



A comparative analysis of the search feature effectiveness of the major English and Chinese search engines

Search feature effectiveness

217

Jin Zhang

School of Information Studies, University of Wisconsin Milwaukee, Milwaukee, Wisconsin, USA

Wei Fei

Suzhou Library, Suzhou, China, and

Taowen Le

Goddard School of Business and Economics, Weber State University, Ogden, Utah, USA

Refereed article received
6 July 2011
Approved for publication
5 May 2012

Abstract

Purpose – The purpose of this paper to investigate the effectiveness of selected search features in the major English and Chinese search engines and compare the search engines' retrieval effectiveness.

Design/approach/methodology – The search engines Google, Google China, and Baidu were selected for this study. Common search features such as title search, basic search, exact phrase search, PDF search, and URL search, were identified and used. Search results from using the five features in the search engines were collected and compared. One-way ANOVA and regression analysis were used to compare the retrieval effectiveness of the search engines.

Findings – It was found that Google achieved the best retrieval performance with all five search features among the three search engines. Moreover Google achieved the best webpage ranking performance.

Practical implications – The findings of this study improve the understanding of English and Chinese search engines and the differences between them in terms of search features, and can be used to assist users in choosing appropriate and effective search strategies when they search for information on the internet.

Originality/value – The original contributions of this paper are that the Chinese and English search engines in both languages are compared for retrieval effectiveness. Five search features were evaluated, compared, and analysed in the two different language environments by using the discounted cumulative gain method.

Keywords Search engines, Search feature, Retrieval effectiveness, Evaluation, Google, Google China, Baidu, Information retrieval

Paper type Research paper

Introduction

As internet information becomes increasingly important to people, effective information retrieval is receiving more attention from researchers worldwide than ever before. With continuing advancements in networking technology, more abundant information is now available to users worldwide on global networks. While locating the needed information from a seemingly unlimited knowledge reservoir would otherwise be impossible for average users, search engines help them reach the needed information through only a few mouse clicks.

A search engine should provide users with various search features to meet their diverse information needs. A special search feature is different from a basic search



Online Information Review
Vol. 37 No. 2, 2013
pp. 217-230

© Emerald Group Publishing Limited
1468-4527
DOI 10.1108/OIR-07-2011-0099

feature in a search engine. Although basic searches are the primary search method for most users, special search features can satisfy special information needs. Basic search features in conjunction with a variety of special search features in a search engine enable users to have an array of search choices. Sometimes it is the special search features in a search engine that differentiate it from other search engines.

The objectives of this study are to identify a major English search engine and the two main Chinese search engines, compare common search features in the search engines, investigate retrieval effectiveness, and provide insight into these search engines. Since Google maintains two search engine versions in China (one for Chinese query searches and one for English query searches), both versions were included in the study. The research question of this study is whether there are significant differences between the major English search engine and the equivalent Chinese search engines in terms of retrieval effectiveness.

Related research

Since their inception researchers and industry professionals alike have been studying and improving search engines. Various studies have been conducted to analyse search engines' search features. For example, one study selected two advanced search features (exact title search and URL search) to examine the impact of non-topical terms and semi-topical terms on query expansion, and found that search results would improve if queries are restricted to the exact title search or URL search (Fattahi *et al.*, 2008). Contrary to expectation in another study Boolean search strategies were ineffective (Ford *et al.*, 2002). Still another study indicated that most query operators such as AND, OR, MUST APPEAR (C), or PHRASE (" ") were not used by the majority of searchers; the researchers examined the effects of query operators on the performance of three major web search engines, and found that the use of most query operators had no significant effect on coverage, relative precision, or ranking (Eastman and Jansen, 2003). Other studies also focused on operators in search engines (Höchstötter and Koch, 2009; Jansen and Spink, 2006). In an automatic process, search queries from three different search engines were compared over a long period, demonstrating that some patterns were stable in a number of different search engines (Höchstötter and Koch, 2009). Jansen *et al.* (2000) analysed transaction logs of the search engine Excite, and investigated changes in queries during a session, the number of search terms, and the use of logic and modifiers, and search term rank/frequency distribution.

As a result of such efforts by academic and industry researchers, numerous search engines have been designed and enhanced to assist users in their search for information, including web search engines, selection-based search engines, meta-search engines, desktop search tools, and web portals and vertical market web sites that have a search facility for online databases (Wikipedia, 2011). While many countries have developed their own search engines to work with online databases or webpages in their own languages such as Baidu for Chinese webpages, certain search engines such as Google have become popular across the globe.

Google is well known due to its dominant market share and its powerful and sophisticated search features. According to Alexa.com (2009) Google surpasses other search engines in terms of traffic. Google is literally dominating the field of web information retrieval with a 56.08 per cent search-engine market share, followed by Bing (10.63 per cent), Yahoo Search (9.96 per cent), Ask (3.31 per cent), and AOL Search (1.80 per cent) (Hitwise, 2012). It is estimated that 81 per cent of global internet users

visit the Google web site (Jarboe, 2007). It is likely that more recent data on Google use may show it to have climbed even higher.

As a search engine Google is rich in search features. In addition to simple keyword search Google provides users with an array of other search options, including title search, exact phrase search, Boolean search, language restriction, file type restriction, country restriction, internet domain restriction, URL search, text search, anchor search, date restriction, similar pages, linked pages, returned result number restriction, numeric range, and safe search filter, etc. These search features or options are designed to assist searchers in meeting their various special information needs and achieving better retrieval performance.

To reach the thriving Chinese search engine market, Google Inc developed a Chinese version of its search engine – Google China – and successfully launched it in China. Other than a different language framework or platform and a different set of character processing mechanisms, Google China had identical functionality and interface to the English version. Soon Google China and the domestically developed Chinese search engine Baidu became the two dominant search engines in the Chinese market with a combined 90 per cent market share. Although Baidu was the most used search engine in the Chinese market and enjoyed a 56.8 per cent share of searches, Google China was second with 32.8 per cent of the search share (Baker, 2006).

Google Inc officially gave Google China a Chinese name – Guge – in April 2006, which means the Song of Sowing and Expectation and the Song of Harvest and Joy. In April 2010, due to some unresolved disputes with China's regulatory agency, Google Inc officially terminated the use of the name Guge, moved its operations from mainland China to Hong Kong, and resumed calling the Chinese version Google China (BaiKe, 2011).

Today, while Google is facing unprecedented challenges in China in comparison to its operations in other countries, Baidu seems to be quite popular among Chinese users. According to the China Internet Network Information Centre (CNNIC) 76.9 per cent of users in China choose Baidu as their preferred search engine, while 16.6 per cent choose Google. While Baidu has 76 per cent of the market share, Google is in second place with 18.9 per cent of the market (CNNIC, 2010).

During the past decade much research effort has been devoted to comparative studies and analyses of search engines in different regions. One study concentrating on two major search engines, FAST and Excite, found differences in search behaviour and search topics between US and European users. While the Americans focus more on e-commerce information, the Europeans use the web more for social information (Spink *et al.*, 2002). Other comparative studies have explored search engines in other countries including Thailand, China, Germany, and Greece (Tongchim *et al.*, 2006; Long *et al.*, 2007; Griesbaum, 2004). In one study Google and other Chinese search engines were compared and analysed in terms of database coverage, retrieval functions, and interfaces (Chen and Qing, 2003). Resources and services of digital library projects in both Google and Baidu were examined (Yan, 2008). In another study three popular Chinese search engines were examined and evaluated from the users' angle (Long *et al.*, 2007). However studies evaluating and comparing retrieval effectiveness of multi-language search engines from the search features perspective are scarce in the existing literature.

Research methodology

Selection of search engines

Google (www.google.com) has been a prestigious and influential search engine in the search engine market in terms of reputation, database coverage, and search features,

and Baidu (www.baidu.com) is the most prominent and popular search engine in China. For these reasons Google and Baidu were selected for this comparative study. As already mentioned, Google has a Chinese version: Google China (www.google.cn). With a Chinese interface Google China is designed to specifically search web sites in Chinese. Users can submit queries in Chinese and the search engine will return webpages in Chinese. This language interface makes Google China a major competitor of Baidu.

This study compares Google, Google China, and Baidu. Google China was included because it served as an effective reference when the retrieval effectiveness of Google was compared with that of Baidu. The retrieval performance of Google was based on queries and webpages in English while that of Baidu was based on queries and webpages in Chinese. When a particular search feature of one search engine outperforms that of another, the performance difference could result from multiple factors such as the search engine retrieval algorithm utilised and the quality of webpages in the respective languages. Referencing the performance of the same search feature in Google China, people would know whether the quality of webpages in the respective languages played a role in the performance difference.

Selection of search features

Five search features – title search, basic search, exact phrase search, Portable Document Format (PDF) search, and URL search – from the above-mentioned three search engines were identified and selected for this study.

Title search requires that keywords in a query should appear in the title of a returned webpage, as embedded in the <title> tag. For basic search, keywords in a query can appear in any part of a returned webpage, such as the title, subtitle, introduction, or conclusion. Exact phrase search requires that phrases specified in a query are exactly matched with phrases in the retrieved webpages. For instance the phrase “school library” in a query can only match “school library” rather than “library school” in a webpage. Webpages on the internet are rich not only in contents but also in formats. PDF is one of the most popular webpage formats. PDF search is particularly important for academic information needs. PDF search limits the retrieved webpage file format to the PDF format. URL search dictates that keywords in a query appear only in the URL of the returned webpage. In other words query terms should show in the URL address of a retrieved webpage in URL search.

Basic search is the default search feature in all three search engines. It is also the most frequently used search feature. Therefore it can serve as a comparison baseline. Title search and PDF search are popular search features. Exact phrase search and URL search are utilised to meet special user requirements. The availability of these five search features in all three engines (Google, Google China, and Baidu) makes the comparative analysis of these three search engines possible.

The retrieval performances of these search features were analysed and retrieval results compared across the three search engines.

Design of search tasks

Search tasks were designed and translated into corresponding queries. These queries were submitted to the selected search engines. The internet information resources are rich and diverse. To ensure that the formulated queries were representative enough, the designed tasks covered four different domains: medicine and health, culture and education, information and technology, and business and economy. Theoretically

speaking, the more search tasks in a study, the more sound the study's results. Some studies have used more than 50 search tasks (Lewandowski, 2008), while other studies have used fewer (Vaughan, 2004). To obtain reasonable experimental results for each of the search features, 20 different search tasks were generated for the study, with five search tasks focusing on each of the four subject domains. Since the comparative study of the selected search engines was conducted in two different language environments, the 20 search tasks were designed in both Chinese and English. In other words the 20 search tasks in English were also translated into Chinese (see Table I). As a result 20 search queries in English were submitted to Google, 20 search queries in Chinese were submitted to Baidu, and 20 search queries in Chinese were also submitted to Google China. For each search task in each subject domain, five queries were formulated and submitted to the search engines using the above-mentioned five search features. Thus a total of 200 queries were utilised: 100 queries in English for Google, and 100 Chinese queries for Baidu and Google China. Among the 200 queries 20 English queries and 20 Chinese queries were for title search, 20 English queries and 20 Chinese queries for basic search, 20 English queries and 20 Chinese queries for exact phrase search, 20 English queries and 20 Chinese queries for PDF search, and 20 English queries and 20 Chinese queries for URL search.

Data collection and result evaluation

After a query was submitted to a search engine using a particular search feature, a ranked results list was returned. According to previous studies on users' search behaviour, users usually browse only the top 10 or 20 retrieval results and ignore the rest of the retrieved webpages on the list (Jacsó, 2008; Jansen and Spink, 2006). Another study pointed out that users tended to read the results from top to bottom, especially the first- and second-ranked results (Joachims *et al.*, 2005). These studies show that users are more interested in the results which are ranked high on a returned results list from a search engine. With those findings in mind the researchers in this study only examined the top 100 returned webpages from each result list. Sponsored webpages were excluded in order to achieve fair comparisons. Dead links on the returned list were excluded because the dead link rate was very low, and without the full text of a webpage, it is difficult to conduct a subjective analysis of that webpage.

In this study four people participated in the relevance judgement for the retrieved webpages collected from the search engines. After reading and understanding a task thoroughly, each of the judges made a relevance judgement on each of the retrieved results in each list, and assigned a relevance score.

Medicine and health	Culture and education	Information and technology	Economy and business
Obesity and heart disease	Cultural diversity	Information retrieval technology	Financial crisis
Diet and diabetes	Distance education	Information management	Economic stimulus
Sports and health	Information literacy	Information security	Industrial pollution
Mental health	World cultural heritage	Intellectual property	Food security
Keep fit	Folk customs	Electronic commerce	International trade and marketing

Table I.
Summary of 20 search queries

Relevance was measured and categorised using a predefined five category scoring system. The five levels of relevance were: not relevant, less relevant, relevant, very relevant, and most relevant. In the scoring system 0 indicates not relevant, 0.25 less relevant, 0.5 relevant, 0.75 very relevant, and 1 most relevant. After the title, the abstract (if included), and full text of a particular webpage were scrutinised and the subject matter of the page compared with the corresponding search task, a relevance score was assigned to the webpage. In other words the relevance judgement for a particular webpage was not based on its length, file type, format, layout, and the angle from which it addressed the subject. The relevance judgement for a webpage depended fully on its subject content.

The effectiveness of a search was calculated based on the raw relevance scores of the items and their positions from the corresponding query. Discounted cumulative gain is defined as the sum of products of a relevance score and its position weight on a returned results list (Järvelin and Kekäläinen, 2000, 2002). DCG is considered the amount of weighted relevance information on a returned results list. The retrieval effectiveness is defined as follows:

$$E(w_i) = \sum_{i=1}^n \frac{R_i(w_i)}{\text{Log}_2(i+1)} \quad (1)$$

In Equation (1) w_i is an examined webpage, $E(w_i)$ is the retrieval effectiveness for a search, $R_i(w_i)$ is a relevance score of a given webpage (w_i) on a retrieval results list, and n equals the length of the results list of the search. Although most users browse the top 10-20 items located on the top of a returned results list, it does not mean that all users only read the top 10-20 items. Some users may browse more items on a retrieval results list. In order to achieve a sound and reliable result, the number of examined items on a results list was defined as 100. In this study, because the top 100 webpages on a search results list were selected for examination and evaluation, the parameter n was 100 for all the queries.

Data analysis

Since the researchers attempted to compare the three search engines regarding the performance of each selected search feature, the resulting retrieval effectiveness data of each search feature were classified into three groups – Google, Google China, and Baidu – and a one-way ANOVA analysis was applied to the three groups to test whether there were significant differences among the three search engines. Therefore five ANOVA analyses were conducted for the five search features, respectively.

If significant differences exist among the three search engines in terms of retrieval effectiveness, then the Tukey HSD analysis method is employed to examine the causes of differences among the groups. The data analysis tool employed was the SPSS statistical analysis software package (Version 17). The significance level (p) or *sig*, which represents the percentage or the probability that the results are due to chance, is set to 0.05 for all the tests in this study. In addition a regression analysis was conducted to examine the ranking quality of a retrieval results list for a particular query. In a well-ranked results list, the relevance score of a returned webpage in a higher position should be higher than that of a returned webpage in a lower position on a retrieval results list. A regression analysis enables people to understand the downward trend of the returned webpages in terms of the relevance to a query. Based on the statistical

analysis results, conclusions on search engine performance among the three search engines were made.

Results

Effectiveness analysis of five search features among the three search engines

For each of the five selected search features, 20 queries were submitted to each search engine. Consequently a total of 100 ($20 \times 5 = 100$) queries were submitted to each search engine. Therefore a total of 300 queries were submitted to the three search engines studied. Since the top 100 webpages from each query were examined and investigated, a total of 30,000 ($100 \times 300 = 30,000$) webpages were processed in this study.

The results are summarised and displayed in Tables II and III. In Table II the means for title search, basic search, exact phrase search, PDF search, and URL search for Google are 14.2827, 13.0910, 11.4940, 14.8360, and 11.2202, respectively. It is evident that the means for all five search features of Google are higher than those of both Google China and Baidu.

With *sig* values of 0.000, 0.000, 0.001, 0.000, and 0.000, respectively, and with all of them lower than 0.05, the ANOVA tests show that there are significant differences among the three search engines in search effectiveness with each of the five search features. The degrees of freedom (df) are 2 between groups, and 57 within groups. *F*-values for the five search features – title search, basic search, exact phrase

	<i>n</i>	Mean	SD	SE	95% confidence interval for mean		Minimum	Maximum
					Lower bound	Upper bound		
<i>Title search</i>								
Baidu	20	12.3220	0.66927	0.14965	12.0088	12.6352	11.17	13.37
Google China	20	13.4528	0.55345	0.12376	13.1938	13.7118	12.29	14.32
Google	20	14.2827	0.65240	0.14588	13.9774	14.5881	13.10	15.44
Total	60	13.3525	1.01817	0.13144	13.0895	13.6155	11.17	15.44
<i>Regular search</i>								
Baidu	20	11.0997	0.78325	0.17514	10.7331	11.4662	9.75	12.08
Google China	20	12.1800	0.61575	0.13769	11.8918	12.4682	10.60	12.87
Google	20	13.0910	0.65924	0.14741	12.7825	13.3995	11.99	14.07
Total	60	12.1236	1.06458	0.13744	11.8485	12.3986	9.75	14.07
<i>Phrase search</i>								
Baidu	20	10.5912	0.95441	0.21341	10.1445	11.0379	9.17	11.65
Google China	20	10.4027	0.84900	0.18984	10.0054	10.8001	7.63	11.35
Google	20	11.4940	0.96606	0.21602	11.0419	11.9462	9.12	13.18
Total	60	10.8293	1.02796	0.13271	10.5638	11.0949	7.63	13.18
<i>PDF search</i>								
Baidu	20	12.5848	0.99142	0.22169	12.1208	13.0488	10.97	14.29
Google China	20	14.2526	0.71306	0.15944	13.9189	14.5863	12.47	15.22
Google	20	14.8360	0.65327	0.14608	14.5302	15.1417	14.09	16.22
Total	60	13.8911	1.24221	0.16037	13.5702	14.2120	10.97	16.22
<i>URL search</i>								
Baidu	20	8.2979	0.92294	0.20638	7.8659	8.7298	6.24	9.81
Google China	20	9.9053	1.41500	0.31640	9.2431	10.5676	7.41	12.26
Google	20	11.2202	1.47337	0.32945	10.5307	11.9098	8.20	13.34
Total	60	9.8078	1.75228	0.22622	9.3552	10.2605	6.24	13.34

Table II.
Descriptive summary of the three search engines

Table III.
ANOVA test results
for the three search
engines

	Sum of squares	df	Mean square	<i>F</i>	Significance
<i>Title search</i>					
Between groups	38.746	2	19.373	49.259	0.000
Within groups	22.417	57	0.393		
Total	61.163	59			
<i>Regular search</i>					
Between groups	39.749	2	19.875	41.776	0.000
Within groups	27.117	57	0.476		
Total	66.867	59			
<i>Phrase search</i>					
Between groups	13.611	2	6.805	7.960	0.001
Within groups	48.734	57	0.855		
Total	62.345	59			
<i>PDF search</i>					
Between groups	54.598	2	27.299	42.696	0.000
Within groups	36.445	57	0.639		
Total	91.042	59			
<i>URL search</i>					
Between groups	85.686	2	42.843	25.579	0.000
Within groups	95.472	57	1.675		
Total	181.158	59			

search, PDF search, and URL search – are 49.259, 41.776, 7.960, 42.696, and 25.579, respectively.

The Tukey HSD test was subsequently applied to the data set to identify the causes of the differences. The detailed results are shown in Table IV. It is clear from the significant mean differences between Google and Google China or Baidu that Google outperforms Google China and Baidu in all five search categories: title search, basic search, exact phrase search, PDF search, and URL search. As shown by the significant mean differences between Google China and Baidu in title search, basic search, PDF search, and URL search, Google China edges out Baidu in those four categories. In summary, Baidu achieved the worst performance among the three search engines.

Figure 1 gives intuitive and visual displays of the retrieval effectiveness results.

Regression analysis of the five search features among the three search engines

The average relevance scores for all five search features and for all three search engines were calculated, and a regression analysis was performed. A greater R^2 score is associated with a better webpage ranking result on a retrieval results list. The regression equation is a logarithmic model and is defined in Equation (2) for all five search features in this study. In Equation (2) Y stands for the relevance score of a webpage, X stands for its position on a results list returned by a search engine, and b_0 and b_1 are two parameters that affect the shape of the curve. However, each search feature corresponds to its own parameters b_0 and b_1 ; therefore, it is no surprise that parameters b_0 and b_1 vary across different search features (see Table V):

$$Y = b_0 + b_1 \times \ln(X) \quad (2)$$

The regression analyses were applied to basic search, exact phrase search, PDF search, title search, and URL search as well. The resultant values of R^2 , F , sig , b_0 and b_1 are tabulated in Table V. A visual display of the R^2 among the three search engines is

Dependent variable	<i>I</i>	<i>J</i>	Mean difference (<i>I</i> - <i>J</i>)	SE	Significance
Title search	Baidu	Google China	-1.13079*	0.19831	0.000
		Google	-1.96072*	0.19831	0.000
	Google China	Baidu	1.13079*	0.19831	0.000
		Google	-0.82992*	0.19831	0.000
	Google	Baidu	1.96072*	0.19831	0.000
		Google China	0.82992*	0.19831	0.000
Regular search	Baidu	Google China	-1.08036*	0.21812	0.000
		Google	-1.99132*	0.21812	0.000
	Google China	Baidu	1.08036*	0.21812	0.000
		Google	-0.91096*	0.21812	0.000
	Google	Baidu	1.99132*	0.21812	0.000
		Google China	0.91096*	0.21812	0.000
Phrase search	Baidu	Google China	0.18850	0.29240	0.796
		Google	-0.90282*	0.29240	0.009
	Google China	Baidu	-0.18850	0.29240	0.796
		Google	-1.09132*	0.29240	0.001
	Google	Baidu	0.90282*	0.29240	0.009
		Google China	1.09132*	0.29240	0.001
PDF search	Baidu	Google China	-1.66779*	0.25286	0.000
		Google	-2.25117*	0.25286	0.000
	Google China	Baidu	1.66779*	0.25286	0.000
		Google	-0.58338	0.25286	0.063
	Google	Baidu	2.25117*	0.25286	0.000
		Google China	0.58338	0.25286	0.063
URL search	Baidu	Google China	-1.60747*	0.40926	0.001
		Google	-2.92234*	0.40926	0.000
	Google China	Baidu	1.60747*	0.40926	0.001
		Google	-1.31487*	0.40926	0.006
	Google	Baidu	2.92234*	0.40926	0.000
		Google China	1.31487*	0.40926	0.006

Note: *The mean difference is significant at the 0.05 level

Table IV. Multiple comparisons in Tukey HSD for the three search engines

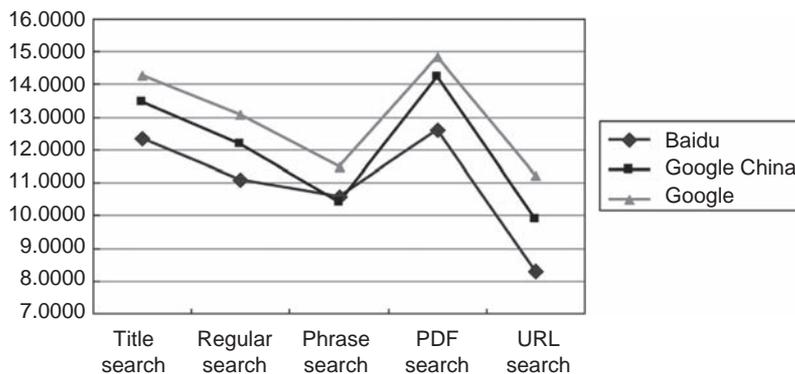


Figure 1. Effectiveness mean plots for the five search features

illustrated in Figure 2. In Figure 2 the X-axis represents various search features, and the Y-axis represents R^2 values.

The regression analysis results in Table V and Figure 2 show that Google once again outperforms both Google China and Baidu in all five search features in terms of

Table V.
Regression analyses
for the five search
features and the three
search engines

Search features	Search engines	R^2	F	Significance	b_0	b_1
Title search	Baidu	0.603	148.913	0.000	0.811	-0.070
	Google China	0.796	382.019	0.000	0.855	-0.075
	Google	0.854	573.233	0.000	1.001	-0.099
Regular search	Baidu	0.449	79.942	0.000	0.712	-0.058
	Google China	0.74	278.582	0.000	0.784	-0.072
	Google	0.897	854.268	0.000	0.971	-0.107
Phrase search	Baidu	0.475	88.505	0.000	0.673	-0.054
	Google China	0.747	288.846	0.000	0.730	-0.071
	Google	0.874	677.898	0.000	0.925	-0.097
PDF search	Baidu	0.662	192.056	0.000	0.871	-0.084
	Google China	0.795	380.242	0.000	0.902	-0.081
	Google	0.815	430.543	0.000	1.088	-0.114
URL search	Baidu	0.207	25.618	0.000	0.487	-0.025
	Google China	0.527	109.111	0.000	0.574	-0.039
	Google	0.728	262.239	0.000	0.750	-0.067

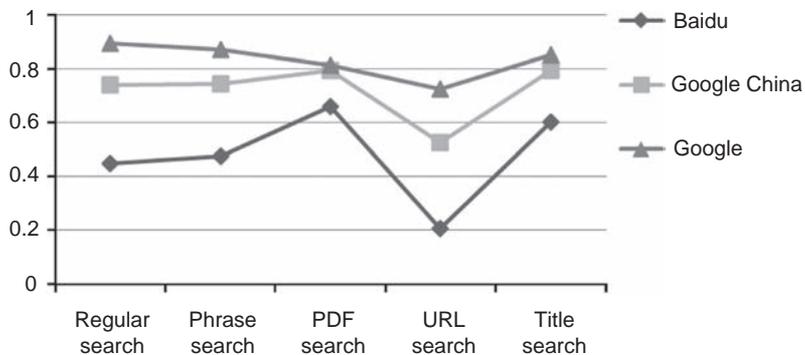


Figure 2.
Comparison of R^2
values among the three
search engines

search results ranking quality. Baidu had the worst performance and Google China achieved second place. Notice that Google's performance in PDF search is almost similar to that of Google China. Also, for all three search engines, performance in URL search is the worst.

Discussion

The findings imply that the quality of English webpages was generally better than the quality of webpages in Chinese. It is interesting that Google outperformed Google China in all five search features in terms of retrieval effectiveness. Google was better than Google China in both the one-way ANOVA analysis and the regression analysis. The mean differences between Google China and Google in title search, basic search, phrase search, PDF search, and URL search are -0.82992 , -0.91096 , -1.09132 , -0.58338 , and -1.31487 , respectively, and their *sigs* are lower than 0.05. Google China and Google, designed and operated by the same company, may adjust the retrieval, indexing, and ranking mechanisms to adapt to different language environments. One of the major differences between Google and Google China is the nature of the language in the webpages. Google primarily processes webpages in English while Google China primarily processes webpages in Chinese. It suggests that the performance differences

between Google and Google China may be caused by the quality differences between English webpages and Chinese webpages. This finding confirms the claim in another study that the quality of webpages affects search engine performance (Vaughan and Thelwall, 2004).

Digital libraries, open access resources, and institutional repositories are not buzzwords. Instead they have been widely created or implemented in libraries, colleges, universities, schools, institutes, and other organisations in western countries, especially in the USA. Unlike average webpages, information resources from these libraries, open access web sites, and institutional repositories are both rich in content and high in quality. They often include scholarly documents, journal papers, proceedings, academic reports, syllabuses, teaching materials, e-books, government documents, and other electronic materials. Such information resources are crawled and indexed by search engines such as Google and have become an extremely important part of search engine databases. However, although internet-related information technology has advanced rapidly in recent years, digital libraries, open access resources, and institutional repositories are still developing. This in part accounts for the lower quality of the webpages in Chinese.

The technology of Web 2.0 has had a profound impact on internet information organisation and utilisation. It takes interaction between the internet and its users to an unprecedented level. Wikipedia, for example, results directly from Web 2.0 technology. English Wikipedia has proven to be a comprehensive, reliable, and rich information resource; however Chinese equivalents to Wikipedia such as Baidu Wikipedia and Baidu Knows are not yet as mature as their English counterpart in terms of content, scope, and quality. This could also account in part for Baidu's poorer performance in this study.

It was known that Google, Google China, and Baidu treat blog information differently when they rank blog webpages in a returned results list. Blog pages generally contain raw information that has not undergone editing or peer-reviewing processes. In other words their quality is typically not controlled. In this study it was found that about 9 per cent of the returned webpages from Baidu were blog-related compared to only about 2.3 per cent of the returned webpages from Google. This may have also accounted in part for the performance differences between Google and Baidu in this study.

Among the five search features, URL search was the least effective for both Google China and Baidu. Although internet protocols support the Chinese language, internet domain names do not support Chinese characters. In other words webpages in Chinese can be displayed and transferred on the internet, but their domain names must be in English. This may explain why URL search achieved the lowest search effectiveness among the five search features both for Google China and for Baidu. It has been reported that the newly revised internet domain names will support Chinese characters (Sayer, 2010). If so the change will benefit more than 420 million Chinese users (CNNIC, 2010) and is likely to have a positive impact on Chinese URL searches for any search engine.

Although only three distinct search engines were explored in this comparative study, some interesting patterns were observed. Within each of the three search engines, PDF search was the most effective, followed in descending order by title search, basic search, exact phrase search, and URL search. If basic search is used as the comparison baseline, two search features performed better than basic search, and two search features were worse than basic search. It suggests that PDF search and title search are two effective search features regardless of the search engine involved, and exact phrase search and URL search are less effective search features for all search

engines. In the regression analysis URL search was found to have the worst performance among the five search features for all three search engines.

A cross-language study involves multiple language environments. It is impossible to obtain exactly the same language environments for a cross-language study. In this cross-language study the different performances between the search engines Google, Google China, and Baidu were investigated. Significant differences between the search engines were found. It is not surprising that the differences may result from many factors such as quality of indexing in the databases, coverage of the databases, quality of webpage contents, interfaces of the search engines, quality of results rankings, retrieval results relevance judgements, and so on. This study shows that quality of results ranking, quality of indexing in the databases, and quality of webpage contents may contribute, in part, to the differences.

Conclusion

In order to meet users' growing information needs, search engines have provided users with many search services such as web search, image search, audio search, and video search. Among these services web search is one of the most frequently used. Many web search features have been developed to fulfil specific user requirements. A prior study has shown that there are differences in effectiveness among search features of English search engines (Zhang and Fei, 2010). In this study five search features (title search, basic search, exact phrase search, PDF search, and URL search) in the main English search engine and two major Chinese search engines were compared in terms of retrieval effectiveness and retrieval result ranking. Google, Google China, and Baidu were selected for the study because of their prominence in the search engine field. The study shows that there are significant differences among the three search engines. Google outperformed Google China in terms of retrieval effectiveness in all five selected search features at the 0.05 significance level. It implies that webpages in Chinese may be of relatively poor quality. Meanwhile Google China surpassed Baidu in terms of retrieval effectiveness in four search features (title search, basic search, PDF search, and URL search) at the 0.05 significance level. The regression analysis results show that Google also exceeded Google China and Baidu in terms of search results ranking quality in all five search features. Google China outperformed Baidu and held the second place in this category. The findings reveal that there are significant differences between the major English and Chinese search engines, and the Google search engines outperformed the Baidu search engine due to factors such as quality of indexing in the databases, coverage of the databases, quality of webpage contents, quality of results rankings, and so on.

Although 40 search tasks (comprising 200 search queries) were employed for this study, it is believed that if more search tasks and more search queries were used, the results would be more convincing and plausible. In this study the relevance judgements for the retrieved webpages were conducted by four people. If more people could participate in the relevance judgements, the relevance judgement results would be more reliable. It is one of the research limitations.

With an understanding of such differences, end-users could select appropriate search features and strategies to improve their search effectiveness, search engine developers could further improve and optimise various search features, and web authors might make greater efforts in designing quality webpages so as to increase the possibility of their webpages' being found and read by the intended readers.

Future research directions include, but are not limited to, comparisons between Chinese search engines and other language search engines in terms of these search

features, comparisons between English search engines and other language search engines in terms of these search features, and comparison of content quality of webpages in Chinese and that of webpages in English. Since both Google and Baidu are comprehensive search engines, they include image search, video search, news search, blog search, etc. The same research method used in this study could be applied to these searches.

References

- Alexa.com (2009), "Google.com – traffic details from Alexa", available at: www.alexa.com/siteinfo/Google.com (accessed 23 February 2011).
- BaiKe (2011), "Google", available at: <http://baike.baidu.com/view/1931.htm#2> (accessed 23 February 2011).
- Baker, L. (2006), "Google and Baidu dominate China search market", available at: www.searchenginejournal.com/google-baidu-dominate-china-search-market/3208/# (accessed 23 February 2011).
- Chen, J. and Qing, X. (2003), "A comparative study of four search engines", *Information Science*, Vol. 21 No. 10, pp. 1084-1087.
- CNNIC (2010), "26th Chinese internet development report", China Internet Network Information Centre, available at: <http://research.cnnic.cn/html/1279173730d2350.html> (accessed 23 February 2011).
- Eastman, C. and Jansen, B.J. (2003), "Coverage, relevance, and ranking: the impact of query operators on web search engine results", *ACM Transactions on Information Systems*, Vol. 21 No. 4, pp. 383-411.
- Fattahi, R., Wilson, C.S. and Cole, F. (2008), "An alternative approach to natural language query expansion in search engines: text analysis of non-topical terms in web documents", *Information Processing & Management*, Vol. 44 No. 4, pp. 1503-1516.
- Ford, N., Miller, D. and Moss, N. (2002), "Web search strategies and retrieval effectiveness: an empirical study", *Journal of Documentation*, Vol. 58 No. 1, pp. 30-48.
- Griesbaum, J. (2004), "Evaluation of three German search engines: Altavista.de, Google.de and Lycos.de", *Information Research*, Vol. 9 No. 4, available at: <http://informationr.net/ir/9-4/paper189.html> (accessed 23 February 2011).
- Hitwise (2012), "Search engine trends", available at: www.experian.com/hitwise/online-trends-search-engine.html (accessed 25 July 2012).
- Höchstötter, N. and Koch, M. (2009), "Standard parameters for searching behaviour in search engines and their empirical evaluation", *Journal of Information Science*, Vol. 35 No. 1, pp. 45-65.
- Jacsó, P. (2008), "How many web-wide search engines do we need?", *Online Information Review*, Vol. 32 No. 6, pp. 860-865.
- Jansen, B.J. and Spink, A. (2006), "How are we searching the World Wide Web? A comparison of nine search engine transaction logs", *Information Processing & Management*, Vol. 42 No. 1, pp. 248-263.
- Jansen, B.J., Spink, A. and Saracevic, T. (2000), "Real life, real users, and real needs: a study and analysis of user queries on the web", *Information Processing and Management*, Vol. 36 No. 2, pp. 207-227.
- Jarboe, G. (2007), "Stats show Google dominates the international search landscape", available at: <http://searchenginewatch.com/3625072> (accessed 23 February 2011).
- Järvelin, K. and Kekäläinen, J. (2000), "IR evaluation methods for retrieving highly relevant documents", *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, ACM Press, New York, NY, pp. 41-48.

-
- Järvelin, K. and Kekäläinen, J. (2002), "Cumulated gain-based evaluation of IR techniques", *ACM Transactions on Information Systems*, Vol. 20 No. 4, pp. 422-446.
- Joachims, T., Granka, L., Pan, B., Hembrooke, H. and Gay, G. (2005), "Accurately interpreting clickthrough data as implicit feedback", paper presented at the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, available at: www.cs.cornell.edu/people/tj/publications/joachims_et_al_05a.pdf (accessed 25 July 2012).
- Lewandowski, D. (2008), "The retrieval effectiveness of web search engines: considering results descriptions", *Journal of Documentation*, Vol. 64 No. 6, pp. 915-937.
- Long, H.Q., Lv, B.F., Zhao, T.B. and Liu, Y. (2007), "Evaluate and compare Chinese internet search engines based on users' experience", *2007 International Conference on Wireless Communications, Networking and Mobile Computing, IEEE Press, Los Alamitos, CA*, pp. 6134-6137.
- Sayer, P. (2010), "Chinese language top-level domains win ICANN approval", available at: www.networkworld.com/news/2010/062510-chinese-language-top-level-domains-win.html (accessed 23 February 2011).
- Spink, A., Ozmutlu, S., Ozmutlu, H.C. and Jansen, B.J. (2002), "US versus European web searching trends", *SIGIR Forum*, Vol. 36 No. 2, pp. 30-37.
- Tongchim, S., Somlertlamvanich, V. and Isahara, H. (2006), "Measuring the effectiveness of public search engines on Thai queries", *Proceedings of the Fifth IASTED International Conference on Communications, Internet, and Information Technology, ACTA Press, St Thomas, US Virgin Islands*, pp. 349-354.
- Vaughan, L. (2004), "New measurements for search engine evaluation proposed and tested", *Information Processing and Management*, Vol. 40 No. 4, pp. 677-691.
- Vaughan, L. and Thelwall, M. (2004), "Search engine coverage bias: evidence and possible causes", *Information Processing and Management*, Vol. 40 No. 4, pp. 693-707.
- Wikipedia (2011), "List of search engines", available at: http://en.wikipedia.org/wiki/List_of_search_engines (accessed 23 February 2011).
- Yan, L. (2008), "Google and Baidu digital library project and university library's countermeasures", *Journal of Library and Information Sciences in Agriculture*, Vol. 20 No. 7, pp. 61-63.
- Zhang, J. and Fei, W. (2010), "Search engines' responses to several search feature selections", *International Information and Library Review*, Vol. 42 No. 3, pp. 212-225.

About the authors

Jin Zhang is a Professor at the School of Information Studies, University of Wisconsin-Milwaukee, USA. He has published more than 40 research papers in international journals such as *JASIST*, *IPM*, *Journal of Intelligent Information Systems*, and *IEEE Transactions on Visualization and Computer Graphics*. His book *Visualization for Information Retrieval* was published by Springer in 2008. His research interests include visualization for information retrieval, information retrieval theory and algorithm, internet information organization, search engine evaluation, consumer health informatics, transaction log analysis, digital libraries, and computer human interface design. Jin Zhang is the corresponding author and can be contacted at: jzhang@uwm.edu

Wei Fei is a librarian at the Suzhou Library, Suzhou, Jiangsu, P. R. China.

Taowen Le is a Professor of Information Systems and Technology at the Goddard School of Business and Economics, Weber State University, Ogden, Utah.