A ANALYSIS OF GOOGLE, YAHOO AND BING WEB SEARCH ENGINE IN PERFORMANCE PARAMETER OF PRECISION AND RELATIVE RECALL

Kamlesh Kumar Pandey¹, Vivek Prakash Sahu², Rajat Kumar Yadu³

¹UTD Amarkantak, MCRPVV Amarkantak Campus, (India) ^{2,3}Department of Computer Science, Mahant Laxmi Narayan Das College Raipur, (India)

ABSTRACT

Search Engine is a special application of internet and most popular in this word. Search Engine are used by everybody because all importance subject related information, data, news is given by Internet through Search Engine. If any internet user can be wanted to any type of information in any time he searches to search engine. Search engine given the searched item related result in from list of link and internet user selected the one link. Search engine given to result List and result are include some result list is related to search item and some result list are not related for search item or Query. In This paper cover to retrieval effectiveness of performance parameter of precision and relative recall of Google, Yahoo and Bing search engine. The searching keyword or queries are divided into three categories. These Categories are a simple one word, simple multi word and complex multi word group and we taken on each group three searching keyword. Main aim of this research paper is find out which web search engine is more relevant.

Keyword: Web Search Engines, Precision, Relative recall, Google, Yahoo, Bing.

I. INTRODUCTION

Web Search engine is a web application program they run on the specific web address and this web address are called to website. Internet User given a searching keyword or queries on the specific web search engine and web search engine are given the dynamically result list and this result are related to searching keyword or query. Every web search engine is using to own database for creating to result list. If any internet user is needed to any type of information at any time he open to search engine and searches this information. In present time many searching technique are available in web search engine and different type of web search engine used to own searching technique. Web Search engines search to all information and all things in World Wide Web like a files, songs, videos, images, web sites, weather information through various interfaces means through web search engine. The National Information Standards society defines to boundary and access management is web search engine. The basically Web search engine are groups in two categories first categories belong to Computer-generated indexes and second categories belong to a directories. Indexes base search engine are used to special search program this program is called to Spiders or robots and it have a large database they able to create a dynamically listing and Directories classify to web documents into a subject classification, yellow pages scheme like a Arts, Business, Computers and Internet, Entertainment, Government they are usually compiled by some type of logical order and

International Journal of Advanced Technology in Engineering and Science Vol. No.4, Special Issue No. 01, February 2016 www.ijates.com

smaller database as a compared to indexes. Google and Bing are most popular Index web search engines and Yahoo is most popular directories web search engines.

A.Google Search Engine:-

Google is an American multinational technology corporation it given to Internet related services and include online advertise technologies, searching, android operating system, Mail, cloud computing, and software. Google is come to internet word in September 4, 1998.

B. Yahoo Search Engine:-

Yahoo is an American international technology business company and it come to January 1994 in internet word. Yahoo product and service is related to internet and this service are a Yahoo News, Mail, Finance, Sports, Search, Messenger, Answer, Flickr, online mapping, video sharing etc.

C.Bing Search Engine:-

Bing is a product of Microsoft and launched on May 28, 2009 .Bing is also known as Live Search, Windows Live Search and MSN Search. Product and service of Bing is Webmaster services, Mobile services, Developer services, searching, Toolbars, Gadgets etc.

The internet user some time cannot get the appropriate result falsity because millions of information is available in internet. Some information is relevant and some information irrelevant is coming. To find the closely information among the huge result is difficult for internet user and expert IT professional. In this study we analysis which search engine are given to how number of relevant and irrelevant document we also analysis to precision and relative recall of Google, Yahoo and Bing.

II. FACTOR OF PERFORMANCE EVALUATING

Performance evaluation of Google, Yahoo and Bing are examined to during November 2015 to January 2016 means we calculation the precision and relative recall for some selected search keyword between November 2015 to January 2016. In this study, Google, Yahoo and Bing are given to own search results and this search result are categorized as five points. These points are first is more relevant, second is less relevant, third is irrelevant, four is links and five point is sites can't be accessed on the basis of the following criteria and this criteria and points are given in the Chu & Rosenthal 1996[1], Leighton 1996[2], Ding & Marchionini 1996[3], Clarke & Willett 1997[4], B.T. Sampath kumar[6]. We define new criteria on the bases on old criteria and calculation on precision and relative recall of present time search. These new criteria are identify the above five point and this criteria are.

- If the web page content is closely matched to subject matter of the search keyword then this type of web document is classify to more relevant and we given to 2 number/score.
- If the web page content is not closely matched to subject topic but consists of some related aspects to the subject topic of the search keyword then this web page is classify to less relevant and we given to 1 number/score.
- If the web page content is not related to the subject topic of the search keyword then this web page is classify as irrelevant and we given to 0 number/score.
- If the web page content is consisted of a complete series of links but some information is required then this web document is classify as link and we given to 0.5 number/score.

International Journal of Advanced Technology in Engineering and Science Vol. No.4, Special Issue No. 01, February 2016 www.ijates.com

• If the web site or web document is can't be accessed or open for a particular URL then its web page is classify to site can't be accessed and we given 0 number/score.

III. PRECISION CALCULATION OF GOOGLE, YAHOO AND BING

In this section we calculate to precision of search engine using to five point and also we analysis how many result link or web page are repeated in many time. We list out how many time(in secoand) is taken to given result in broad Band Connection and also calculate the how many parented of relevant sites, Irrelevant sites, links and Sites cannot be accessed.

First factor of performance is precision so this section we calculation on precision of search engines for each of the search keyword using this formula and used to five criteria (Eq. 1).

Precision is a ratio of the number of relevant records retrieved by search engine to the total number of irrelevant and relevant records retrieved by search engine.

Precision=Sum of the scores (number) of sites retrieved by a Search Engine / Total number of sites retrieved (Eq. 1)

a. Precision of Google

Total numbers of 5,29,68,00,000 sites are founded for different nine keyword and we selected to 900 sites for precision calculated. Table 1 showed that 34.78% of the sites retrieved by Google were less relevant followed by links 8.56% and irrelevant sites 15.44%. It was also observed that 33% sites were more relevant and only a small percentage of the sites 4.22% can't be accessed. The precision of the Google was calculated using the above formula. The overall precision of the Google was 1.09.

b. Precision of yahoo

Total numbers of 4,968,400,000 sites are founded for different nine keyword and we selected to 900 sites for precision calculated. Table 2 showed that 37.78% of the sites retrieved by Yahoo were less relevant followed by links 6.22% and irrelevant sites 13.11%. It was also observed that 38.56% sites were more relevant and only a small percentage of the sites 4.33% can't be accessed. The precision of the Yahoo was calculated using the above formula. The overall precision of the Yahoo was 1.18.

c. Precision of Bing

Total numbers of 4,344,200,000 sites are founded for different nine keyword and we selected to 900 sites for precision calculated. Table 3 showed that 36.78% of the sites retrieved by Bing were less relevant followed by links 7.78% and irrelevant sites 12.44%. It was also observed that 38.78% sites were more relevant and only a small percentage of the sites 3.89 % can't be accessed. The precision of the Bing was calculated using the above formula. The overall precision of the Bing was 1.18.

International Journal of Advanced Technology in Engineering and Science Vol. No.4, Special Issue No. 01, February 2016 www.ijates.com



			TABLE	1- Precision	calculation	n of Google				
Search	Total number	Select	More	Less	Irrelevan	links	Sites	Repeate	Search	Prec
keywor	of sites	ed	relevant	relevant	t sites		cannot	d	ing	isio
d		sites	sites	sites			be	link	Time	n
							accesse			
							d			
				Simple one	word querie	s				
Comput	2,30,00,00,000	100	43	23	16	12	6	7	0.40 se	1.15
er										
Databas	99,80,00,000	100	39	36	18	6	1	4	0.58 se	1.17
e										
Multim	64,70,00,000	100	36	32	14	9	9	4	0.52 se	1.09
edia										
				Simple mult	i word queri	es				
What is	35,90,00,000	100	32	41	18	5	4	2	0.60 se	1.07
search										5
engine										
Comput	28,90,00,000	100	45	39	12	3	1	1	0.39 se	1.44
er										
science										
Digital	43,30,00,000	100	31	46	13	6	4	2	0.53 se	1.38
India										
Complex	multi word queri	es								
Internet	3,23,00,000	100	26	42	19	11	2	9	0.52 se	0.99
and web										
designin										
g										
Evaluati	1,45,00,000	100	23	46	14	13	4	8	0.52 se	0.98
on of										
digital										
library										
Comput	22,40,00,000	100	22	44	15	12	7	8	0.51 se	0.94
er										
science										
&										
enginee										
ring										
Total	5,29,68,00,000	900	297	349	139	77	38	33	1	1.09
			(33 %)	(38.78%)	(15.44%)	(8.56%)	(4.22%)	(3.67%)		
	1	<u> </u>	I	1	<u> </u>		<u> </u>	132	Daga	1

g

International Journal of Advanced Technology in Engineering and Science

Vol. No.4, Special Issue No. 01, February 2016

www.ijates.com



Search	lotal	Selected	More	Less	Irrelevant	links	Sites	Repeated	Searching	precision
keyword	number of	sites	relevant	relevant	sites		cannot	link	Time	
	sites		sites	sites			be			
							accessed			
				Simple or	1e word quer	ies				
Computer	539,000,000	100	42	44	9	3	2	1	0.38 se	1.29
Database	436,000,000	100	36	36	13	8	7	4	0.55 se	1.12
Multimedia	89,800,000	100	35	29	17	9	10	4	0.51 se	1.03
				Simple mu	ılti word que	ries				
What is	3,740,000,000	100	29	36	14	11	10	7	0.54 se	0.99
search										
engine										
Computer	16,300,000	100	36	48	8	7	1	5	0.38 se	1.23
science										
Digital	12,700,000	100	32	38	17	9	4	3	0.58 se	1.06
India										
				Complex m	ulti word qu	eries				
Internet	109,000,000	100	42	36	17	3	2	0	0.54 se	1.21
and web										
designing										
Evaluation	10,300,000	100	46	36	12	4	2	0	0.55 se	1.3
of digital										
library										
Computer	16,300,000	100	49	37	11	2	1	0	0.52 se	1.36
science &										
engineering										
Total	4,968,400,000	900	347	340	118	56	39	24		1.18
			(38.56%)	(37.78%)	(13.11%)	(6.22%)	(4.33%)	(2.67)		
										L

TABLE 2- Precision calculation of Yahoo

TABLE 3- Precision calculation of Bing

Searc	Total number	Sel	More	Less	Irrelevan	links	Sites	Repeated	Searchin	Precisi
h	of sites	ect	relevant	relevant	t sites		cannot	link	g Time	on
keywo		ed	sites	sites			be			
rd		site					accesse			
		s					d			
	Simple one word queries									
Comp	5,370,00,000	100	39	36	13	11	1	3	0.49 se	1.19
uter										
Multi	539,00,000	100	41	37	13	7	2	2	0.36 se	1.22
media										
Multi	458,000,000	100	38	34	12	6	10	4	0.52 se	1.13
media										
Simple multi word queries										
Social	308,00,000	100	37	39	13	7	4	4	0.35 se	1.06
<u>L</u>	1	1	1	1	1		1	1		1

133 | P a g e

International Journal of Advanced Technology in Engineering and Science

Vol. No.4, Special Issue No. 01, February 2016

ijates ISSN 2348 - 7550

www.ijates.com

Marke										
ting										
Comp	167,00,000	100	39	43	12	3	3	3	0.35 se	1.22
uter										
scienc										
e										
Digital	28,800,000	100	34	36	16	12	2	4	0.36 se	1.10
India										
				Complex	multi word o	queries	I			1
Evalu	1,890,00,000	100	36	32	11	13	8	0	0.47 se	1.10
ation										
of										
Online										
Marke										
ting										
Evalu	116,00,000	100	42	35	10	9	4	3	0.43 se	1.25
ation										
of										
digital										
India										
Comp	18,400,000	100	43	39	12	4	2	1	0.57se	1.27
uter										
scienc										
e &										
engine										
ering										
Total	4,344,200,000	900	349	331	112	70	35	24		1.18
			(38.78%)	(36.78%)	(12.44%)	(7.78%)	(3.89%)	(2.67 %)		

We selected the first top 100-result link given by Google, Yahoo and Bing web search engine. We try to show comparative precision analysis of Google, Yahoo and Bing show in graph figure 1 in the bases of searching Keyword. Comparative Performance analysis of Google, Yahoo and Bing show in graph figure 2 in the base of searching Keyword and precision. Finally Table 4 are summaries the total precision of simple one word, simple multi word and complex multi word group of Google, Yahoo and Bing and graph figure 3 are show to comparative precision on the basis this three group. We try in graph figure 4 are show to repeated link in the basis of searching link and graph figure 5 are show to searching time and according to searching Keyword.





Figure 1- Comparative precision analysis of Google, Yahoo and Bing

Figure 2- Comparative Performance analysis of Google, Yahoo and Bing



Table 4 – Comparative precision of Google, Yahoo and Bing

Search Engine	Total number of	Total number of	Total number of	Total Precision
	Simple one word	Simple one word Simple multi		
		word	word	
Google	1.36	1.16	0.97	1.09
Yahoo	0.79	1.06	1.29	1.18
Bing	1.2	1.16	1.18	1.18





Figure 4- Repeated link Analysis of Google, Yahoo and Bing



Figure 5- Searching Time Analysis of Google, Yahoo and Yahoo



International Journal of Advanced Technology in Engineering and ScienceVol. No.4, Special Issue No. 01, February 2016ijateswww.ijates.comISSN 2348 - 7550

IV. RELATIVE RECALL CALCULATION OF GOOGLE, YAHOO AND BING

Second factor of performance is relative recall. Recall is a retrieval system and it achieve all or most relevant documents in the collection means recall is the ratio of the amount of relevant records retrieve to the search engine and total number of relevant records in the database. Calculating on relative recall using this formula and this formula (Eq. 2) given by Shafi & Rather 2005[5] Recall is the ratio of the number of relevant records retrieved by search engine to the total number of relevant records in the database.

Relative Recall =Total number of web sites retrieve by a search engine/ Sum of sites retrieved by the all search engine (Eq. 2)

- *a.* **Relative Recall of Google:** Total numbers of 5,29,68,00,000 sites are founded for different nine keyword. Google is given to relative recall is 0.36 in all group but it given in Simple one word group have it recall is 0.65, Simple multi word group have it recall is 0.21 and Complex multi word have it recall is 0.44.
- *b.* **Relative Recall of Yahoo:-** Total numbers of 4,968,400,000 sites are founded for different nine keyword. Yahoo is given to relative recall is 0.34 in all group but it given in Simple one word group have it recall is 0.17, Simple multi word group have it recall is 0.76 and Complex multi word have it recall is 0.22.
- c. Relative Recall of Bing:- Total numbers of 4,344,200,000 sites are founded for different nine keyword. Bing is given to relative recall is 0.29 in all group but it given in Simple one word group have it recall is 0.17, Simple multi word group have it recall is 0.01 and Complex multi word have it recall is 0.33.

The relative recall of the Google, Yahoo and Bing is calculated and show the Table 5 in the base of searching keyword and graph figure 6 shows to comparative analysis. We also try to summaries the total relative recall of simple one word, simple multi word and complex multi word group of Google, Yahoo and Bing in Table 6 and graph figure 7 are show to comparative relative recall on the basis this three group.

Searching Goog		ogle	Ya	hoo	Bing					
Keyword	Total No. of	Relative Recall	Total No. of	Relative Recall	Total No. of	Relative Recall				
	Sites		Sites		Sites					
	Simple one word queries									
Computer	2,30,00,00,000	0.68	539,000,000	0.16	5,370,00,000	0.16				
Database	99,80,00,000	0.67	436,000,000	0.29	539,00,000	0.04				
Multimedia	64,70,00,000	0.54	89,800,000	0.08	458,000,000	0.38				
		Simp	ole multi word que	eries						
What is search	35,90,00,000	0.09	3,740,000,000	0.90	308,00,000	0.07				
engine										
Computer	28,90,00,000	0.89	16,300,000	0.05	167,00,000	0.05				
science										
Digital India	43,30,00,000	0.91	12,700,000	0.03	28,800,000	0.06				
	Complex multi word queries									

Table -5 Relative recall of the Google, Yahoo and Bing

137 | P a g e

International Journal of Advanced Technology in Engineering and Science Vol. No.4, Special Issue No. 01, February 2016 ija

www.ijates.com

ijates
ISSN 2348 - 7550

Internet and	3,23,00,000	0.1	109,000,000	0.33	1,890,00,000	0.06
web designing						
Evaluation of	1,45,00,000	0.39	10,300,000	0.28	116,00,000	0.31
digital India						
Computer	22,40,00,000	0.86	16,300,000	0.06	18,400,000	0.07
science &						
engineering						
Total	5,29,68,00,000	0.36	4,968,400,000	0.34	4,344,200,000	0.29

Figure 6- Comparative Relative Recall analysis of Google, Yahoo and Bing



Table -6 Comparative Relative recall of Google, Yahoo and Bing

Search Engine	Total number of Simple one word	Total number of Simple multi word	Total number of Complex multi word	Relative recall
Google	0.65	0.21	0.44	0.36
Yahoo	0.17	0.76	0.22	0.34
Bing	0.17	0.01	0.33	0.29

Figure 7- Comparative Relative Recall analysis according to word group



International Journal of Advanced Technology in Engineering and Science Vol. No.4, Special Issue No. 01, February 2016 ijates www.ijates.com ISSN 2348 - 7550

V. CONCLUSION

Main aim of this study is evaluating the performance and comparing to Google, Yahoo and Bing web search engine on their performance. In this paper we presented the overview web search engine and Google, Yahoo and Bing search engine. Finally result of this study shows on precision of Google is 1.09, Yahoo is 1.18 and Bing is 1.18 of all group. Then over all Bing and Yahoo precision is higher than Google, Google precision higher than Yahoo and Bing and Bing precision higher then Yahoo in Simple one word. Google and Bing is given to high precision then Yahoo in Simple multi word. Yahoo is given to high precision in Complex multi word but Bing is given to higher Precision then Google. Second factor of performance is Relative Recall (most relevant document) and this result is of Google is 0.36, Yahoo is 0.34 and Bing is 0.29 of all group so Google have a higher Relative Recall and Bing is given to lower Relative Recall in all over group but we taken on group then Google is given to high Relative Recall 0.65 in Simple one word, Yahoo is given to high Relative Recall 0.76 in Simple multi word and Google is given to high Relative Recall 0.44 in Complex multi word group. At last Google is given to best searching Result in simple and Multi Word but we try to search in complex word then it's fail then time we used to Yahoo and Bing Because it's Database Size is small.

REFERENCES

- [1] Chu, H., & Rosenthal, M. (1996), "Search engines for the World Wide Web: A Comparative study and evaluation methodology" Proceedings of the ASIS 1996 Annual Conference, vol- 33, pp- 127-35.
- [2] Leighton, H. (1996)," Performance of four WWW index services, Lycos, Infoseek, Webcrawler and WWW Worm" Retrieved from http://www.winona.edu/library/webind.htm
- [3] Ding, W., & Marchionini, G. (1996)."A Comparative study of the Web search service performance" Proceedings of the ASIS 1996 Annual Conference, vol 33, pp- 136-142
- [4] Clarke, S., & Willett, P. (1997) "Estimating the recall performance of search engines", ASLIB Proceedings, vol 49 (7), pp- 184-189.
- [5] Shafi, S. M., & Rather, R. A. (2005)," Precision and recall of five search engines for retrieval of scholarly information in the field of biotechnology" Webology, 2 (2), Retrieved from http://www.webology.ir/2005/v2n2/a12.html
- [6] B.T. Sampath kumar and S.M.Pavithra "Evaluating the searching capabilities of search engine and metasearch engine: a comparative study" annals of library and information studies vol.57, june 2010, pp-87-97
- [7] Kamlesh Kumar Pandey, "Internet Search Engine: A Comparative and Performance Evaluation of Web Search Engine and Meta Search Engine", published in the Proc. ICERCSE 2015 21 November 2015, Published by : UICSA, Rani Durgavati University, Jabalpur, pp 138 143,ISSN No: 2393-9931.
- [8] Biradar B S, and Sampath Kumar B T, "Internet search engines:a comparative study and evaluation methodology", SRELS Journal of Information Management, volume 43(3),2006 pp-231-241
- [9] Hossein Jadidoleslamy," INTRODUCTION TO METASEARCH ENGINES AND RESULT MERGING STRATEGIES: A SURVEY" International Journal of Advances in Engineering & Technology, Nov 2011, ISSN: 2231-1963, pp-30-40

International Journal of Advanced Technology in Engineering and ScienceVol. No.4, Special Issue No. 01, February 2016ijateswww.ijates.comISSN 2348 - 7550

- [10] K.Srinivas, P.V.S.Srinivas and A.Govardhan. "A Survey on the "Performance Evaluation of Various Meta Search Engines" IJCSI Volume 8, Issue 3, No. 2, May 2011 Pages 359-364.
- [11] Lu Y., Meng W., Shu L., Yu C. and Liu K."Evaluation of result merging strategies for metasearch engines" 6th International Conference on Web Information Systems Engineering (WISE Conference); New York;2005.
- [12] Yiyao Lu, Weiyi Meng, Liangcai Shu, Clement T. Yu, and King-Lup Liu" Evaluation of result merging strategies for
- [13] Meng W., Yu C. and Liu K. "Building efficient and effective metasearch engines" In ACM Computing Surveys; 2002
- [14] MARK LEVENE, "AN INTRODUCTION TOSEARCH ENGINES AND WEB NAVIGATION", ISBN 978-0-470-52684-2 (pbk.) year-2010, Published by John Wiley & Sons, Inc., Hoboken, New Jersey
- [15] Kim-An Phan, "Meta-Search Engine Analysis" Seminar Thesis, University of Fribourg Faculty of Economics & Social Sciences Information Systems Research Group
- [16] http://en.wikipedia.org/wiki/search_engine
- [17] http://www.searchengineguide.com
- [18] https://www.bing.com
- [19] https://www.yahoo.com
- [20] https://www.google.com