International Journal of Electronic Commerce Studies
Vol.4, No.1, pp. 47-62, 2013
doi: 10.7903/ijecs.1111

INTERFACING GOOGLE SEARCH ENGINE TO CAPTURE USER WEB SEARCH BEHAVIOR

Fadhilah Mat Yamin
Universiti Utara Malaysia
School of Technology Management & Logistics, UUM COB, 06010
Sintok, Kedah, Malaysia
fmy@uum.edu.my

T. Ramayah
Universiti Sains Malaysia
School of Management, 11800 Minden, Penang, Malaysia.
ramayah@usm.my

ABSTRACT

The behaviour of the searcher when using the search engine especially during the query formulation is crucial. Search engines capture users' activities in the search log, which is stored at the search engine server. Due to the difficulty of obtaining this search log, this paper proposed and develops an interface framework to interface a Google search engine. This interface will capture users' queries before redirect them to Google. The analysis of the search log will show that users are utilizing different types of queries. These queries are then classified as breadth and depth search query.

Keywords: Search Engine, Search Interface, Search Log Transaction

1. INTRODUCTION

To date, millions repositories, websites and directories that contain billion of electronic documents and web pages have been made available for public access. WorldWideWebSize.com (http://www.worldwidewebsize.com/) estimates that at least 20.26 billion web pages (until Thursday, 18 March, 2010) have been indexed by three major search engines such as Google, Bing, Yahoo Search and Ask[1].

Information on the Web is organized and indexed by an Internet based software agent called a web crawler[2]. The crawler explores the web through the uniform resource locator (URL), retrieves the content of the website, categorizes the information based on the definition created by the webmaster, and then indexes the website in the database[3]. The user-search interface is used to communicate the user and the database. The common user-search interfaces that are available on the Web are online directories and search engines[4, 5, 6]. A search engine is a computer program that retrieves information based on the queries entered by the users. The search engine is the most popular search tool for information searching[5].

Searchers' activities and their interaction with the Web search engine are usually recorded in the search engine server log. The log contains searchers' details such as Internet protocol (IP) number, session number (ID) and information searched, which includes the search streams, terms, operators and other information. A search log provides valuable information to the researchers who study the searchers' searching behavior, search patterns, usage mining and other items. However obtaining the log is expensive as it was not intended for public access. Therefore, an alternative method is proposed to record the usage and create a researcher's own search log. In this paper, a framework of an interface system for interfacing Google search engine is proposed to capture user search activities. A method for analyzing user search behaviour from the search log is also discussed.

2. GOOGLE SEARCH ENGINE

Google is the general purpose search engine and one of the widely used search tools on the Internet[6, 7, 8]. According to Nazim[6], Google popularity is due to the number of reasons such as wide coverage and updated regularly, fast in access, provide user friendly interface, provide links to other websites and a separate interface for searching journals, images, news, and audio.

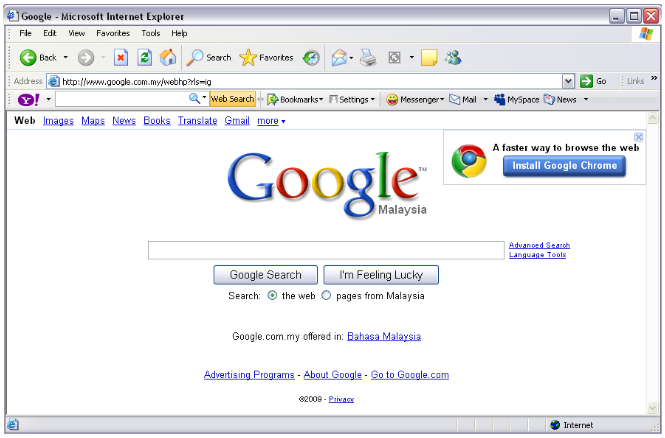

Google - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites
Address http://www.google.com.my/webhp?rls=ig Go Links
Web Search Bookmarks Settings Messenger Mail MySpace News
Web Images Maps News Books Translate Gmail more
A faster way to browse the web
Install Google Chrome
Google Malaysia
Advanced Search
Language Tools
Google Search
I'm Feeling Lucky
Search: ● the web ○ pages from Malaysia
Google.com.my offered in: Bahasa Malaysia
Advertising Programs - About Google - Go to Google.com
©2009 - Privacy
Internet

Figure 1. Google search engine

3. FRAMEWORK OF INTERFACING GOOGLE SEARCH ENGINE

In order to study the searchers' search behaviour, the searching activities that represent the searching behaviour needs to be recorded in the search log. Typically, search behavior is influenced by the user knowledge[9]. Fadhilah and Ramayah[10] have shown that the user's knowledge is significant to the search satisfaction. According to Jansen[11], the search log involves three major stages:

- Collection: the process of collecting the interaction data for a given period in a transaction log.
- Preparation: the process of cleaning and preparing the transaction log data for analysis
- Analysis: the process of analyzing the prepared data

To record the search activities, an interface has been designed and developed. The interface is a layer between the searcher and the Google. As defined by Wang et al.[12] interface is a layer between the user and the search system that facilitates human computer communication.

Figure 2 shows a model of the proposed search interface. A search interface consists of search interface engine and reporting module. A query entered by the searcher will be stored into a database and forwarded to Google. The query will not be modified. It will be forwarded as it is.

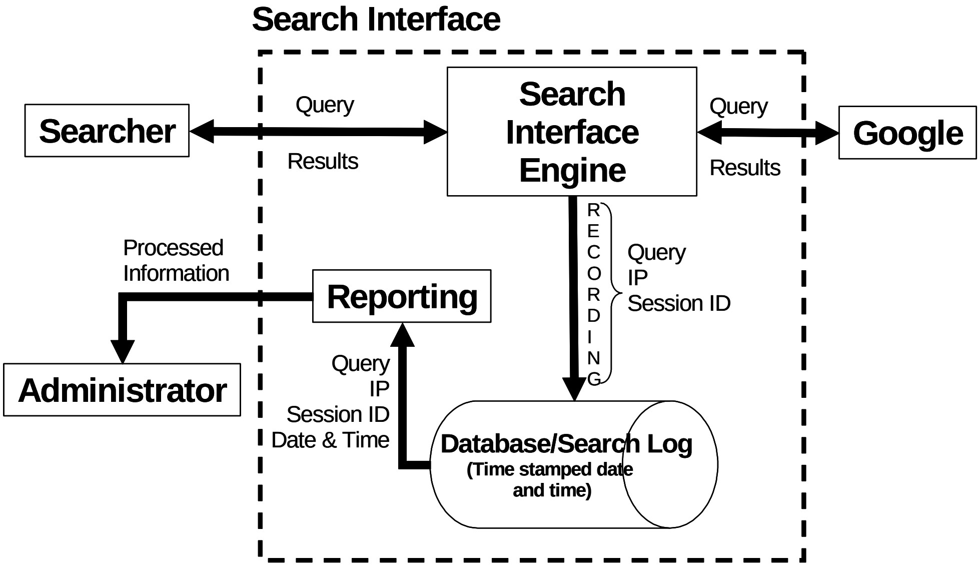
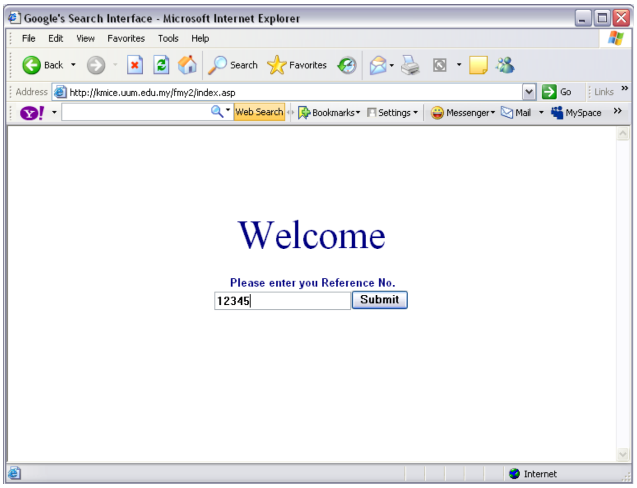


Figure 2. A model of search interface

The search interface consists of two parts, namely a reference number page (Figure 3) and the searching interface (Figure 4). The reference number page is an interface that accepts the user’s reference number. In this study, the student’s matric number was used as the unique reference number to group and index the user’s queries information.


Google's Search Interface - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites
Address http://kmice.uum.edu.my/fmy2/index.asp
Go Links
Web Search Bookmarks Settings Messenger Mail MySpace
Submit
Internet

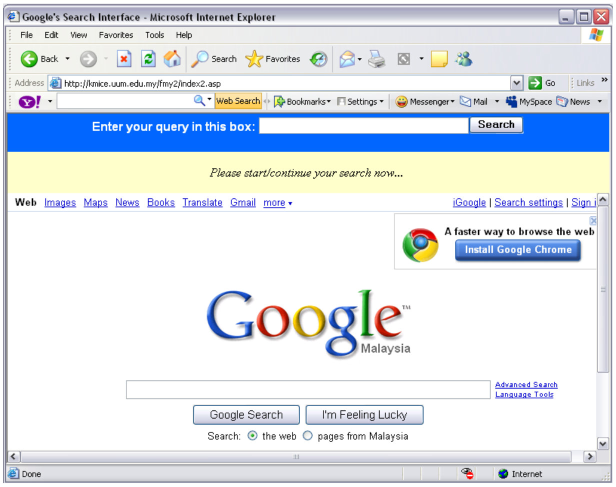
Welcome

Please enter you Reference No.

12345

Figure 3. Search interface - reference number page

The searching interface (Figure 4) will receive the user’s query and forward it to Google for processing and displaying results. This interface does not modify the query or delay the search process as it only records the query before redirecting the query to the Google search engine. This interface consists of two main parts. The upper part, with the blue background, is a section where students can enter their queries. The lower part is where the Google interface and results are displayed. When the students enter queries in the blue area, the queries will be time stamped and stored in the database. The query is submitted to Google, which then returns a list of search results.

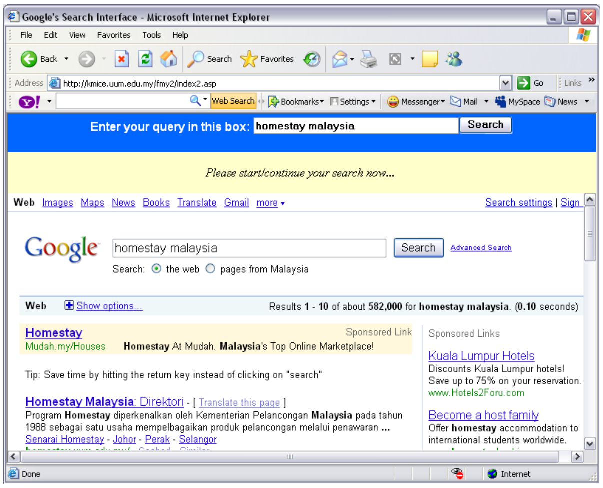

Google's Search Interface - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites
Address http://kmice.uum.edu.my/fmy2/index2.asp
Go Links
Web Search Bookmarks Settings Messenger Mail MySpace News
Search
Web Images Maps News Books Translate Gmail more
iGoogle | Search settings | Sign i
A faster way to browse the web
Install Google Chrome
Google Malaysia
Advanced Search
Language Tools
Google Search
I'm Feeling Lucky
Done
Internet

Enter your query in this box:

Please start/continue your search now...

Search: ● the web ○ pages from Malaysia

Figure 4. Search interface - searching section


Google's Search Interface - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites
Address http://kmice.uum.edu.my/fmy2/index2.asp
Go Links
Web Search Bookmarks Settings Messenger Mail MySpace News
Search
Web Images Maps News Books Translate Gmail more
Search settings | Sign
Google
Search
Advanced Search
Web
Show options...
Sponsored Link
Sponsored Links
Done
Internet

Enter your query in this box: homestay malaysia

Please start/continue your search now...

homestay malaysia

Search: ● the web ○ pages from Malaysia

Results **1** - **10** of about **582,000** for **homestay malaysia**. (**0.10** seconds)

Homestay
Mudah.my/Houses **Homestay** At Mudah. **Malaysia**'s Top Online Marketplace!

Tip: Save time by hitting the return key instead of clicking on "search"

Homestay Malaysia: Direktori - [Translate this page]
Program **Homestay** diperkenalkan oleh Kementerian Pelancongan **Malaysia** pada tahun 1988 sebagai satu usaha mempelbagaikan produk pelancongan melalui penawaran **...**
Senarai Homestay - Johor - Perak - Selangor

Kuala Lumpur Hotels
Discounts Kuala Lumpur hotels!
Save up to 75% on your reservation.
www.Hotels2Foru.com

Become a host family
Offer **homestay** accommodation to international students worldwide.

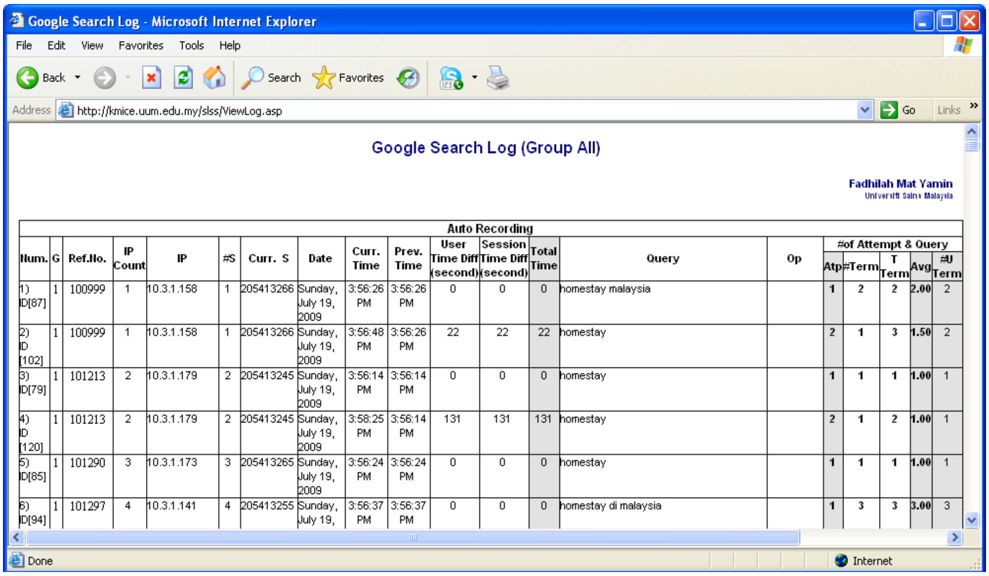
Figure 5. Example of search session

Users' queries and other information from the searching session are recorded in the transaction log. A Web transaction log is a file where all activities on the Web are recorded. A transaction log is a common source of information to investigate the Web search behaviour[13, 14, 15]. Figure 6 shows the example of the search log that was used in this study. The search log contains information about the user and the computer used such as user ID and computer IP and information about the search session which includes the session ID, date and time. Other items in the log such as time difference, IP and session counter, number of attempts and queries, number of terms, terms average and number of unique terms were calculated by the system.

Table 1 shows the list of items in the log and their descriptions. In this study, only queries were taken for analysis. Other information was used as a reference.

Table 1. Log item and description

Column	Item	Description
1	Num (and record ID)	Num is a continuous line number and the record ID is a reference number of the record in database
2	G (Group number)	Indicate the group number
3	Ref. No.	Ref. No. is the user ID that is used as a reference for the particular user.
4	IP Count	Counting the number of IP –the counter increase when new IP found
5	# S (Session)	Counting the number of session - the counter increase when new session found
6	Curr. S	Shows the current session
7	Date	Shows the date
8	Curr. Time	Shows the time of the current search session
9	Prev. Time	Shows the time of the previous search session
10	User Time Diff (second)	Shows the time different (in second) for each user based on current and previous search session
11	Session Time Diff (second)	Shows the time different (in second) for each session based on current and previous search session
12	Total Time	Total time taken by each user to complete the search task
13	Query	Query entered by user
14	Op (Operator)	Boolean operator used
15	# of Attempt & Query	Summarize the query used by each user
15 (a)	Atp (Attempt)	Shows the number of attempt made by user
15 (b)	# Term	Shows the number of term used
15 (c)	T Term	Shows the total number of the terms
15 (d)	Avg (Average)	The query average.
15 (e)	# U Term	Number of unique terms in the query



Address: http://kmice.uum.edu.my/slss/ViewLog.asp

Google Search Log (Group All)

Fadhilah Mat Yamin
Universiti Sains Malaysia

Auto Recording															#of Attempt & Query				
Num.	G	Ref.No.	IP Count	IP	#S	Curr. S	Date	Curr. Time	Prev. Time	User Time Diff (second)	Session Time Diff (second)	Total Time	Query	Op	Atp	#Term	T Term	Avg	#U Term
1) ID[87]	1	100999	1	10.3.1.158	1	205413266	Sunday, July 19, 2009	3:56:26 PM	3:56:26 PM	0	0	0	homestay malaysia		1	2	2	2.00	2
2) ID [102]	1	100999	1	10.3.1.158	1	205413266	Sunday, July 19, 2009	3:56:48 PM	3:56:26 PM	22	22	22	homestay		2	1	3	1.50	2
3) ID[79]	1	101213	2	10.3.1.179	2	205413245	Sunday, July 19, 2009	3:56:14 PM	3:56:14 PM	0	0	0	homestay		1	1	1	1.00	1
4) ID [120]	1	101213	2	10.3.1.179	2	205413245	Sunday, July 19, 2009	3:58:25 PM	3:56:14 PM	131	131	131	homestay		2	1	2	1.00	1
5) ID[85]	1	101290	3	10.3.1.173	3	205413265	Sunday, July 19, 2009	3:56:24 PM	3:56:24 PM	0	0	0	homestay		1	1	1	1.00	1
6) ID[94]	1	101297	4	10.3.1.141	4	205413255	Sunday, July 19,	3:56:37 PM	3:56:37 PM	0	0	0	homestay di malaysia		1	3	3	3.00	3

Figure 6. Example of search log

4. METHOD OF SEARCH BEHAVIOUR ANALYSIS

The user search behaviour measurements are based on the users' activities as captured in the transaction log. A transaction log stores queries that were used during the searching session. The search log is undergoing processes to clean the data. Users who did not perform queries were removed. These users were recognized based on the query entered. Typically, users are expected to enter more than one query, which indicates that they formulate and reformulate the queries. Otherwise, the users are browsing, or visiting each link exhaustively. Browsing activity is not the scope of this study. The irrelevant queries were also identified and removed. Then query is classified either as breadth or depth search query[16]. Breadth query strategy is a broad usage of query. The query formulated is general, wide and not specific to the domain. Depth query strategy is a narrow usage of query. The query is narrowed into the domain, and the use of keyword is more specific towards the search task. The classification was based on the criteria in Table 2 and Table 3.

Table 2. Criteria for breadth search query

Breadth search query			
Coding Symbol	Strategy	Description	Example
B1	Keyword search	Directly typing the query subject	Typing the words Homestay
B2	Wide search definition	Searching using a broad query	Searching for Ministry of Tourism to find the Homestay
B3	General knowledge	Using information that is not mentioned in the search task	Searching for the Homestay mentioning Guest House.

Table 3. Criteria for depth search query

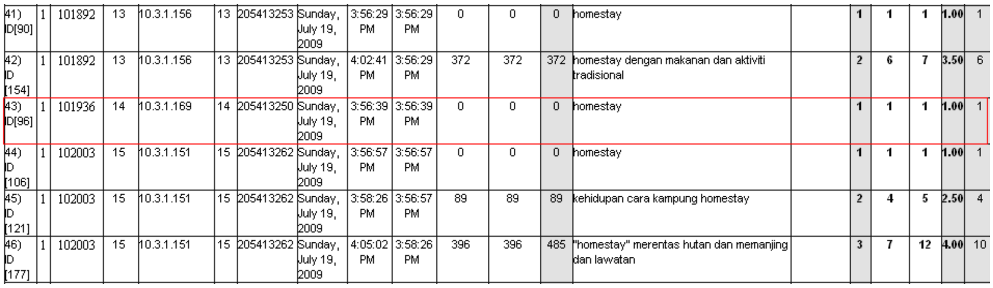
Depth search query			
Coding Symbol	Strategy	Description	Example
D1	Boolean search	Using Boolean syntax	Homestay AND Pahang
D2	Computer convention	Using a computer convention	Homestay.gif, homestay. jpeg
D3	Complex search	Cross searching with more than one query	Homestay, jungle tracking, fishing, etc.

The following are the steps of the coding process.

Step 1: The transaction log was retrieved from the database. Example of the log is shown in Figure 6.

Step 2: The transaction log was sorted according to the group, date/time, computer IP and session number. The example in Figure 6 has been sorted based on these criteria.

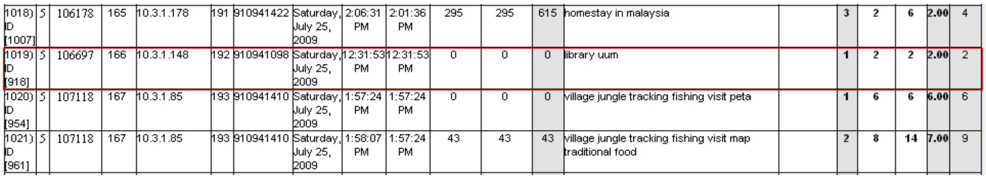
Step 3: Data cleaning was performed to remove single query users and irrelevant queries. Figure 7 shows example user with reference no. 101936 at line 43 only entered one query. This user was suspected to be browsing, which was not recorded in the transaction log. This user was discarded from the list.



41) ID[90]	1	101892	13	10.3.1.156	13	205413253	Sunday, July 19, 2009	3:56:29 PM	3:56:29 PM	0	0	0	homestay		**1**	**1**	**1**	**1.00**	1
42) ID [154]	1	101892	13	10.3.1.156	13	205413253	Sunday, July 19, 2009	4:02:41 PM	3:56:29 PM	372	372	372	homestay dengan makanan dan aktiviti tradisional		**2**	**6**	**7**	**3.50**	6
43) ID[96]	1	101936	14	10.3.1.169	14	205413250	Sunday, July 19, 2009	3:56:39 PM	3:56:39 PM	0	0	0	homestay		**1**	**1**	**1**	**1.00**	1
44) ID [106]	1	102003	15	10.3.1.151	15	205413262	Sunday, July 19, 2009	3:56:57 PM	3:56:57 PM	0	0	0	homestay		**1**	**1**	**1**	**1.00**	1
45) ID [121]	1	102003	15	10.3.1.151	15	205413262	Sunday, July 19, 2009	3:58:26 PM	3:56:57 PM	89	89	89	kehidupan cara kampung homestay		**2**	**4**	**5**	**2.50**	4
46) ID [177]	1	102003	15	10.3.1.151	15	205413262	Sunday, July 19, 2009	4:05:02 PM	3:58:26 PM	396	396	485	"homestay" merentas hutan dan memanjing dan lawatan		**3**	**7**	**12**	**4.00**	10

Figure 7. Example of single query user

During the search process, users were expected to formulate and use queries that were related to the homestay. Unrelated queries were removed from the list. Figure 8 shows examples of irrelevant queries. Users with reference no 106697 at line 1019 used the word "library uum" as the query. This query was not relevant to homestay, and therefore, this particular record was removed from the list.



1018) ID [1007]	5	106178	165	10.3.1.178	191	910941422	Saturday, July 25, 2009	2:06:31 PM	2:01:36 PM	295	295	615	homestay in malaysia		**3**	**2**	**6**	**2.00**	4
1019) ID [918]	5	106697	166	10.3.1.148	192	910941098	Saturday, July 25, 2009	12:31:53 PM	12:31:53 PM	0	0	0	library uum		**1**	**2**	**2**	**2.00**	2
1020) ID [954]	5	107118	167	10.3.1.85	193	910941410	Saturday, July 25, 2009	1:57:24 PM	1:57:24 PM	0	0	0	village jungle tracking fishing visit peta		**1**	**6**	**6**	**6.00**	6
1021) ID [961]	5	107118	167	10.3.1.85	193	910941410	Saturday, July 25, 2009	1:58:07 PM	1:57:24 PM	43	43	43	village jungle tracking fishing visit map traditional food		**2**	**8**	**14**	**7.00**	9

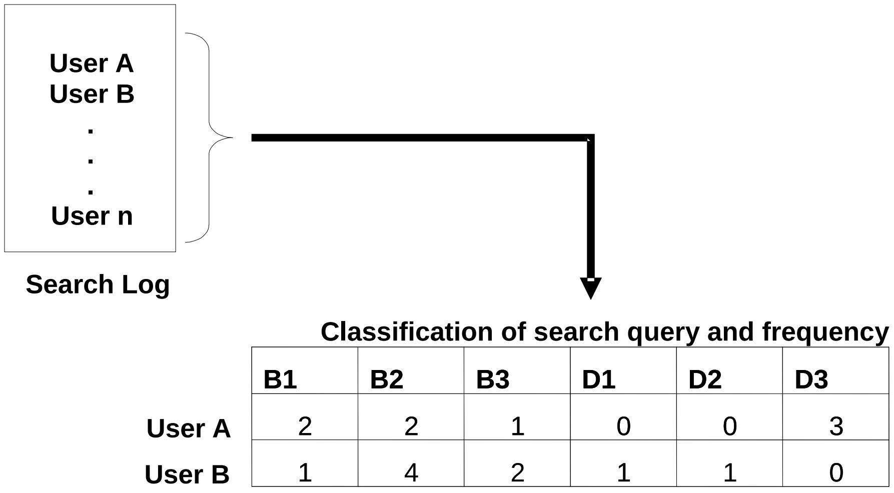
Figure 8. Example of irrelevant query

Step 4: Query classification or marking was performed. In this step, each query was examined and classified either as breadth or depth query. Table 4 shows examples of queries for users A and B. During the coding, each query strategy was denoted with a symbol. The breadth search queries are represented by symbols B1, B2 and B3, where each symbol represents keyword search, wide search and general knowledge, respectively. Depth search queries are represented by symbols D1, D2, and D3. D1 represents the Boolean search, while D2 and D3 represent computer convention and complex search, respectively.

Step 5: The classified queries were transferred into the table and the frequency of each query type was calculated. For ease of the analyses in Statistical Package for the Social Sciences (SPSS), the search logs were transferred into a table (Figure 9). For example, user A and user B were among the respondents in this study. Queries by both users were transferred into the table. Based on the query occurrences, user A formulated eight queries, showing that user A had eight attempts. Attempt is conceptualize as number of times the query entered by user. Out of these attempts, five queries fell under breadth search strategy, while the other three were depth search strategy. Further classification has shown that five queries identified as breadth search strategy can be divided into direct keyword (B1), wide search (B2), and general knowledge (B3) by which each strategy represents 2, 2, and 1 queries respectively. Three queries under the depth search query are classified as complex search strategy (D3). Other types of queries, Boolean operator (D1) and computer convention (D2), were not used by this user. User B, on the other hand, has formulated nine queries. Seven of the queries fall under breadth search query and the other two are depth search query.

Table 4. Example of queries and classification

User	Query	Classification
User A	Homestay	B1
	Homestay Malaysia	B2
	Homestay	B1
	Homestay in Kedah	B2
	Rumah tumpangan	B3
	Homestay, kedah, jungle tracking, makanan tradisional	D3
	Homestay, Malaysia, tradisional	D3
	Sarawak, jungle traking, package	D3
User B	Website Homestay	B1
	Homestay in Malaysia	B2
	Homestay Selangor	B2
	Perak homestay	B2
	Guest house	B3
	Kelantan Guest House	B3
	Malaysia Homestay Aktiviti	B2
	Homestay AND Terengganu	D1
	Peta ke homestay Kuala Medang	D2



Classification of search query and frequency

	B1	B2	B3	D1	D2	D3
User A	2	2	1	0	0	3
User B	1	4	2	1	1	0

Figure 9. Example of the classification of search query and frequency

Step 6: The mean for both breadth and depth search query was calculated. An example is shown in Table 5. These mean values are then used in the statistical analysis.

Table 5. Mean value for breadth and depth search query

User	Mean Breadth	Mean Depth
User A	1.667	1
User B	2.333	0.667

4. FINDINGS

In this study, a total 1,072 queries were extracted from the transaction log and analyzed. All queries are marked to determine their category (either breadth or depth search query). After the marking, 596 queries are classified as breadth search query and 476 are depth search query. Based on these findings, it might be obvious to say that the breadth search query is moderately higher compared to depth search query.

Table 6 shows specific mean values and standard deviation for each user search query categories. For the breadth search query, the mean values ranged from .97 to 2.36. Among the query categories in the breadth search query, “general knowledge” achieved the highest mean value, 2.36. The lowest mean value is the wide search category with the mean value .97.

In the second group of search query, depth search query, the mean values ranged from .62 to 1.63 and the standard deviation score is from .77 to 1.84. In this category, the complex search shows the highest mean value, followed by computer convention and Boolean operator.

Table 6. Mean values and standard deviation for query search formulation

Query categories	Query	Total	Mean	Std. deviation
Breadth search query	Direct search	162	1.24	1.39
	Wide search	127	0.97	1.21
	General knowledge	307	2.36	2.86
Depth search query	Boolean operator	81	0.62	0.77
	Computer convention	181	1.38	1.40
	Complex search	214	1.63	1.84

5. DISCUSSION AND CONCLUSION

User search behaviour contains two dimensions; breadth search query and depth search query. These findings show that a user had performed more than one attempt during the completion of the search task. This is in line with Spink et al.[17], which indicated that users often repeated the search process in their searching activities. The mean for the breadth search query is slightly higher compared to the depth search query. This indicates that the users have fully utilized breadth search query in their searching. This finding is in_line with Park et al.[15], which indicated that Internet users do not fully utilized advanced search strategies as opposed to basic Internet searching. In this study, an advanced search strategy is categorized under depth search query category. However, a small difference of mean value between the breadth search and depth search shows that the users tend to improve their searching with depth query.

The details of the descriptive analysis for breadth search query indicate that the most popular types of query used by the respondents are general knowledge. This finding is in line with Wildemuth[18] which indicates that user will begin with general knowledge and tend to focus on the search topic. Direct search is the second-most popular search. Direct search strategy is influenced by the search task, in which a user can extract the keyword from the search task. This scenario is parallel with a study by Jansen[19]. The third strategy is wide search. Wide search is the lowest in popularity, as those using this strategy require some general knowledge related to the search topic.

The analysis for depth search query indicates that the popular types of query are complex search and computer convention. The analysis also reveals that the usage of Boolean operator is the lowest. This is in line with several findings which indicate that the usage of Boolean operator is not a popular strategy among Internet users[20, 21].

Based on the above analysis, it appears that the respondents are utilizing both behaviors during their searching. It is suspected that the respondents switched their strategy from broad to narrow in order to achieve their goal.

6. REFERENCES

[1] M. Kunder, *The size of the World Wide Web*. Retrieved on March 19, 2010, from http://www.worldwidewebsize.com.

[2] V. Shkapenyuk, and T. Suel, Design and implementation of a high performance distributed web crawler. In Rakesh Agrawal, Klans Dittrich, and Anne H.H Ngu (Eds.), *Proceedings of the 18th International Conference on Data Engineering (ICDE)* (p357-368). San Jose, California: IEEE CS Press, 2002. http://dx.doi.org/10.1109/ICDE.2002.994750.

[3] S. Cazalens, E. Desmontils, C. Jacquin, and P. Lamarre, A web site indexing process for an internet information retrieval agent system. In Qing Li, Z. Meral Ozsoyoglu, Roland Wagner, Yashikl Kambayashi, Yanchun Zhang (Eds.), *Proceedings of the First International Conference on Web Information Systems Engineering* (p254-258). Hong Kong: The Institute of Electrical and Electronics Engineers, Inc, 2000. http://dx.doi.org/10.1109/WISE.2000.882400.

[4] C. Jenkins, M. Jackson, P. Burden, and J. Wallis, Searching the World Wide Web: An evaluation of available tools and methodologies. *Information and Software Technology*, 39(14-15), p985-994, 1998. http://dx.doi.org/10.1016/S0950-5849(97)00061-X.

[5] J. Day, The quest for information: A guide to searching in the Internet. *Journal of Contemporary Dental Practice*, 2(4), p33-43, 2001.

[6] M. Nazim, Information searching behaviour in the Internet age: A users' study of Aligarh Muslim University. *The International Information & Library Review*, 40(1), p73-81, 2008. http://dx.doi.org/10.1016/j.iilr.2007.11.001.

[7] A. Spink, and B.J. Jansen, A study of web search trends. *Webology*, 1(2). Retrieved on April 24, 2007, from http://www.webology.ir/2004/v1n2/a4.html.

[8] S. Brin, and L. Page, The anatomy of a large scale hyper textual web search engine. *Computer Networks and ISDN Systems*, 30(1-7),

p107-117, 1998. http://dx.doi.org/10.1016/S0169-7552(98)00110-X.

[9] M.Y. Fadhilah, and T. Ramayah, User knowledge on web search strategy: The importance of topic and web search system understanding. *Paper presented at* 12th *International Business Information Management Association (IBIMA09) Conference*, Kuala Lumpur, Malaysia, June 29-30, 2009.

[10] M.Y. Fadhilah, and T. Ramayah, Searching the web: The impact of user knowledge on search satisfaction. *Paper presented at Knowledge Management International Conference and Exhibition (KMICE2010)*, Kuala Terengganu, Malaysia, May 25-27, 2010.

[11] B.J. Jansen, Search log analysis: What is, what's been done, how to do it. *Library & Information Science Research*, 28(3), p407-432, 1998. http://dx.doi.org/10.1016/j.lisr.2006.06.005.

[12] P. Wang, W.B. Hawk, and C. Tenopir, Users' interaction with World Wide Web resources: An exploratory study using a holistic approach. *Information Processing and Management*, 36(2), p229-251, 2000. http://dx.doi.org/10.1016/S0306-4573(99)00059-X.

[13] B.J. Jansen, The effect of query complexity on web searching results. *Information Research*, 6(1). Retrieved on April 23, 2007, from http://informationr.net/ir/6-1/paper87.html.

[14] B.J Jansen, A. Spink, and T. Saracevic, Real life, real users, and real needs: A study and analysis of user queries on the Web. *Information Processing and Management*, 36(2), p207-227, 2000. http://dx.doi.org/10.1016/S0306-4573(99)00056-4.

[15] S. Park, J.H. Lee, and H.J. Bae, End user searching: A Web log analysis of NAVER, a Korean web search engine. *Library of Information Science Research*, 27(2), p203-221, 2005. http://dx.doi.org/10.1016/j.lisr.2005.01.013.

[16] M.Y. Fadhilah, and T. Ramayah, User web search behaviour and query formulation. *Paper presented at 2011 International Conference on Semantic Technology and Information Retrieval (STAIR11)*, Putrajaya, June 28-29, 2011.

[17] A. Spink, J. Bateman, and B.J. Jasen, Searching heterogeneous collections on the Web: Behavior of excite users. *Information Research,* 4(2). Retrieved on April 24, 2007, from http://informationr.net/ir/4-2/paper53.html.

[18] B.M. Wildemuth, The effects of domain knowledge on search tactic formulation. *Journal of the American Society for Information Science and Technology*, 55(3), p246-258, 2004. http://dx.doi.org/10.1002/asi.10367.

[19] B.J. Jansen, The effect of query complexity on web searching results. *Information Research*, 6(1). Retrieved on April 23, 2007, from http://informationr.net/ir/6-1/paper87.html.

[20] D. Wolfram, A. Spink, B.J. Jansen, and T. Saracevic, VOX POPULI: The public searching of the Web. *Journal of the American Society for Information Science and Technology*, 52(12), p1073-1074, 2001. http://dx.doi.org/10.1002/asi.1157.

[21] M.SM. Saad, and A.N. Zainab, Undergraduates in computer science and information technology using the internet as a resource. *Malaysia Journal of Library & Information Science*, 9(1), p1-16, 2004.