COMPARATIVE STUDY OF SOME POPULAR WEB SEARCH ENGINES

Olalekan S. Akinola
Department of Computer Science,
University of Ibadan,
Ibadan,
Nigeria
Email: solom202@yahoo.co.uk

ABSTRACT

Web search engines are veritable tools for information mining. These search engines differ in effectiveness at retrieving relevant documents from the web. A big question is "which search engine(s) is/are best to use in specific instances?" This work evaluates effectiveness of five search engines on fifteen Computer Science queries. Measures obtained from this study include search output quantity, search time, relevant documents retrieved, search precision and the quantity of African contents in the search. Briefly the study shows that (1) the fact that a Web search engine is very popular and able to retrieve large number of documents does not mean that it has high precision for retrieving relevant documents for its users and (2) Africa has contributed little or nothing to the global Web contents in the field of Computer Science. However, this work needs to be extended to other academic areas in order to generalize these results.

Keywords: Information Retrieval, Effectiveness, Web, Search Engine

I. INTRODUCTION

The World Wide Web (Web) is one of the largest publicly available databases of documents on the Internet [9]. It is a good testing ground for most information retrieval techniques.

The growth of the Web is an unprecedented phenomenon. Four years after its birth in 1990, a million or more copies of the first well-known Web browser, Mosaic, were in use [1]. This growth was a result of the exponential increase of Web servers and the value and number of web pages made accessible by these servers. In 1999 the number of Web servers was estimated at about 3 million and the number of web pages at about 800 million [18], and three years later, in June 2002, the search engine AlltheWeb, announced that its index contained information of about 2.1 billion web pages [9].

There are millions of Web users and about 85% of them use search engines to locate information on the Web [17]. It is discovered that search engine use is the second most popular Internet activity next to e-mail [15]. Due to high demand there are hundreds of general purpose and thousands of specialized search engines on the Internet today.

A Web search engine is an information retrieval system [20], which is used to locate the Web pages relevant to user queries. Search engines for the general web do not really search the Web directly. Each one searches a database of the full text of Web Pages selected from the billions of Web pages out there residing on servers. When the web is searched using a search engine such as Google or Alta Vista, a somewhat stale copy of the real web page is searched. When we click on links provided in a search engine's search results, we retrieve the current version of the Web page [6].

Search engine databases are selected and built by computer robot programs called *Spiders* [22]. Although it is said that they "crawl" the web in their hunt for pages to include, in truth they stay in one place. They find the pages for potential inclusion by following the links in the pages they already have in their database (that is, already "know about"). They cannot think or type a Universal Resource Locator (URL) or use judgment to "decide" to go and look something up and see what is on the web about it.

If a Web page is never linked to any other page, search engine spiders cannot find it. The only way a brand new page, one that no other page has ever linked to, can get into search engine companies is by requesting that the new page be included. All search engine companies offer ways to do this. After spiders find pages, they pass them on to another computer program for "indexing". This program identifies the text, links and other content in the page and stores it in the search engine database's files so that the database can be searched by keyword and whatever more advanced approaches are offered, and the page will be found if the search query matches its content.

The motivation we have for this study is based on the fact that identifying the most effective Web search engines (among the popular ones we use everyday) satisfying the current information needs is important both at personal and business levels. The results of this research can be used for (a) finding more relevant documents with less effort; (b) finding out which of these popular search engines is most appropriate and effective for searching documents for academic and research purposes and (c) motivating search engine providers for higher standards.

In this paper, our case study was based on five search engines, which are frequently used by Internet users when sourcing for information on the Internet [4]. These are Google (www.google.com), Microsoft Network (MSN, www.msn.com), Alta vista (www.altavista.com), 37 (www.37.com) and Yahoo (www.yahoo.com) search engines. A list of fifteen computer science – related keywords was searched with these search engines and the following information obtained from them: Quantity of search output (that is, effectiveness), time used for the search, the advert content, relevant results and the number of African contents.

Results from the study have great implications for both students and academics using Internet for finding information for their academic endeavours.

This paper is organized as follows. In section 2 we give an overview of some related works. We then explain the methodology and / or experimental set up in terms of queries and search engines involved in Section 3. Detailed experimental results and discussion on them are presented in Section 4. Conclusion and recommendations are finally presented in Section 5.

II.   LITERATURE REVIEW

The explosive growth of the Internet has rendered the World Wide Web as the primary tool for information retrieval today [12]. Adeyemo and Osunade [2] assert that in Africa research institutions, universities and media houses are the major users of the Internet. A survey of 1000 academics in 10 of the universities in Nigeria shows that the web, one of the internet resources, is used by 92.5% of the academics surveyed [3]. According to Choo, *et al*. [23], people use the Web as information resource to support their daily work activities and engage in a range of complementary modes of information seeking.

For information retrieval purposes on the Web, most people use search engines like Google, Yahoo and MSN. Sullivan [24] and Griesbaum [11] assert that Google is the dominant and most successful search engine among its contemporaries. For assessing the performance of these search engines, there are various measures such as database coverage, query response time, user effort and retrieval effectiveness, that can be used [9]. The dynamic nature of the Web also brings some more performance measure concerns regarding index freshness and availability of the Web pages as time passes [5]. The most common effectiveness measures are

precision (ratio of retrieved relevant documents to the total number of retrieved documents) and ***recall*** (ratio of retrieved relevant documents to the total number of relevant documents in the database) [9].

Measuring the search engine effectiveness is expensive due to the human labour involved in judging relevancy. This has been reported in several studies [9, 13]. Evaluation of search engines may need to be done often due to changing needs of users or the dynamic nature of search engines (for example, their changing Web coverage and ranking technology) and therefore it needs to be efficient [9].

Per Thousand Precision approach was formulated and used for computing the precision of the search engines used in this study. Since we based our human judgment on the first 1000 top documents retrieved for the relevant documents retrieved by the search engines, we therefore adopted this approach, although

the first top 20 documents are commonly used [16].

There are two types of search engine evaluation approaches in the literature: testimonial and shootout [9]. Testimonials are casual studies and state the general impression obtained after executing a few queries. Shootouts are rigorous studies and follow the information retrieval measures for evaluation purposes. Our approach in this study is of the type shootout. The Jansen and Pooch [15] study provides an extensive review and analysis of current Web searching results. It provides a rich reference list and proposes a framework for future research on Web searching.

In addition, the Gordon and Pathak [10] study, which measures the performance of eight search engines using 33 information needs also recommends seven features to maximize the accuracy and informative content of such studies (see Table 1 below). Hawkins et al., [13] also proposes some more features in addition to the seven items specified in Gordon and Pathak [10] (also in Table 2.1)

Table 2.1: Desirable features of Web search evaluation according to Gordon and Pathak [10] : features 1 to 7 and Hawking et al.,[13]: features 8-11.

---

1. The searches should be motivated by genuine information needs of Web users.
2. If a search intermediary is employed, the primary searcher's information need should be captured as fully as and with as much context possible and transmitted to the intermediary.
3. A sufficiently large number of searches must be conducted to obtain meaningful evaluations of search engines' effectiveness.
4. Most major search engines should be considered.
5. The most effective combination of specific features of each search engine should be exploited (i.e. the queries submitted to the engines may be different).
6. The user who needs the information must make relevance judgments.
7. Experiments should (a) prevent bias towards search engines (e.g. by blinding or randomizing search outputs), (b) use accepted information retrieval measures, (c) employ statistical tests to measure performance differences of search engines.
8. The search topics should represent the range of information needs over which it is desired to draw conclusions.

9. Result judging should be appropriate to the type of query submitted (e.g. some queries may need a one-line answer).
10. Document presentation should be like that of a Web browser (images should be viewable, if necessary, it should be possible to follow links).
11. Dead links should count as useless answers.

Source: (Fazli, et al, [9]).

In this study, we satisfy most of the features given in Table 1 above. Features 2 and 5 are exceptionally disregarded in this study. Feature 2 does not really apply to us and contrary to the suggestions of the authors, in such studies we believe that genuine user queries should be used. By this way, the true everyday behaviour of the search engines will be measured [9]. Feature 5 needs expert searchers claiming that the best query formulation should be used for a specific search engine. Since most Web searchers in real life are casual users, we therefore disagree with this point. In the Web environment, there is no standardized test collection with complete relevance judgments [13]. Therefore we could only satisfy 6 to a little extent in this study. Being the dynamic and huge environment that it is, makes it almost impossible to determine such a collection and to obtain human relevance judgments, since it would be too cumbersome for people to judge so many Web pages.

Literature shows that several studies have been carried out by researchers in line with all these features with some modifications and / or enhancements [7, 19, 21]. For instance, Griesbaum [11] compares the performances of three German search engines – Altavista, Google and Lycos – in terms of relevance and precision using 50 queries. Results from his study shows that Google reached the highest values, followed by Lycos and then AltaVista. Hananzita and Kiran [12] also compare the performances of four Malaysian web search engines using Google as the benchmark search engine. Their results also show that Google outperforms the four engines considered.

III. METHODOLOGY / EXPERIMENTAL SET UP

III.1 User Queries

Measuring retrieval effectiveness requires user queries. For the construction of the queries used in this study, fifteen Computer Science – related information needs (queries) were described in terms of keywords that were to be used for the study. These keywords covered a broad range of topics in Computer Science profession. The study used AND and OR Boolean operators when the keywords were given to the search engines to search. The following keywords were used for this study:

1. Algorithms(ALG)
2. Data Structures (DS)
3. J2Me
4. Operating Systems (OS)
5. Database Management Systems (DBMS)
6. Hacking (HK)
7. Microprocessors (MP)
8. Recursion (RC)
9. Software (SW)
10. Compilers (CP)
11. Optical Character Recognition (OCR)
12. Networking (NW)
13. World Wide Web (WWW)
14. Multimedia (MM)
15. Artificial Intelligence (AI)

III.2 Search Engines

Five popular search engines among students and academic researchers [4] were selected in conducting our experiment. They are as mentioned in the

penultimate paragraph of Section 1 of this paper.

III.3 Performance Evaluation Measures

Based on the keywords used in the experiment, we aim at assessing the individual search engines based on:

i.   **Quantity** of documents retrieved by the engines;

ii.   **Time taken** for each keyword search to be completed by the engines (Response time): while three of the search engines (Yahoo, MSN and Google) reported their time of search as part of their results we measured this manually by a digital timer for those that do not have the timing facility (Alta vista and 37.com);

iii.   **Per Thousand Precision** of documents retrieved: we observed and recorded the number of results that gave us relevant information about the keywords used, divide this by the total number of documents retrieved and finally multiplied by 1000 (reason aforementioned);

iv.   **Advert Content**: for each search engine, we observe the amount of commercial- oriented documents produced at each keyword searched; and

v.   **African content**: we enlisted this criterion to determine how much has the African continent as a whole been able to establish herself in the world of computer science today. We determine this by scrutinizing the information given by the links to see if we can find any African related author's name and address.

III.4   The Search Process

In this step, we used the queries formulated and submitted them to the chosen search engines. For each keyword given to the search engines, we diligently took all the performance evaluation measures as indicated in subsection 3.3 above. All the experiments were done on a Pentium 4 MMX, 40MB, 2.0MHz computer located in the Ayo Rosiji Research Laboratory, at the Kola Daisi Building, in the Department of Computer Science, University of Ibadan, Nigeria on $11 - 16$ October, 2008. For the measures of Advert content and African content, we based our human judgment on the top 1000 documents retrieved by the search engines. The only exception is the 37.com search engine, of which total documents retrieved are always less than 50.

IV.   EXPERIMENTAL RESULTS
IV.1   The   Search   Engines' Performances

Table 4.1 gives the quantity of documents (Q), retrieved and time taken (T) in seconds for each query given to the search engines. The keywords or queries are abbreviated in the table simply for space purposes but they were used directly as normal texts for the search study.

Table 4.1: Quantity of Documents retrieved (Q) and Time in Seconds (T)

S/N	Query (Key Words)	Google		MSN		Yahoo		Alta Vista		37.com	
		Q	T	Q	T	Q	T	Q	T	Q	T
1.	ALG	9240000	0.16	1583408	0.19	110000	0.15	224813	0.50	34	0.54
2.	DS	7270000	0.19	1739483	0.23	435115	0.10	7516787	0.59	11	0.64
3.	J2ME	726000	0.16	207605	0.19	67500	0.40	125157	0.50	33	0.54
4.	OS	11600001	0.18	6255462	0.22	1000000	0.75	3674348	0.56	29	0.61
5.	DBMS	8820020	0.17	171494	0.20	105423	0.20	177320	0.53	34	0.57
6.	HK	5370000	0.15	1305195	0.18	795112	0.60	1992876	0.47	31	0.51
7.	MP	1180120	0.18	243040	0.22	124133	0.15	266029	0.56	33	0.61
8.	RC	753000	0.16	113116	0.19	100000	0.65	161083	0.50	27	0.54
9.	SW	23500000	0.16	8123	0.19	1315000	0.15	48117366	0.50	38	0.54
10.	CP	5910000	0.15	364205	0.18	652000	0.70	525593	0.47	34	0.51
11.	OCR	277000	0.19	74879	0.23	50000	0.80	58903	0.59	30	0.64
12.	NW	22500000	0.14	3390131	0.17	1100000	0.60	6606411	0.44	7	0.47
13.	WWW	8230000	0.14	34783984	0.17	85000000	0.15	71309208	0.44	32	0.47
14.	MM	35900000	0.17	13588351	0.20	73000000	0.30	46	0.53	15	0.57
15.	AI	3400000	0.16	776817	0.19	330000	0.40	625816	0.50	31	0.54
MEAN		23745076	0.16	4307019	0.20	10945619	0.41	9560338	0.51	28	0.55

Mean Quantity(Q) vs. Search Engines
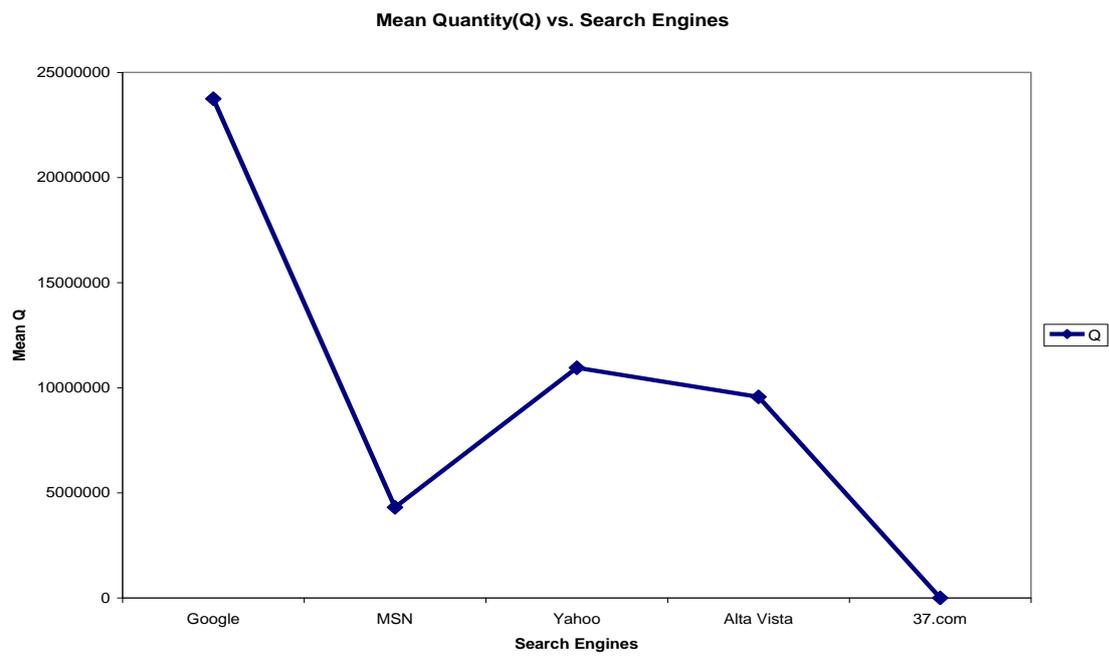
Fig. 4.1: Mean Quantity (Q) of Documents Retrieved by the Search Engines
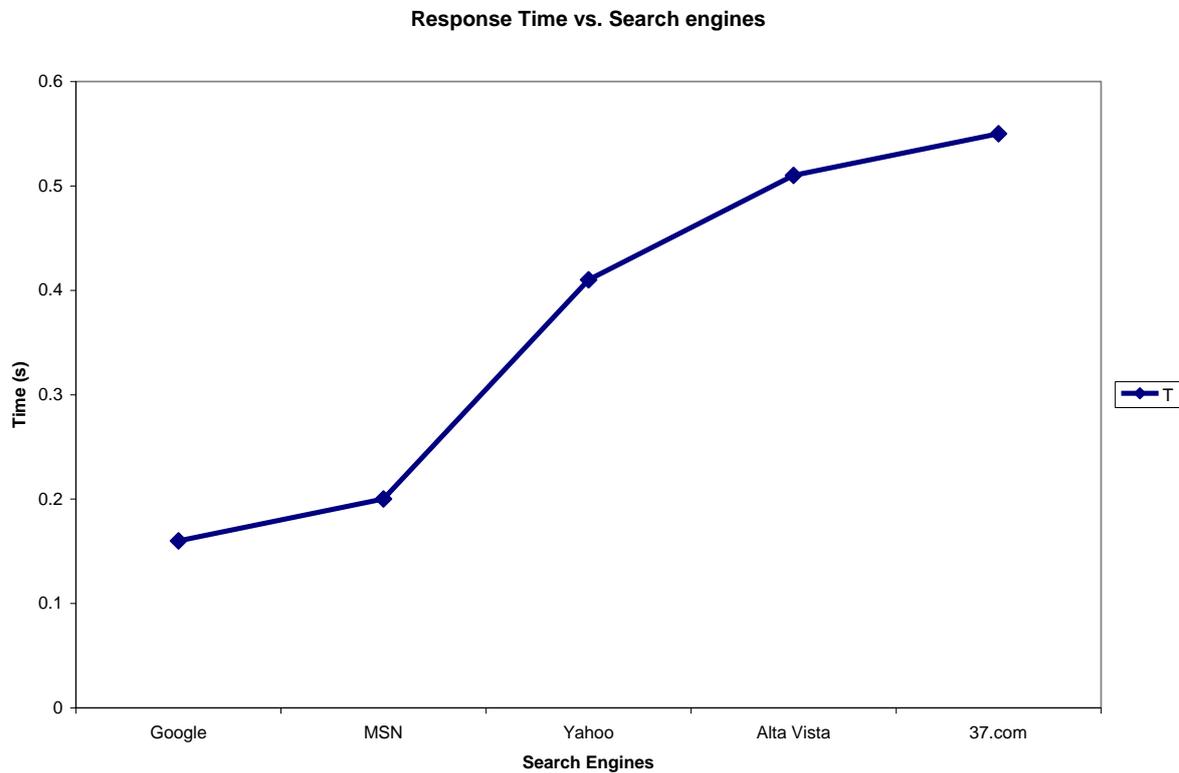
Response Time vs. Search engines

Fig. 4.2: Mean Response Time (T) in Seconds versus Search Engines

The data presented in Table 4.1 and Figs. 4.1 and 4.2 above reveals that Google search engine has the highest document retrieval efficiency in terms of total number of documents retrieved (mean = 23,745,076) for all the queries (key words) given. This trend is followed by Yahoo (mean = 10,945,619), Alta vista (mean = 9,560,338), MSN (mean = 4, 307,019) and 37.com (mean = 28) in that order. This result is in line with the ones obtained by Griesbaum [11] and Hananzita and Kiran [12] where Google was reported to have the highest document retrieval from the Web. Time, on the other hand followed a reverse trend. The search engine with the minimum document retrieval efficiency (37.com) has the highest time taken (mean = 0.55s) for retrieving the few number of documents. This is followed by Alta vista (mean = 0.51s) and Yahoo (mean = 0.41s), which have more or less the same mean number of documents retrieval efficiency. Google and MSN have comparable mean time values of 0.16s and 0.20s respectively. Hong and David [14] also

demonstrate that Google always take a very small time for document retrieval on the Web.

IV.2   Relevant Documents Retrieved (Precision) and Advert Content

Table 4.3 gives the quantity of relevant results from the searches, which we eventually computed as Per Thousand Precision (PTP) and the advert contents in the top 1000 documents retrieved by the search engines. The only exception is the 37.com, which does not seem to have good efficiency for all the queries used in the study. We only considered the top 1000 documents retrieved by the search engines in order to determine these measures since it is time and effort costly to go through all the documents retrieved by these search engines one by one, less for the 37.com. Even the top 1000 documents chosen for this study was very time consuming and strenuous as we have to click on each of their links to see what contents they have.

In the table, the advert content (AC) gives the total number of documents that are advertising in nature possibly for selling books or software or asking the web user to subscribe to journals or publishers in charge of the documents. Per Thousand Precision (PTP) is calculated thus:

$$PTP = \frac{\text{Total number of retrieved relevant documents (R)} \times 1000}{\text{Total number of retrieved documents}}$$

There are some overlaps in the advert and relevant documents retrieved in that some

documents contain advertisement and at the same time having some useful introduction and / or information about the subject matter of the query. Some of the documents that we cannot classify into any of these two categories were simply disregarded for these measures in the study.

**Table 4.3: Advert Content (AC), Relevant Documents Retrieved (R) and
Per Thousand Precision (PTP)**

S/N	Query (Key Words)	GOOGLE			MSN			YAHOO			ALTA VISTA			37.COM		
		AC	R	PTP	AC	R	PTP	AC	R	PTP	AC	R	PTP	AC	R	PTP
1	ALG	406	563	0.061	250	267	0.169	597	184	1.673	430	575	0.256	6	18	529.41
2	DS	350	584	0.080	267	533	0.306	248	176	0.405	540	554	0.074	1	7	636.36
3	J2ME	337	711	0.979	505	190	0.915	254	211	3.126	421	604	4.826	5	26	787.88
4	OS	200	654	0.056	378	467	0.075	753	145	0.145	240	756	0.206	17	8	275.86
5	DBMS	350	874	0.099	433	233	1.359	708	167	1.584	298	697	3.931	22	12	352.94
6	HK	354	550	0.102	133	383	0.293	654	188	0.237	350	658	0.330	2	28	903.23
7	MP	267	685	0.581	417	317	1.304	456	174	1.402	350	651	2.447	18	11	333.33
8	RC	369	728	0.967	117	600	5.304	427	141	1.410	670	328	2.036	2	23	851.85
9	SW	324	597	0.003	525	133	16.37	850	207	0.157	411	613	0.013	19	36	947.37
10	CP	417	627	0.106	433	256	0.703	750	159	0.244	710	287	0.546	7	20	588.24
11	OCR	413	741	2.675	343	419	5.596	507	174	3.480	630	375	6.366	14	14	466.67
12	NW	350	650	0.029	495	381	0.112	457	138	0.126	200	797	0.121	0	7	1000.0
13	WWW	206	655	0.079	200	200	0.001	508	147	0.002	250	752	0.105	3	12	375.00
14	MM	401	597	0.017	362	257	0.109	447	189	0.003	10	36	782.61	7	8	533.33
15	AI	250	751	0.221	105	543	0.699	457	137	0.415	240	779	1.245	5	22	709.68
MEAN		333	664	0.404	331	345	2.221	538	169	0.96	383	564	53.674	9	17	619.41

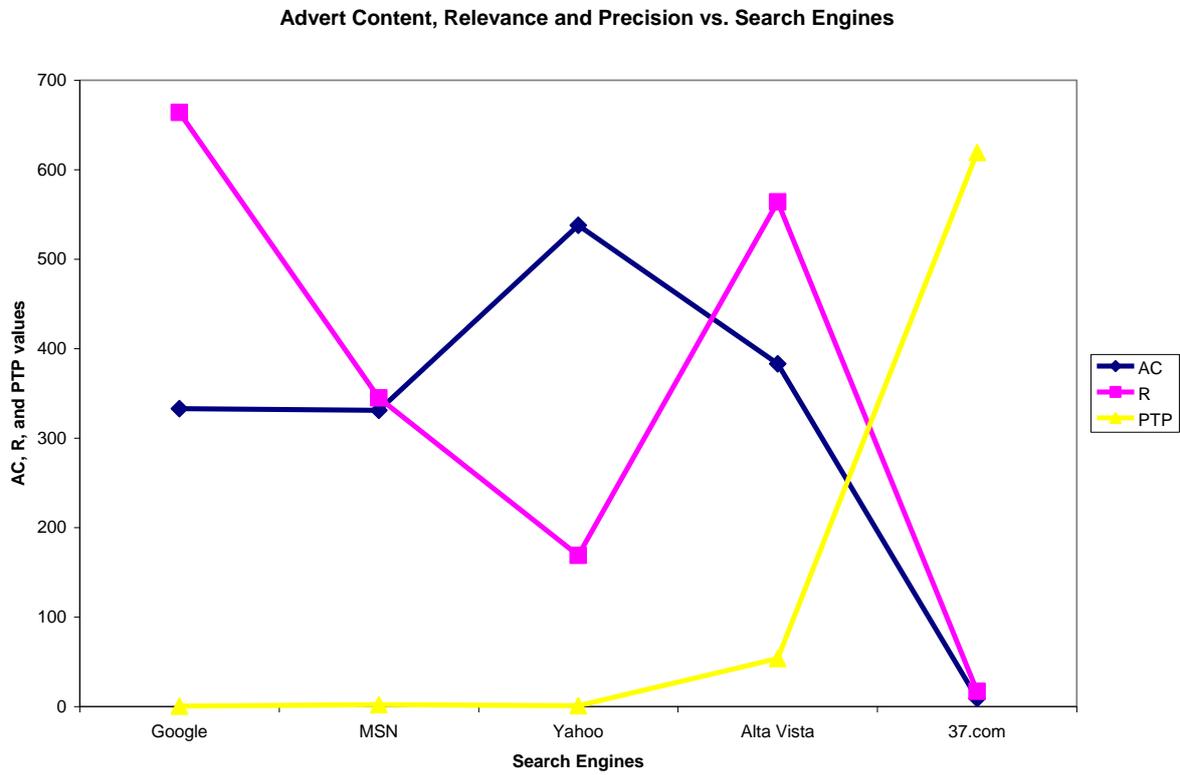Advert Content, Relevance and Precision vs. Search Engines

Fig. 4.3: Chart of Advert Content (AC), Relevant Documents (R), and Precision (PTP) versus the Search Engines

Table 4.3 and Fig. 4.3 above reveals that Yahoo search engine has the highest Mean Advert Content (AC, mean = 538) followed by Alta vista (mean = 383), Google (mean = 333), MSN (mean = 331) and 37.com (mean = 9) in that order.

In the case of Mean Relevant Documents Retrieved (R), Google has the highest (mean = 664) followed by Alta vista (mean = 564), MSN (mean = 345), Yahoo (mean = 169), and 37.com (mean = 17), in that order. This result is in consonance with the work of Griesbaum [11] in which Google had the highest relevance followed by Lycos and AltaVista in that order.

In the case of Mean Per Thousand Precision, which is one of the universal methods of determining the effectiveness of web search engines, 37.com happens to have the highest (mean = 619.41), followed by Alta vista (mean = 53.67), MSN (mean = 2.22), Yahoo (mean = 0.96) and Google (mean = 0.40) in that order. Griesbaum [11] shows that differences of precision between Google, lycos and Alta Vista are very low. Therefore, the result obtained for Google having highest relevance initially and now with the lowest precision is fairly justified. Hananzita [12] asserts that recall by itself is not a good measure of the quality of a search engine. The precision gives more information about a search engine's effectiveness at retrieving relevant documents from the Web.

IV.3   African Content

The African content (AF) is the total number of documents we adjudged to be African in nature, that is, documents we found that the author's name or the author's address comes from Africa. Only

two of the keywords searched were able to be African in nature. These are Hacking and Multimedia. Yahoo search engine was able to retrieve 2 documents for the Hacking keyword, which we really found to be African in nature. One of the documents was a Nigerian Newspaper report on electronic hacking problem in Nigeria. The other document seems to come from Ghana from the author's address seen. With Multimedia keyword, MSN retrieved one document, which actually comes from South Africa.

V.   DISCUSSION OF RESULTS

Web search engines are meant for information retrieval for academic and research purposes among others. These search engines are numerous and new ones are springing up everyday on the Internet. The main question is "which of the existing search engines is / are very effective for information retrieval on the web for a specific purpose?" Effective measures are done via the total documents retrieved, time taken for the retrieval, effort expended and most importantly the precision and recall of the documents retrieved.

Google seems to have the best documents retrieval performance in terms of mean number of documents retrieved and relevance with little time of search. However, it has the least effectiveness in terms of precision. The results show that the fact that a search engine brings out large quantity of documents when used for Web search does not mean that it is highly effective. What this suggests are that (1) possibly some of the documents retrieved may have 'dead links' since stale copies of the universal resource locator (URL) addresses of the documents are also

indexed and stored in the search engine's database. It is when the links to them are clicked that the search engine will go and look for the current version of the documents, which might have been moved or no more valid; and (2) possibly most of the documents retrieved are advertising in nature. They may not contain the information needed by the searcher at a particular period of time. More so, Google is a very poplar search engine and since many people, institutions and organizations are associating themselves with it, it therefore means that the search engine would be a good avenue for advertising software products and books on the Web.

Yahoo being another popular search engine follows the same trend as Google. It also has good document retrieval capacity and response time. However, its average precision is very low (0.96). Its advert content is the highest in the study. This means that if a searcher is looking for journal subscription or buying books and products online, Yahoo and Google will be the better search engines to use.

MSN does not seem to have good document retrieval efficiency but its relevant documents retrieved as well as precision are encouraging. Alta Vista has good document retrieval efficiency as well as relevant documents retrieved and precision. This means that it is both useful for advertisement and relevant information retrieval purposes.

37.com search engine performs worst in terms of quantities of documents retrieved and response time for the keywords used in this study. However, it has the best relevance and precision. Although this search engine has the least document retrieval efficiency, this may be attributed

to the fact that the main focus of the organization in charge of the search engine may not be in Computer Science profession or that its database might not have been fortified with documents in this domain as at the time this study was done. Actually, we are able to find out practically that most of the few documents retrieved by this search engine is always relevant to our information need. As such, we conclude that 37.com would be a very good search engine for researchers in the nearest future, although we need to verify this fact with information needs of some other professions.

By and large, our results in this study are in line with most of the previous findings in the literatures [9, 10, 13]. In a survey conducted by Enid [8] Google, Yahoo and MSN continue to be the most searched engines with Google taking 48.3%, Yahoo, 27.5% and MSN 10.9% United States (US) search market share in 2007. The main point is that results from this study has demonstrated that the fact that a Web search engine is very popular and able to retrieve large number of documents does not mean that it has high precision for retrieving relevant documents for its users.

A very interesting result in this study is the fact that African contents on the web are nothing to write home about. This means that Africa has not contributed meaningfully to the global web content in the field of Computer Science.

VI. CONCLUSION AND RECOMMENDATIONS

In this study we present an experiment to evaluate the performance of some popular web search engines. We measure the performance of the search engines after examining various numbers of top pages

returned by the search engines and then determine the mean document quantities retrieved, time taken (response time) for the retrieval, documents' relevance, advert content, African content and precision for each search engines. In the experiment we use 15 Computer Science – related queries and look at their performances on five different search engines based on human relevance judgments of users.

Our experiments show that there are differences in the performance of these search engines in the research domain (Computer Science). Summarily, the study shows that:

(i)   Google, MSN and Yahoo search engines have high document quantity retrieval capacity with low response time but their effectiveness (precision) in retrieving relevant documents is very low. They are good avenues for advertisement on the Web;
(ii)   Alta vista search engine is very good for both relevance and advertisement;
(iii)   37.com search engine, although may bring out low document quantity retrieval but most of these documents will be relevant to the information needs of the searcher; and
(iv)   Africa is yet to contribute meaningfully to the global web contents especially in the field of Computer Science.

Knowing the most effective Web search engines satisfying the current information need is important both at personal and business enterprise levels. However, the definition of "the best" changes due to both the changing information needs of users and the changing quality and nature

of search engines. Accordingly, for different information needs from different professionals, the most effective search engine may be different. Hence, search engine performance needs to be tested and should be done quite often. Based on this recommendation, we therefore conclude that web users should vigilantly look for the best search engine on which their information needs will be fully met with less or little response time and distractions. Results obtained in this study cannot be generalized for now, since it is centered around the keywords or queries from Computer Science profession. Further works in some other professions are still needed to be done in order to clarify and establish the claims reported in this study.

Acknowledgements

REFERENCES

1. Abbate, J. (1999) *Inventing the Internet,* Massachusetts: MIT press.
2. Adeyemo, A. B. and Osunade, O. (2002) Examining University of Ibadan ICT infrastructure, In Amos David and Osofisan Nike (Eds.), *Proceedings of Information and communication technology Applied to economic Intelligence,* University of Ibadan, Nigeria, 20 -23August, 2002, pp. 122 – 126.
3. Agbonlahor, Rosemary O. (2005) Factors affecting the adoption of

4. information technology in Nigerian Universities: Implications for human resource development, in Abiodun O Bada (Ed.), *Proceedings of the Eighth International Working Conference of International Federation for Information Processing, (IFIP WG 9.4),* Abuja Nigeria, 26 – 28 May, 2005.

5. Akinola, S. Olalekan, Yara, P. Olatunde and Osunade O. Oluwaseyitanfunmi (2005). Internet Information Seeking Behaviour of Nigerian Students: University of Ibadan as a Case Study, in Abiodun O Bada (Ed.), *Proceedings of the Eighth International Working Conference of International Federation for Information Processing, (IFIP WG 9.4),* Abuja Nigeria, 26 – 28 May, 2005.

6. Bar-llan, J. (2002). Methods for measuring search engine performance over time. *Journal of the American Society for Information Science and Technology, 53(4), 308 – 319.*

7. Berkeley, U. C. (2002). Teaching Library Internet Workshops http://www.lib.berkeley.edu/Teaching Lib/Guides/Internet/ Downloaded in June 2004.

8. Chowdhurry, A. and Soborroff, I. (2002). Automatic evaluation of World Wide Web search Services. In *Proceedings of the ACM SIGIR Conference*, vol.25, pp. 421 – 422.

9. Enid Burns (2007) US search Engine Ranking, The ClickZ Network, May 31 2007, www.clickz.com/showpage.html?page =3626020, Downloaded in March, 2008.

10. Fazli, Can, Rabia, Nuray, Ayisigi, B. Sevdik (2003). Automatic Performance Evaluation of Web Search Engines, *Information Processing and Management*, xxx, xxx – xxx.

11. Gordon, M. and Pathak, P (1999). Finding information on the World Wide Web: the retrieval Effectiveness of search engines, *Information Processing and Management*, 35(2) , 141 – 180.

12. Griesbaum Joachim (2004) Evaluation of three German search engines: Altavista.de, Google.de and Lycos.de, *Information Reearch*, 9(4) paper 189.

13. Hananzita Halim and Kiran Kaur (2006) Malaysian Web Search Engines: A Critical Analysis, *Malaysian Journal of Library and Information Science,* Vol.11, No. 1, pp. 103 -122.

14. Hawking, D., Craswel, N., Bailey, P., and Griffiths, K. (2001). Measuring search engine Quality, *Information Retrieval*, 4(1) 35 – 39.

15. Hong Yu and David Kaufman (2007). A cognitive evaluation of four online search engines for answering definitional questions posed by Physicians, Pacific Symposium on Biocomputing, 12:328-339.

16. Jansen, B. J. and Pooch, U. (2001). A review of Web searching and a framework for future research. *Journal of the American Society for Information Science and Technology,* 52(3), 235 – 246.

17. Jansen, B., 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,* 36(2), 207 – 227.

18. Kobayashi, M. and Takeda, K. (2000). Information retrieval on the Web. *ACM Computing Surveys,* 32(2), 144 – 173.

19. Lawrence, S. and Giles, C.L. (1999). Accessibility of information on the Web. *Nature* 400, 107 –109.

20. Mowshowitz, A. and Kawaguchi, A. (2002). Assessing bias in search engines, *Information Processing and Management,* 35(2), 141 – 156.

21. Salton, G. and McGill, M. (1983). *Introduction to modern information retrieval.* New York: McGraw-Hill.

22. Soboroff, I., Nicholas, C. and Cahan, P. (2001). Ranking retrieval systems without relevance judgments. In *Proceedings of the ACM SIGIR Conference,* vol. 24, pp. 66 – 73.

23. Web Skills (2002). Learn Web Skills, http://www.learnwebskills.com/search/engines.html. Downloaded August 2008.

24. Choo, Chun Wei, Detlor, Brian and Turnbull, Don (1999). Information seeking on the Web – an integrated model of browsing and searching. In Larry Woods (Ed.) *ASIS '99*: *Proceedings of the 62nd ASIS Annual Meeting,* Vol. 36, pp. 3-16.

25. Sullivan, D. (2004). Major search engines and directories. *SearchEngineWatch.com,* http//www..searchenginewatch.com/links/article.php/2156221. Retrieved May, 2009.

AUTHOR'S BIODATA

Olalekan S. Akinola is currently a lecturer of Computer Science at the University of Ibadan, Nigeria. He had his PhD Degree in Software Engineering from the same University in Nigeria. He is currently working on Software Process Improvement models for the Nigeria software industry.