results demonstrate that the precision rates of information retrieval systems can be improved by using the suggested terms of the proposal presented herein. The proposed approach also performs better than another famous Chinese term suggestion system, OPENFIND.

References

[1] S. Brin, L. Page, "The Anatomy of a Large-Scale Hypertextual Web Search Engine," *Computer Networks and ISDN Systems*, Vol. 30, pp. 107-117, 1998.

[2] S. Brin, L. Page, "The PageRank Citation Ranking: Bringing Order to the Web," Technical Report, 1998.

[3] S. Chakrabarti, B. E. Dom, D. Gibson, J. Kleinberg, P. Raghavan, S. Rajagopalan, "Automatic Resource Compilation by Analysing Hyperlink Structure and Associated Text," *Proceedings of the 7th World Wide Web Conference, Brisbane,* 1998.

[4] S. Chakrabarti, B. E. Dom, S. R. Kumar, P. Raghavan, S. Rajagopalan, A. Tomkins, D. Gibson, J. Kleinberg, "Mining the Web's Link Structure," *IEEE Computer*, Vol. 32(8), pp. 60-67, August 1999.

[5] C. H. Chang, C. C. Hsu, "Enabling Concept-Based Relevance Feedback for Information Retrieval on the WWW," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 11(4), pp. 595-609, July/August 1999.

[6] B. Chidlovskii, N. S. Glance, M. Antonietta Grasso, "Collaborative Re-Ranking of Search Results," *AAAI-2000 Workshop on Artificial Intelligence for Web Search*, pp. 18-22, July 2000.

[7] L. F. Chien H. T. Pu, "Important Issues on Chinese Information Retrieval," *Computational Linguistics and Chinese Language Processing*, Vol. 1(1), pp. 205-221, 1996.

[8] W. B. Frakes, R. Baeza-Yates, Information Retrieval: Data Structures and Algorithms, Prentice-Hall, 1992.

[9] F. Guozhen, C. Xueqi, B. Shuo, "SAInSE: An Intelligent Search Engine Based on WWW Structure Analysis," *Proceedings of 15th International Parallel and Distributed Processing Symposium*, pp. 1734-1740, 2001.

[10] L. Huang, "A Survey on Web Information Retrieval Technologies," Working Paper, http://citeseer.nj.nec.com/336617.html.

[11] B. M. Kim, J. Y. Kim, J. Kim, "Query Term Expansion and Reweighting using Term Co-Occurrence Similarity and Fuzzy Inference," *IFSA World Congress and 20th NAFIPS International Conference*, Vol. 2, pp. 715-720, 2001.

[12] W. S. Li, D. Agrawal, "Supporting Web Query Expansion Efficiently using Multi-Granularity Indexing and Query Processing," *Data and Knowledge Engineering*, Vol. 35(3), pp. 239-257, 2000.

[13] H. Lieberman, "Letizia: An Agent That Assists Web Browsing," *International Joint Conference on Artificial Intelligence (IJCAI-95)*, pp. 924-929, August 1995.

[14] T. X. Lin, HitRank: Search Ranking Based on Mining User-oriented Feedback, Master Thesis, National Taiwan University of Science and Technology, 2002.

[15] M. Mitra, A. Singhal, C. Buckley, "Improving Automatic Query Expansion," *Proceedings of 21th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 206-214, August 1998.

[16] C. Silverstein, M. Henzinger, H. Marais, M. Moricz, Analysis of a Very Large AltaVista Query Log, SRC Technical Note 1998-014.

[17] J. Wei, S. Bressan, B. C. Ooi, "Mining Term Association Rules for Automatic Global Query Expansion: Methodology and Preliminary Results," *Proceedings of 1st International Conference on Web Information Systems Engineering*, Vol. 1, pp. 366-373, 2000.

[18] J. R. Wen, J. Y. Nie, H. J. Zhang, "Query Clustering Using User Logs," *ACM Transactions on Information Systems*, Vol. 20(1), pp. 59-81, January 2002.

[19] B. Xu, W. Zhang, H. Yang, W. C. Chu, "A Rough Set Based Self-Adaptive Web Search Engine," *Computer Software and Applications Conference (COMPSAC 2001)*, pp. 377-382, 2001.

[20] J. Xu, W. B. Croft, "Query Expansion using Local and Global Document Analysis," *Proceedings of 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 4-11, 1996.

[21] C. C. Yang, J. Yen, H. C. Chen, "Intelligent Internet Searching Agent Based on Hybrid Simulated Annealing," *Decision Support Systems*, Vol. 28, pp. 269-277, 2000.

[22] URL: http://www.google.com/, Google.

[23] URL: http://www.OPENFIND.com.tw/, OPENFIND.

[24] URL: http://www.teoma.com/, Teoma.

[25] URL: http://www.yam.com/, YAM.