Statistical Evaluation of Web Search Engines using User-based Characteristics

P. I. Ogunyinka[1]; L. I. Aigbogun[1]; A. A. Sodipo

Department of Mathematical Sciences,
Olabisi Onabanjo University, Ago-Iwoye, Nigeria.
E-mail: ogunyinka.peter@gmail.com*

Abstract — **Different web search engines had been rated based on different metrics. However, almost none had considered the search query length, the retrieved quantity, and retrieval time for evaluation of web search engines. This study had rated five web search engines (Google, Yahoo, WOW, AOL and Bing) using non-parametric Kruskal-Wallis test for significant mean difference and single-phase sampling for regression estimation and examination of internal error. The retrieval time was used as the study variable while the retrieved quantity of the organic search results and the search query length were used as the auxiliary variables. The correlation coefficient, mean square error, percentage coefficient of variation and percentage relative efficiency were used for the evaluation and comparison of the estimated population mean of the retrieval time. Results revealed that Google was the most rated web search engine with the highest significant retrieved quantity and significant retrieval time while Bing was the least rated web search engine. It was recommended that web search engine users should use short search query length to obtain significant retrieved quantity at the lowest retrieval time. In general, the results of the percentage relative efficiency was used to rate Google, Yahoo, WOW, AOL, and Bing as the first, second, third, fourth, and fifth best-rated web search engines based on the maximization of the query length and retrieved quantity to obtain minimum retrieval time.**

Keywords: *Web search engines, retrieved quantity, retrieval time, query length, survey statistics.*

I. Introduction

The web search engine is any application software that explores the World Wide Web (W3 model) based on the search terms (query) and displays its results to the user. The internet contains billions of web pages and files which are almost impossible for an individual to manually search for information. These web pages and files can best be accessed through web search engines. Gupta and Sharma (2014) reported that 86% of internet web users use the internet through web search engines while 85% of the users prefer to go through hyperlinks on other web pages. There are many web search engines used by web users. These include Google, Yahoo, AOL, Bing and ASK web search engines. Advantages of web search engines include the ability to locate unique phases, quotations, and information on the web pages and documents on the web. Voluminous search results, many types of search instructions, out-of-date search results, and inconsistency among web search engines are some of the limitations of web search engines.

Sir Tim Berners-Lee (Berners-Lee, 1992) hosted the first list of web server on the CERN web server while Archie in 1990 kick-started the official milestone of web search engines. Table 1 summarizes the milestones of web search engines as documented by Seymour *et al.* (2011). Web search engine is divided into three types namely individual search engine, directory search engine and meta-search engine (Wu and Li, 2004). Individual search engine specializes in either full text or non-full text search. Directory search engine explores the already built hierarchical list or indexes of internet resources on a special sever. However, the meta-search engine, also known as the multi-search engine, explores through the combined individual search engines.

A web search engine, structurally, is a collection of programs. The programs that form the web search engine are web crawler, indexer, and agent or searchers. Figure 1 shows the structural flowchart of a web search engine. The crawler is also known as spider, robot, bot, worm, and ant automatically crawls the most popular pages on the web-based on the search terms. The crawler then follows all the links sighted on the popular pages by following this pattern from level one (most popular pages) to the advanced level through all the sighted links. Consequently to this, the crawling system would have formed a spider that had traveled around the WWW within the shortest time.

Crawlers have different crawling methods which determine their speed. Some of the crawlers may use webpage header, title, and meta keywords (Khriste *et al.*, 2011).

Table 1: History of Web search engines

SN	Web search engine	Founded year	SN	Web search engine	Founded year
1	Archier	1990	12	Google	1998
2	Gopher	1991	13	Teoma, Vivisimo	1999 - 2000
3	Veronica and Jughead	1991	14	Yahoo	2004
4	Webcatalog and Wanderer	1993	15	MSN and GoodSearch	2005
5	Aliweb	1993	16	Wikiseek, Guruji, Sproose and Blackle	2006 - 2007
6	Jump Station	1993	17	Powerset, Picollator, Viewzi	2008
7	Webcrawler	1994	18	Cuil, LeapFish, Valdo	2008
8	Matacrawler	1994	19	Bing	2009
9	Alta Vista	1995	20	Sperse, Yebol, Goby	2009 - 2010
10	Excite	1995	21	Exalead	2011
11	Dogpile, Inktomi and HotBot	1996			

The indexer receives from the crawler. Indexer categories the pages, remove duplicates, and organizes the pages into a structured system. The indexer combines the structure indexes of the crawled pages with the existing pages on the database and saves the combined results on the database, either in reproducing or cumulative order. Factors that determine the speed or effectiveness of an indexer are web coverage (data sizes), up-to-dateness of search engine databases, WWW contents, the invisible web (webpages that are pass-worded), and spam (junk pages) (Lewandowski, 2005).

The indexer ranks the crawled web pages using the query-dependent and query-independent factors. The query-dependent ranking factors include word document frequency, search term distance, search term order, meta-tags, anchor text, language and Geo-targeting. The query-independent factors include link popularity, click popularity, up-to-dateness, document length, file format and website size (Lewandowski, 2005).

Table 2: Overview of evaluation metrics for web search engines

SN	Category	Metrics
1	Recall/Precision and their direct descendants	Precision, Recall, Recall-Precision graph, F-measure, F_i-Measure and Mean Average Precision (MAP)
2	Other system-based metrics	Reciprocal Rank (RR), Mean Reciprocal Rank (MRR), Quality of Result Ranking (QRR), Bpref, Sliding Ratio (SR), Cumulative Gain (CG), Discounted Cumulative Gain (DCG) and Average Distance Measure (ADM)
3	User-based metrics	Expected Search Length (ESL), Expected Search Time (EST), Maximal Marginal Relevance (MMR), α-Normalized Discounted Cumulative Gain (α-NDCG) and Expected Reciprocal Rank (ERR)

The agent retrieves and sorts the data from the database based on the search query. The agent determines the structure of the Search Engine Result Pages (SERPs). Babayigit and Groppe (2018) reported the three classifications of SERPs as advertising for proportional purposes, latest news, and organic search results. The advertising for proportional purposes are pages that have been advertised with the search engine host. Organic search results are the ranked natural search results. Only the organic search results can be influenced in their ranking using the power of Search Engine Optimization (SEO). Figure 2 shows the SERP of Yahoo search engine as at June 30, 2020.

Information Retrieval (IR) and electronic mail (e-mail) services are two primary services rendered by the internet (Wu and Li, 2004). Information is retrieved from the internet through the web search engine. Google, Yahoo, AOL, Bing, AlterVista, AlltheWeb, ASK among others are some of the web search engines available in the last decade. However, since there are many web search engines that render this service, there is a need to always compare the efficiency of these available web search engines. There are three major metrics for evaluating web search engines. These metrics are Recall/Precision metric, system-based metric, and user-based metric. Table 2 summarizes these three categories of evaluation metrics, as presented by Sirotkin (2012).

Many literature have evaluated search engines using user-based metrics. Spink and Jansen (2004) evaluated two web search engines (AlterVista and AlltheWeb) using the attitude in searching names on the search attitude (user-based metrics). It was concluded that personal name search was a common attitude but it was not the major activity on the web search engine. Wu and Li (2004) compared the efficiency of four web search engines (Google, AlltheWeb, Hotbot, and AltaVista) and four meta-search engines (MataCrawler, ProFusion, MetaFind and MetaEUREKA) using the relevance and average precision of the harvested document metrics. It was concluded that both search engine and meta-search engine equally perform better with short queries than long queries which contradict the existing claim that meta-search engines performed better than web search engines. Jansen and Molina (2006) examined the effectiveness of five search engines (Excite, Google, Overture, Froogle and Yahoo) in retrieving relevant e-commerce links using a one-way ANOVA statistical method for evaluation. It was concluded that web links from e-commerce search engines are significantly more relevant than web links from web directory domains.

Lewandowski (2008) compared four search engines (Google, Yahoo, MSN and ASK) based on language restriction features on these engines. Results revealed that Google and MSN were not recommended for foreign language search while Yahoo and ASK were recommended for foreign language document search. Jain (2013) highlighted the on-page and off-page factors as the metrics used by Google, Yahoo, Bing, MSN, AltaVista and ASK. It was concluded that Google used the most effective search algorithm on the World Wide Web (W3 model). Dwivedi and Yadav (2018) compared five search engines (Google, Yahoo, AOL, Bing and ASK) to obtain the in-depth coverage, thought and clarity and number of related web links corresponding to thirty-two search keywords. ANOVA test revealed that the five search engines were not significantly different with respect to the three metrics considered. Google Trends (in figure 3) shows the comparison of Google, Yahoo, WOW, AOL and Bing search engines between December 6 and 15, 2014 (the same date range used in this study). Google, Yahoo and Bing were rated as the first, second and third most searched engines on the web, based on the Google server.

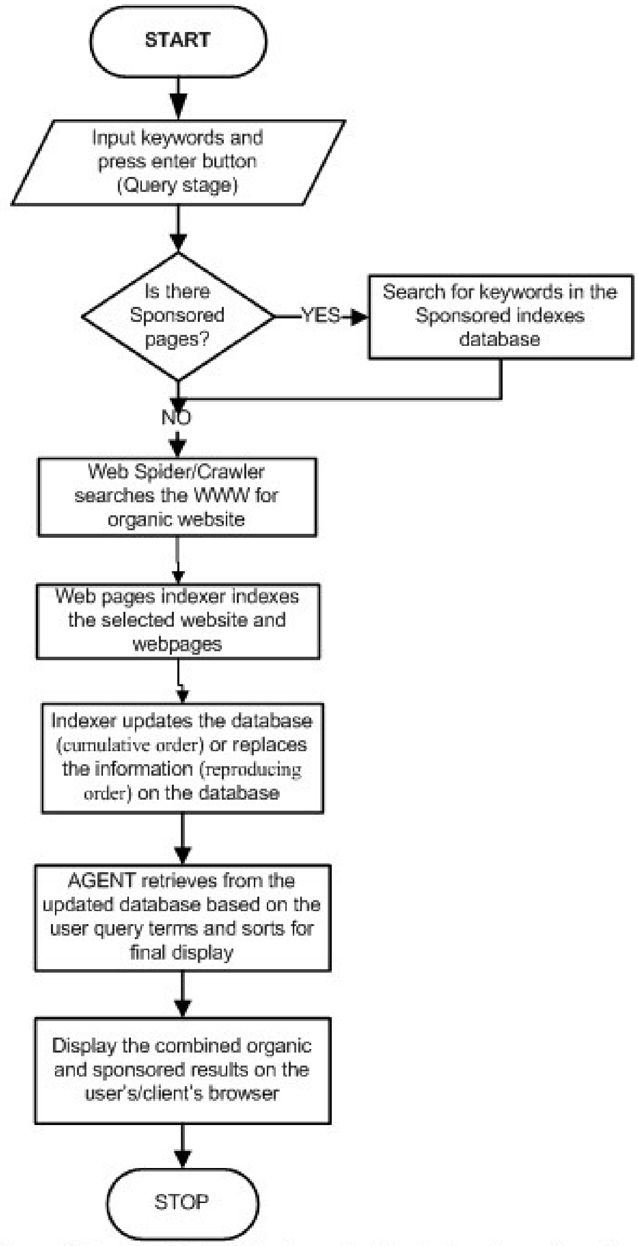

Figure 1: Flowchart of the Spider-based or Crawler-based search engine

Sequel to the reviewed literature, no literature had considered the retrieved quantity, retrieval time, and the length of search queries for the evaluation of web search engines. This study focuses on the comparison of five selected web search engines by comparing the means of the user-based metrics and comparing the estimated internal error of the retrieval time.

II. Methodology

This study had compared five web search engines (Google, Yahoo, WOW, AOL and Bing) using data collected about the user-based characteristics such as Retrieved Quantity (RQ) of the organic links, the Retrieval Time (RT) and length of the search query (query length) as the metrics of comparison. The data collection process selected 200 science-related search queries. Some of the search queries included analysis, calcium, fermentation, heat, nutrition, petroleum, solidity, urology, antenna, doctor, ferrous, irritation and lithium. Each of the queries was entered into the search textbox of each of the five web search engines and the corresponding retrieved quantity and retrieval time were obtained. Celeron(R) Dual-Core, 3.00GB, 2.10GHz Laptop technology was used to collect the data from the five web search engines between December 6 to 15, 2014, inclusively. Google automatically displayed the retrieved quantity and retrieval time while Yahoo, WOW, Bing and AOL only displayed the retrieved quantity automatically. However, stopped watch was used to manually obtain the retrieval time while the length function in the Microsoft Excel application was applied to the queries to obtain the query length.

SPSS version 23 and Ms. Excel version 2010 were used for the statistical analyses. Preliminary tests revealed that both the retrieved quantity and retrieval time distribution violated the normality and homogeneity assumptions at $p<0.05$. This violation could be associated with the presence of extreme values (outliers) in the datasets. All data transformations were not appropriate for the correction of these violations. Hence, the Kruskal-Wallis test (non-parametric tests) was used to test if the mean for each of retrieved quantities and retrieval time, from the five search engines, were significantly different at 5% significant level. The null hypotheses were set thus:

H_{10}: There is no statistically significant difference among the retrieved quantities for the fives web search engines.

H_{20}: There is no statistically significant difference among the retrieval times for the fives web search engines.

The post-hoc test was used to confirm which search engines were significantly different. Similarly, single-phase regression estimator sampling in survey statistics was used to estimate the population means of the retrieval time and the associated mean square error. Furthermore, the associated percentage coefficient of variation and percentage relative efficiency was used for comparison of the five web search engines.

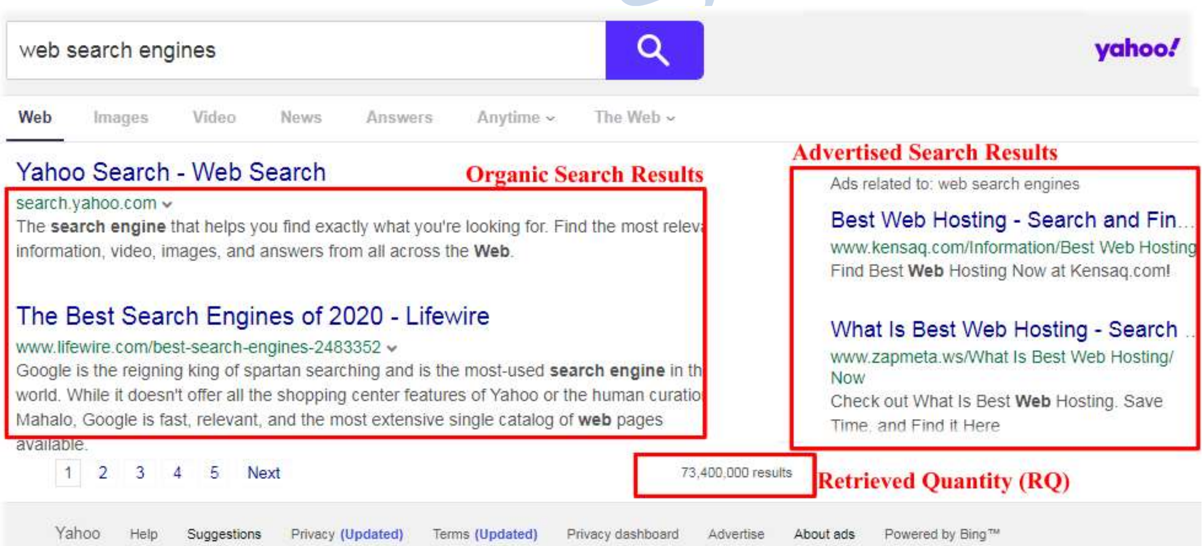

Figure 2: The organic and advertised search results

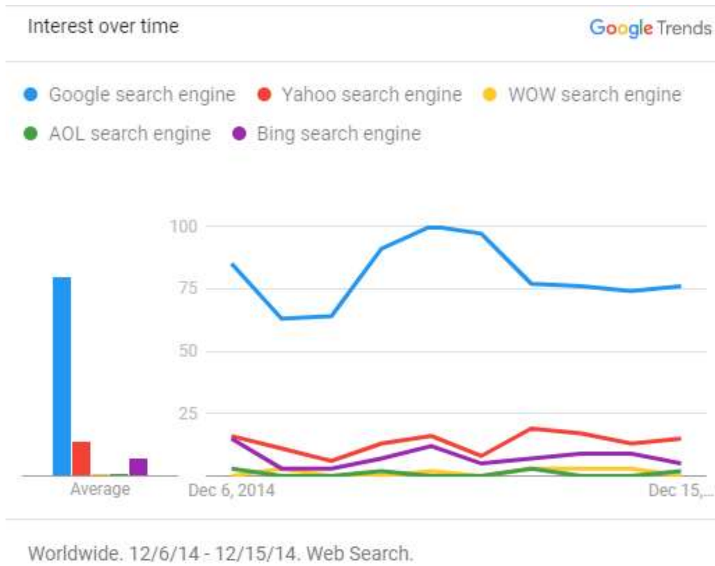

Figure 3: Google Trends analysis on the five considered web search engines between December 6 and 15, 2014

2.1 Review of Kruskal-Wallis test

The Kruskal-Wallis test statistic is giving as

$$H = (N-1)\frac{\sum_{i=1}^{g} n_i(\bar{r}_i - \bar{r})^2}{\sum_{i=1}^{g}\sum_{j=1}^{n_i}(r_{ij} - \bar{r})^2} \quad (1)$$

Where n_i is the number of observation in group i. r_i is the rank (among all observations) of observation j from group i, N is the total number of observations across all groups, $\bar{r}_i = \left(\sum_{j=1}^{n_i} r_{ij}/n_j\right)$ is the average rank of all observations in group i and $\bar{r} = \frac{(N+1)}{2}$ is the average of all the r_{ij}. If the data contain no ties the denominator of the equation for H is exactly $\frac{(N-1)(N+1)}{12}$ and $\bar{r} = \frac{(N+1)}{2}$. Hence,

$$H = \frac{12}{N(N+1)}\sum_{i=1}^{g} n_i\left(\bar{r}_{i.} - \frac{(N+1)}{2}\right)^2. \quad (2)$$

2.2 Review of Regression Estimator with extreme value correction factor using two auxiliary variables in Survey Statistics

Al-Hossain and Khan (2014) had developed improved regression estimator in survey statistics. This estimator uses the method of Sarndal (1972) correction factor for the correction of the extreme (outliers) values in the survey data using two auxiliary variables. The estimator is presented as

$$\hat{\bar{Y}}_{dl} = \bar{y}_{dl} = \bar{y}_{c11} + b_1(\bar{X}_1 - \bar{x}_{1c21}) + b_2(\bar{X}_2 - \bar{x}_{2c31}). \quad (3)$$

Where $\bar{y}_{c11} = (\bar{y} \pm C_1)$; $\bar{x}_{1c21} = (\bar{x}_1 \pm C_2)$ and $\bar{x}_{2c31} = (\bar{x}_2 \pm C_3)$. $\hat{\bar{Y}}_{dl}$ is the estimated population mean, $\bar{y}$ is the sample mean of the study variable (y), $\bar{x}_1$ is the sample mean of the first auxiliary variable, $\bar{x}_2$ is the sample mean of the second auxiliary variable, $\bar{X}_1$ is the population mean of the first auxiliary variable, $\bar{X}_2$ is the population mean of the second auxiliary variable, b_1 is the regression coefficient of y on $\bar{x}_1$ and b_2 is the regression coefficient of y on $\bar{x}_2$. The optimum value of C_1, C_2 and C_3 are obtained as $C_{1opt} = \frac{(y_{max} - y_{min})}{2n} = \frac{\Delta_y}{2n}$, $C_{2opt} = \frac{(x_{1max} - x_{1min})}{2n} = \frac{\Delta_{x_1}}{2n}$, and $C_{3opt} = \frac{(x_{2max} - x_{2min})}{2n} = \frac{\Delta_{x_2}}{2n}$.

The corresponding estimated Mean Square Error ($\widehat{MSE}$) of $\hat{\bar{Y}}_{dl}$ is presented as

$$\widehat{MSE}_{min}\left(\hat{\bar{Y}}_{dl}\right) = \theta S_y^2[1 - \hat{\rho}_{yx_1}^2 - \hat{\rho}_{yx_2}^2 + 2\hat{\rho}_{yx_1}\hat{\rho}_{yx_1}\hat{\rho}_{x_1x_2}] - \frac{\theta}{2(N-1)}[\Delta_y - \hat{b}_1\Delta_{x_1} - \hat{b}_2\Delta_{x_2}]^2, \quad (4)$$

where $\theta = \left(\frac{1}{n} - \frac{1}{N}\right)$, $\hat{\rho}_{yx_1}$ is the estimated population correlation coefficient between y and x_1, $\hat{\rho}_{y2}$ is the estimated population correlation coefficient between y and x_2, $\hat{\rho}_{x_1x_2}$ is the estimated population correlation coefficient between x_1 and x_2, N is the population size, $\hat{b}_1$ is the estimated b_1 and $\hat{b}_2$ is the estimated b_2.

The Percentage Coefficient of Variation (PCV) of $\hat{\bar{Y}}_{dl}$ is presented as

$$PCV\left(\hat{\bar{Y}}_{dl}\right) = \frac{\widehat{MSE}_{min}\left(\hat{\bar{Y}}_{dl}\right)}{\hat{\bar{Y}}_{dl}} * 100\%. \qquad (5)$$

Similarly, the Percentage Relative Efficiency (PRE) of the estimator $\hat{\bar{Y}}_{dl}$ is presented as

$$PRE(1/2) = \frac{\widehat{MSE}_{min}\left(\hat{\bar{Y}}_{dl2}\right)}{\widehat{MSE}_{min}\left(\hat{\bar{Y}}_{dl1}\right)} *$$

The result of the Kruskal-Wallis test (at 5% significant level) and post-hoc test are presented in tables 3 and 4, respectively. Similarly, the correlation coefficient analysis result is presented in table 5, the estimated Mean Square Error ($\widehat{MSE}$) and the Percentage Coefficient of Variation (PCV) results are presented in table 6 while the percentage relative efficiency result is presented in table 7.

III. RESULTS

Table 3: Kruskal-Wallis test results

		N	Mean	Std. Error	*P-value*	*Decision*
Retrieved Quantity	Google	200	301494170.00	1.162E8	0.000*	Reject the Null hypothesis
	Yahoo	200	19649345.00	4818354.391		
	WOW	200	32983846.37	12292719.423		
	AOL	200	240015580.00	36741866.192		
	Bing	200	38764840.00	11458519.905		
	Total	1000	126581556.27	24865059.191		
Retrieval Time	Google	200	0.2905	.00630	0.000*	Reject the Null hypothesis
	Yahoo	200	0.2518	.00826		
	WOW	200	0.2816	.00873		
	AOL	200	0.3706	.00935		
	Bing	200	0.4574	.16885		
	Total	1000	0.3304	.03394		

*: There exists statistical significant difference among the web search engines.

Table 4: Post-Hoc test of mean comparison for Retrieved Quantity (RQ) and Retrieval Time (RT)

Retrieved Quantity (RQ)			Retrieval Time (RT)		
WSE1-WSE2	Adj. Sig.	Decision	WSE1-WSE2	Adj. Sig.	Decision
WOW - Yahoo	1.000	WOW ≅Yahoo	Yahoo - WOW	0.118	Yahoo ≅ WOW
WOW - Bing	0.183	WOW ≅Bing	Yahoo - Bing	0.055	Yahoo ≅Bing
WOW - Google	0.000*	WOW < Google	Yahoo - Google	0.004*	Yahoo < Google
WOW - AOL	0.000*	WOW < AOL	Yahoo - AOL	0.000*	Yahoo < AOL
Yahoo - Bing	1.000	Yahoo ≅Bing	WOW - Bing	1.000	WOW ≅Bing
Yahoo - Google	0.000*	Yahoo < Google	WOW - Google	1.000	WOW ≅ Google
Yahoo - AOL	0.000*	Yahoo < AOL	WOW - AOL	0.000*	WOW < AOL
Bing - Google	0.000*	Bing< Google	Bing - Google	1.000	Bing≅ Google
Bing - AOL	0.000*	Bing< AOL	Bing - AOL	0.000*	Bing> AOL
Google - AOL	1.000	Google ≅ AOL	Google - AOL	0.000*	Google < AOL

*: There exists statistical significant difference between the two web search engines.

Table 5: Correlation Matrix of Variables (Quantity/Time) with respect to the query length

SN	Search Engine	QUANTITY	TIME
1	Google	Positive	Negative
2	Yahoo	Negative	Negative
3	WOW	Negative	Negative
4	AOL	Negative	Negative
5	Bing	Negative	Negative

Table 6: Computation of the estimated population mean, $\widehat{MSE}$ and the PCV

Web Search Engine	$\widehat{\bar{Y}}_{dl}$	$\widehat{MSE}$	PCV (%)
Google	0.2926	0.000046	2.3171
Yahoo	0.2533	0.000067	3.2269
WOW	0.2880	0.000071	2.9291
AOL	0.3663	0.000098	2.7071
Bing	0.5124	0.044662	41.2436

Table 7: Computation of the Percentage Relative Efficiency (PRE)

	Google	Yahoo	WOW	AOL	Bing
Google	100.00%	145.35%	154.82%	213.80%	97137.26%
Yahoo		100.00%	106.52%	147.10%	66831.83%
WOW			100.00%	138.10%	62743.53%
AOL				100.00%	45433.80%
Bing					100.00%

IV. Discussion

Kruskal-Wallis test results revealed that the distributions of the Retrieved Quantity (RQ) were statistically significantly different for the five web search engines at $p < 0.05$ (see table 3). Similarly, the Retrieval Time (RT) for the five web search engines, were statistically and significantly different at $p < 0.05$ (see table 3). Hence, there was need to investigate to ascertain which pair of the search engines differ from each other for both the RQ and the RT.

Table 4 shows the post-hoc test results. Google had significant high RQ mean over WOW, Yahoo and Bing at $p < 0.05$. Similarly, AOL retrieves significantly high average RQ over WOW, Yahoo and Bing at $p < 0.05$ while the RQs for (Google and AOL), (WOW and Yahoo), (WOW and Bing) and (Yahoo and Bing) were significantly the same for each pair at 5% significant level. This implies that Google and AOL search engines had significantly highest RQs as against the other three search engines while Yahoo, Bing and WOW had significantly the same RQs. Hence, Google and AOL were significantly better than Yahoo, WOW and Bing with respect to RQ results.

Similarly on table 4, the post-hoc test results revealed that AOL had significant high RT over Yahoo, WOW and Google at 5% significant level while Bing had high RT mean over AOL at 5% significant level. Yahoo and WOW had the significant least RT over the remaining three search engines at 5% significant level. This implies that Yahoo and WOW were significantly faster than Google and AOL while Bing is the least fast search engine. However, Google is significantly faster than AOL search engine.

The Correlation Coefficient (CC) analysis results are presented in table 5. It was revealed that CC between RQ and search Query Length (QL) for Google is positive while CC results for Yahoo, WOW, AOL and Bing are negative. This implies that in Google web search engine, the higher the search QL the higher the RQ results and vice versa. On the contrary, for Yahoo, WOW, AOL and Bing, the higher the search QL, the lower the RQ results and vice versa. Google users are advised to increase the QL in order to get high RQ or vice versa. However, users of Yahoo, WOW, AOL and Bing search engines are advised to reduce the QL in order to get more RQs or vice versa. However, the negative CC analysis results for RT with search QL implies that the for the five web search engines, the higher the search QL, the lower the RT and vice versa. This means that the five considered web search engine users are advised to increase their search QL as much as possible in order to reduce the retrieval time of any of the five web search engines.

Equations (3), (4) and (5) were, respectively, used to estimate the populations mean, estimated Mean Square Error ($\widehat{MSE}$) and the Percentage Coefficient of Variation (PCV) of the RT for each of the five web search engines. The estimated population mean (in equation 3) has accounted for the correction of the extreme values (outlier) in the distribution. It was assumed that RT depends on the RQ and the QL and the sample size (n) is 150. Hence, RT was used as the study variable while the RQ and the QL were used as the auxiliary variables.

Since the PCV accounts for both the MSE and $\hat{\bar{Y}}_{dl}$, the PCV was used for decision making. Google (with PCV of 2.32%) is rated the best search engine among the five search engines. AOL, WOW, Yahoo and Bing were rated second, third, fourth and fifth web search engines, respectively. This implies that even when Yahoo and WOW were significantly faster than Google (see table 4), yet Google proves to maximize the RT (with the least PCV) than the remaining four web search engines. PCV results proves that Bing is not only the least fast web search engine but it, also, spends too much time in accomplishing this. Perhaps, we conclude that Bing web search engine is the least coordinated search engine among the five considered web search engines.

Finally, table 7 shows the results for the Percentage Relative Efficiency (PRE) of one web search engine over the other. PRE measures the quantitative efficiency of one web search engine over the other. It is revealed that Google is 145%, 155%. 214% and 97137% efficient over Yahoo, WOW, AOL and Bing, respectively. Similarly, Yahoo is 107%, 147% and 66832% relative efficient over WOW, AOL and Bing, respectively. WOW is 138% and 62744% relatively efficient over AOL and Bing, respectively. Finally, AOL is 45434% relatively efficient over Bing. These results, finally, affirms the rating of the five web search engines based on the maximization of the QL and RQ to attain minimum RT. Google. Yahoo, WOW, AOL and Bing are rated as the first, second, third, fourth and fifth respective best rated web search engines based on the aforementioned condition.

V. Conclusion

This study concentrates on the rating of five web search engines (Google, Yahoo, WOW, AOL and Bing). The results had revealed that, with respect to Retrieved Quantity (RQ), Google and AOL retrieve the highest amount of web pages on its database. Perhaps, Google and AOL had the best web crawler/spider that updates the database most often. However, with respect to Retrieval Time (RT), Yahoo and WOW had the smallest search time over the remaining three web search engines.

It was, also, concluded that web search engine users should reduce their search query length in order to obtain large RQ from Yahoo, WOW, AOL and Bing while Google users are advised to increase the QL in order to obtain high RQ and vice versa.

It was concluded that Google is the best-rated web search engine that harvests the maximum RQ in the shortest RT considering the QL while Bing is the least-rated web search engine that harvests the minimum RQ in the highest RT considering the QL.

References

Al-Hossain Y. Abdullah and Khan Mursala (2014). Efficiency of Ratio, Product and Regression Estimators under Maximum and Minimum Values using Two Auxiliary Variables. Journal of Applied Mathematics, Article ID 693782, doi:10,1155/2014/693782, 1-6.

Babayigit Sinan and Groppe Sren (2018). Webpage Ranking Analysis of various search engines with special focus on Country-specific search. Open Journal of Web Technologies (OJWT), 5(1), 44-47.

Dwivedi Nripendra and Yadav Monoj (2018). Web Search Engines (Google, Bing, Yahoo, ASK and AOL) Analysis for their search Result. International Journal of Innovations and Advancement in Computer Science, 7(3), 102-107.

Gupta Trilok and Sharma Archana (2014). Anatomy of Web Search Engines. International Journal of Advanced Research in Computer and Communication Engineering, 3(4), 6244-6249.

Jain Ayush (2013). The Role and Importance of Search Engine and Search Engine Optimization, 2(3), 99-102.

Jansen J. Bernard and Molina R. Paulo (2006). The effectiveness of Web search engines for retrieving relevant e-commerce links. Information Processing and Management, 42, 1075-1098, doi:10.1016/j.ipm.2005.09.003.

Khriste P. Gajanan, Veer K. Dharamraj and Meske D. B. (2011). Role of Search Engines in Library at a Glance. International Journal of Information Dissemination and Technology, 1(2), 86-91.

Lewandowski Dirk (2005). Web searching, search engines and Information Retrieval. Information Services and Use, 25, 137-147.

Lewandowski Dirk (2008). Problems with the use of Web search engines to find results in foreign languages. Online Information Review, 32, 1-5.

Sarndal C. E. (1972). Sample Survey theory vs. general statistical theory: Estimation of the population mean, International Statistical Institute, 40, 1-12.

Seymour Tom, Frantsvog Dean and Kumar Satheesh (2011). History of Search Engines. International Journal of Management and Information Systems, 15(4), 47-58.

Sirotkin Pavel Vorgelegt Von (2012). On Search Engine Evaluation Metrics. Unpublished Ph.D. Dissertation, Heinrich-Heine-Universitat, 1-193.

Spink Amanda and Jansen J. Bernard (2004). Searching for people on Web search engines. Journal of Documentation, 60(3), 266-278, doi:10.1108/00220410410534176.

Berners-Lee T, Cailliau R., Groff J-F & Pollermann B. (1992). World-wide web: The Information Universe. Electronic Networking Research, Applications and policy. 2(1): 52-58.

Wu Shengli and Li Jiegu (2004). Effectiveness Evaluartion and Comparison of Web Search Engines and Meta-Search Engines. Springer Lecture Notes in Computer Science, 3129, 303-314.