IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 14, Issue 5 (Sep. - Oct. 2013), PP 135-141
www.iosrjournals.org

Enhancement in Weighted PageRank Algorithm Using VOL

Sonal Tuteja[1]

[1](Software Engineering, Delhi Technological University,India)

Abstract: *There are billions of web pages available on the World Wide Web (WWW). So there are lots of search results corresponding to a user's query out of which only some are relevant. The relevancy of a web page is calculated by search engines using page ranking algorithms. Most of the page ranking algorithm use web structure mining and web content mining to calculate the relevancy of a web page. In this paper, the standard Weighted PageRank algorithm is being modified by incorporating Visits of Links(VOL).The proposed method takes into account the importance of both the number of visits of inlinks and outlinks of the pages and distributes rank scores based on the popularity of the pages. So, the resultant pages are displayed on the basis of user browsing behavior.*

Keywords: *inlinks, outlinks, search engine, web mining, World Wide Web (WWW).*

I. Introduction

WWW has ample number of hyperlinked documents and these documents contain heterogeneous information including text, image, audio, video, and metadata. Since WWW's evolution, it has expanded by 2000% and its size is doubling in every six to ten months [1]. For a given query, there are lots of documents returned by a search engine out of which only few of them are relevant for a user. So the ranking algorithms are indispensable to sort the results so that more relevant documents are displayed at the top.

Various ranking algorithms have been developed such as PageRank, Weighted PageRank, Page Content Ranking, and HITS etc. These algorithms are based on web structure mining or web content mining or combination of both. But web structure mining only considers link structure of the web and web content mining is not able to cope up with multimedia such as images, mp3 and videos [2]. The proposed method incorporates VOL with Weighed PageRank to calculate the value of page rank. In this way, the algorithm provides better results by merging web usage mining with web structure mining.

The organization of the paper is as follows: In section 2, a brief idea about research background has been given. In Section 3, the proposed work has been described with algorithm and example. Section 4 describes the advantages and disadvantages of proposed work. In section 5, the results of proposed method have been compared with Weighted PageRank and Weighed PageRank using visits of links. Conclusion and future work have been given in section 6.

II. Research Background

Data mining can be defined as the process of extracting useful information from large amount of data. The application of data mining techniques to extract relevant information from the web is called as web mining [3] [4]. It plays a vital role in web search engines for ranking of web pages and can be divided into three categories: web structure mining (WSM), web content mining (WCM) and web usage mining (WUM) [3]. WSM is used to extract information from the structure of the web, WCM to mine the content of web pages and WUM to extract information from the server logs.

Brin and Page [5] came up with an idea at Stanford University to use link structure of the web to calculate rank of web pages. PageRank algorithm is used by Google to prioritize the results produced by keyword based search. The algorithm works on the principle that if a web page has important links towards it then the links of this page to other pages are also considered important. Thus it relies on the backlinks to calculate the rank of web pages. The page rank is calculated by the formula given in equation 1.

$$PR(u) = c \sum_{v \varepsilon \mathrm{B(u)}} \frac{PR(v)}{\mathrm{N_v}} \qquad (1)$$

Where u represents a web page, $PR(u)$ and $PR(v)$ represents the page rank of web pages u and v respectively, B(u) is the set of web pages pointing to u, N_v represents the total numbers of outlinks of web page v and c is a factor used for normalization.

Original PageRank algorithm was modified by taking into consideration that not all users follow direct links on WWW. The modified formula for calculating page rank is given in equation 2.

$$PR(u) = (1-d) + d \sum_{v\varepsilon \mathrm{B(u)}} \frac{PR(v)}{\mathrm{N_v}} \qquad (2)$$

Where d is a dampening factor which represent the probability of user using direct links and it can be set between 0 and 1.

Wenpu Xing and Ali Ghorbani [6] proposed an algorithm called Weighted PageRank algorithm by extending standard PageRank. It works on the principle that if a page is important, more linkages from other web pages have to it or are linked to by it. Unlike standard PageRank, it does not evenly distribute the page rank of a page among its outgoing linked pages. The page rank of a web page is divided among its outgoing linked pages in proportional to the importance or popularity (its number of inlinks and outlinks).

$\mathrm{W^{in}(v,u)}$, the popularity from the number of inlinks, is calculated based on the number of inlinks of page u and the number of inlinks of all reference pages of page v as given in equation 3.

$$\mathrm{W^{in}(v,u)} = \frac{I_u}{\sum_{p\varepsilon \mathrm{R(v)}} I_p} \qquad (3)$$

Where I_u and I_p are the number of inlinks of page *u* and *p* respectively. R(v) represents the set of web pages pointed by v.

$\mathrm{W^{out}(v,u)}$, the popularity from the number of outlinks, is calculated based on the number of outlinks of page *u* and the number of outlinks of all reference pages of page *v* as given in equation. 4.

$$\mathrm{W^{out}(v,u)} = \frac{O_u}{\sum_{p\varepsilon \mathrm{R(v)}} O_p} \qquad (4)$$

Where O_u and O_p are the number of outlinks of page *u* and *p* respectively and R(v) represents the set of web pages pointed by v. The page rank using Weighted PageRank algorithm is calculated by the formula as given in equation 5.

$$PR(u) = (1-d) + d \sum_{v\varepsilon \mathrm{B(u)}} PR(v)\mathrm{W^{in}}(v,u)\mathrm{W^{out}}(v,u) \qquad (5)$$

Gyanendra Kumar et. al. [7] came up with a new idea to incorporate user's broswsing behavior in calculating page rank. Previous algorithms were either based on web structure mining or web content mining but none of them took web usage mining into consideration. A new page ranking algorithm called Page Ranking based on Visits of Links (VOL) was proposed for search engines. It modifies the basic page ranking algorithm by taking into consideration the number of visits of inbound links of web pages. It helps to prioritize the web pages on the basis of user's browsing behavior.

In the original PageRank algorithm, the rank of a page p is evenly distributed among its outgoing links but in this algorithm, rank values are assigned in proportional to the number of visits of links. The more rank value is assigned to the link which is most visited by user. The Page Ranking based on Visits of Links (VOL) can be calculated by the formula given in equation 6.

$$PR(u) = (1-d) + d \sum_{v\varepsilon \mathrm{B(u)}} L_u \frac{PR(v)}{TL(v)} \qquad (6)$$

Where $PR(u)$ and $PR(v)$ represent page rank of web pages u and v respectively, d is dampening factor, *B(u)* is the set of web pages pointing to *u*, L_u is number of visits of links pointing from v to u, $TL(v)$ is the total number of visits of all links from *v*.

Neelam Tyagi and Simple Sharma [8] incorporated user browsing behavior in Weighted PageRank algorithm to develop a new algorithm called Weighted PageRank based on number of visits of links (VOL). The algorithm assigns more rank to the outgoing links having high VOL .It only considers the popularity from the number of inlinks and ignores the popularity from the number of outlinks which was incorporated in Weighted PageRank algorithm.

In the original Weighted PageRank algorithm, the page rank of a web page is divided among its outgoing linked pages in proportional to the importance or popularity (its number of inlinks and outlinks) but in this algorithm, number of visits of inbound links of web pages are also taken into consideration. The rank of web page using this algorithm can be calculated as given in equation 7.

$$WPR_{VOL}(u) = (1-d) + d \sum_{v\varepsilon \mathrm{B(u)}} \frac{L_u WPR_{VOL}(v)\mathrm{W}^{\mathrm{in}}(v,u)}{TL(v)} \qquad (7)$$

Where $WPR_{VOL}(u)$ and $WPR_{VOL}(v)$ represent page rank of web page u and v respectively, d is the dampening factor, *B(u)* is the set of web pages pointing to u, L_u is number of visits of links pointing from v to u, $TL(v)$ is the total number of visits of all links from v, $\mathrm{W}^{\mathrm{in}}(v,u)$ represents the popularity from the number of inlinks of u. Table gives a brief description of above algorithm using some parameters from [9].

Table 1: Comparison of Ranking Algorithms

Algorithm	PageRank	Weighted PageRank	PageRank with VOL	Weighted PageRank with VOL
Web mining technique used	Web structure mining	Web structure mining	Web structure mining, web usage mining	Web structure mining, web usage mining
Input Parameters	Backlinks	Backlinks, Forward links	Backlinks and VOL	Backlinks and VOL
Importance	More	More	More	More
Relevancy	Less	Less	More	More

III. Proposed work

The original Weighted PageRank algorithm distributes the rank of a web page among its outgoing linked pages in proportional to their importance or popularity. $\mathrm{W}^{\mathrm{in}}(\mathrm{v,u})$, the popularity from the number of inlinks and $\mathrm{W}^{\mathrm{out}}(\mathrm{v,u})$, the popularity from the number of outlinks does not include usage trends. It does not give more popularity to the links most visited by the users. The weighted PageRank using VOL makes use of web structure mining and web usage mining but it neglects the popularity from the number of outlinks i.e., $\mathrm{W}^{\mathrm{out}}(\mathrm{v,u})$. In proposed algorithm,$\mathrm{W}^{\mathrm{in}}_{\mathrm{VOL}}(\mathrm{v,u})$, the popularity from the number of visits of inlinks and $\mathrm{W}^{\mathrm{out}}_{\mathrm{VOL}}(\mathrm{v,u})$, the popularity from the number of visits of outlinks are used to calculate the value of page rank. $\mathrm{W}^{\mathrm{in}}_{\mathrm{VOL}}(\mathrm{v,u})$ is the weight of link(v, u) which is calculated based on the number of visits of inlinks of page *u* and the number of visits of inlinks of all reference pages of page *v* as given in equation 8.

$$\mathrm{W}^{\mathrm{in}}_{\mathrm{VOL}}(\mathrm{v,u}) = \frac{I_{u(VOL)}}{\sum_{p\varepsilon\mathrm{R(v)}} I_{p(VOL)}} \qquad (8)$$

Where $I_{u(VOL)}$ and $I_{p(VOL)}$ represents the incoming visits of links of page u and p respectively and R(v) represents the set of reference pages of page v. $\mathrm{W}^{\mathrm{out}}_{\mathrm{VOL}}(\mathrm{v,u})$ is the weight of link(v, u) which is calculated based on the number of visits of outlinks of page u and the number of visits of outlinks of all reference pages of page v as given in equation 9.

$$\mathrm{W}^{\mathrm{out}}_{\mathrm{VOL}}(\mathrm{v,u}) = \frac{O_{u(VOL)}}{\sum_{p\varepsilon\mathrm{R(v)}} O_{p(VOL)}} \qquad (9)$$

Where $O_{u(VOL)}$ and $O_{p(VOL)}$ represents the outgoing visits of links of page u and respectively and R(v) represents the set of reference pages of page v. Now these values are used to calculate page rank using equation 10.

$$EWPR_{VOL}(u) = (1-d) + d \sum_{v\varepsilon\mathrm{B(u)}} WPR_{VOL}\mathrm{W}^{\mathrm{in}}_{\mathrm{VOL}}(\mathrm{v,u})\mathrm{W}^{\mathrm{out}}_{\mathrm{VOL}}(\mathrm{v,u}) \qquad (10)$$

Where d is a dampening factor, B(u) is the set of pages that point to u, $WPR_{VOL}(u)$ and $WPR_{VOL}(v)$ are the rank scores of page u and v respectively, $\mathrm{W}^{\mathrm{in}}_{\mathrm{VOL}}(\mathrm{v,u})$ represents the popularity from the number of visits of inlinks and $\mathrm{W}^{\mathrm{out}}_{\mathrm{VOL}}(\mathrm{v,u})$ represents the popularity from the number of visits of outlinks.

3.1. Algorithm to calculate EWPR$_{\mathrm{VOL}}$

1. *Finding a website*: The website with rich hyperlinks is to be selected because the algorithm depends on the hyper structure of website.
2. *Generating a web graph*: For selected website, web graph a generated in which nodes represent web pages and edges represent hyperlinks between web pages.

3. Calculating number of visits of hyperlinks: Client side script is used to monitor the hits of hyperlinks and information is sent to the web server and this information is accessed by crawlers.

4. Calculate page rank of each web page: The values of $\mathrm{W}^{\mathrm{in}}_{\mathrm{VOL}}\,(\mathrm{v},\mathrm{u})$, the popularity from the number of visits of inlinks and $\mathrm{W}^{\mathrm{out}}_{\mathrm{VOL}}\,(\mathrm{v},\mathrm{u})$, the popularity from the number of visits of outlinks are calculated for each node using formulae given in equation 8 and 9 and these values are substituted in equation 10 to calculate values of page rank.

5. Repetition of step 4: The step 4 is used recursively until a stable value of page rank is obtained. The Fig. 1 shown below explains the steps required to calculate page rank using proposed algorithm.

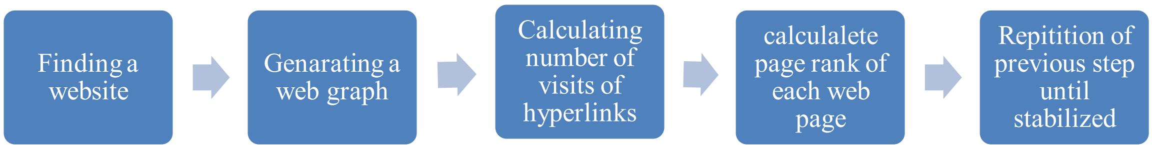
Finding a website
Genarating a web graph
Calculating number of visits of hyperlinks
calculalete page rank of each web page
Repitition of previous step until stabilized

Fig 1: Algorithm to calculate $\mathrm{EWPR}_{\mathrm{VOL}}$

3.2. Example to illustrate the working of proposed algorithm

The working of proposed algorithm has been illustrated via taking a hypothetical web graph having web pages A, B and C and links representing hyperlinks between pages marked with their number of visits shown in Fig. 2.

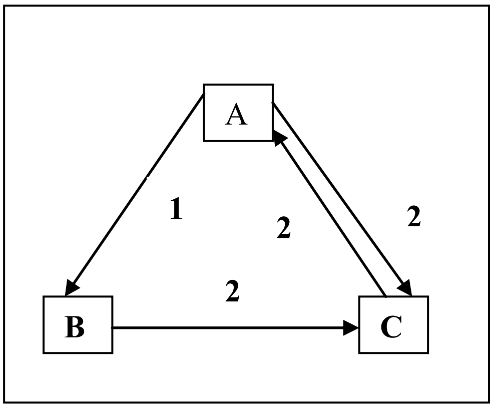

Fig. 2: A web graph

The value of pagerank for web pages A, B and C are calculated using equation 10 as:

$$EWPR_{VOL}(A) = (1-d) + dWPR_{VOL}(C)\mathrm{W}^{\mathrm{in}}_{\mathrm{VOL}}\,(C,A)\mathrm{W}^{\mathrm{out}}_{\mathrm{VOL}}\,(C,A)$$

$$EWPR_{VOL}(B) = (1-d) + dWPR_{VOL}(A)\mathrm{W}^{\mathrm{in}}_{\mathrm{VOL}}\,(A,B)\mathrm{W}^{\mathrm{out}}_{\mathrm{VOL}}\,(A,B)$$

$$EWPR_{VOL}(C) = (1-d) + d(WPR_{VOL}(A)\mathrm{W}^{\mathrm{in}}_{\mathrm{VOL}}\,(A,C)\mathrm{W}^{\mathrm{out}}_{\mathrm{VOL}}\,(A,C) + WPR_{VOL}(\mathrm{B})\mathrm{W}^{\mathrm{in}}_{\mathrm{VOL}}\,(B,C)\mathrm{W}^{\mathrm{out}}_{\mathrm{VOL}}\,(B,C)$$

Each intermediate values $\mathrm{W}^{\mathrm{in}}_{\mathrm{VOL}}\,(v,u)$ and $\mathrm{W}^{\mathrm{out}}_{\mathrm{VOL}}\,(v,u)$ are calculated using equation 8 and 9.

$$\mathrm{W}^{\mathrm{in}}_{\mathrm{VOL}}\,(C,A) = \frac{I_{A(VOL)}}{I_{A(VOL)}} = \frac{2}{2} = 1$$

$$\mathrm{W}^{\mathrm{out}}_{\mathrm{VOL}}\,(C,A) = \frac{O_{A(VOL)}}{O_{A(VOL)}} = \frac{3}{3} = 1$$

$$\mathrm{W}^{\mathrm{in}}_{\mathrm{VOL}}\,(A,B) = \frac{I_{B(VOL)}}{I_{B(VOL)} + I_{C(VOL)}} = \frac{1}{1+2} = \frac{1}{3}$$

$$\mathrm{W}^{\mathrm{out}}_{\mathrm{VOL}}\,(A,B) = \frac{O_{B(VOL)}}{O_{B(VOL)} + O_{C(VOL)}} = \frac{2}{2+2} = \frac{2}{4}$$

$$\mathrm{W}^{\mathrm{in}}_{\mathrm{VOL}}\,(A,C) = \frac{I_{C(VOL)}}{I_{C(VOL)} + I_{B(VOL)}} = \frac{4}{4+1} = \frac{4}{5}$$

$$\mathrm{W}^{\mathrm{out}}_{\mathrm{VOL}}\,(A,C) = \frac{O_{C(VOL)}}{O_{C(VOL)} + O_{B(VOL)}} = \frac{2}{2+2} = \frac{2}{4}$$

$$\mathrm{W}_{\mathrm{VOL}}^{\mathrm{in}}(B,C)=\frac{I_{C(VOL)}}{I_{C(VOL)}}=\frac{4}{4}=1$$

$$\mathrm{W}_{\mathrm{VOL}}^{\mathrm{in}}(B,C)=\frac{O_{C(VOL)}}{O_{C(VOL)}}=\frac{2}{2}=1$$

The calculated values are put in above equations to calculate the values of page ranks. For d = 0.35, page rank values for A, B and C can be calculated as:

$$EWPR_{VOL}(A)=0.65+0.35\left(1*\frac{2}{2}*\frac{3}{3}\right)=1$$

$$EWPR_{VOL}(B)=0.65+0.35\left(1*\frac{1}{3}*\frac{2}{4}\right)=.70833$$

$$EWPR_{VOL}(C)=0.50+0.50\left(1*\frac{4}{5}*\frac{2}{4}+.70833*\frac{4}{4}*\frac{2}{2}\right)=1.03792$$

These values are calculated iteratively until the values get stabilized and the final values of page ranks are: A=1.01406, B=0.70915 and C=1.04017. For d = 0.50, page rank values for A, B and C can be calculated as:

$$EWPR_{VOL}(A)=0.50+0.50\left(1*\frac{2}{2}*\frac{3}{3}\right)=1$$

$$EWPR_{VOL}(B)=0.50+0.50\left(1*\frac{1}{3}*\frac{2}{4}\right)=0.58333$$

$$EWPR_{VOL}(C)=0.50+0.50\left(1*\frac{4}{5}*\frac{2}{4}+.58333*\frac{4}{4}*\frac{2}{2}\right)=0.99167$$

These values are calculated iteratively until the values get stabilized and the final values of page ranks are: A=0.99527, B=0.58294 and C=0.99052. For d = 0.85, page rank values for A, B and C can be calculated as:

$$EWPR_{VOL}(A)=0.15+0.85\left(1*\frac{2}{2}*\frac{3}{3}\right)=1$$

$$EWPR_{VOL}(B)=0.15+0.85\left(1*\frac{1}{3}*\frac{2}{4}\right)=0.29167$$

$$EWPR_{VOL}(C)=0.50+0.50\left(1*\frac{4}{5}*\frac{2}{4}+.29167*\frac{4}{4}*\frac{2}{2}\right)=0.73792$$

These values are calculated iteratively until the values get stabilized and the final values of page ranks are: A=0.63531, B=0.24 and C=0.57001. The value of page rank at various values of d has been given in Table.

Table 2: Value of page ranks at different d values

d	A	B	C
0.35	1.01406	0.70915	1.04017
0.50	0.99527	0.58294	0.99052
0.85	0.63531	0.24001	0.57001

IV. Advantages and Disadvantages

The proposed method includes web usage mining to calculate the page ranks of web pages and has following advantages.

- The page rank using original WPR remains unaffected whether the page has been accessed by the users or not. i.e.; the relevancy of a web page is ignored. But the page rank using proposed method $EWPR_{VOL}$ assigns high rank to web pages having more visits of links.
- The page rank using original WPR depends only on the link structure of the web and remains same whether the web page has been accessed by the user or not. Although the algorithm WPR_{VOL} makes use of web structure mining and web usage mining to calculate the value of page rank but it ignores the popularity from the number of outlinks $\mathrm{W}^{\mathrm{out}}(\mathrm{v},\mathrm{u})$. On the other side, our proposed method $EWPR_{VOL}$ makes use

of $W^{in}_{VOL}(v,u)$, the popularity from the number of visits of inlinks and $W^{out}_{VOL}(A,C)$, the popularity from the number of visits of outlinks to calculate page rank.

- The proposed method uses number of visits of links to calculate the rank of web pages. So the resultant pages are popular and more relevant to the users need.

The proposed method includes number of visits of links of web pages to calculate page ranks. Suppose there is a junk page whose initial page rank is high then the users will access it and it will lead to increase in VOL which will further improve page rank. So the pages which are actually relevant will have less page rank than junk pages. So some other usage behavior factors must be introduced in addition to VOL. These factors are:

- *Time spent on web page corresponding to a link:* The algorithm must assign more weight to the link if more time is spent by the users on the web page corresponding to that link. Most of the times, the time spent on the junk pages is very less as compared to relevant pages. So this factor will help in lowering the rank of junk pages.
- Most recent use of link: The link which is used most recently by users should have more priority than the link which has been not used so far. So most recent use of link can also be used to calculate the page rank.
- Information about the user: A web page is not equally relevant for all the users. Due to different requirements of different users, a web page may be more important for one but not for other. Some kind of users' information like age, gender, educational background can be used to categorize web pages according to different users' need.

V. Result Analysis

This section compares the page rank of web pages using standard Weighted PageRank (WPR), Weighted PageRank using VOL (WPR_{VOL}) and the proposed algorithm. We have calculated rank value of each page based on WPR, WPR_{VOL} and proposed algorithm i.e. $EWPR_{VOL}$ for a web graph shown in Fig. 2.

As the value of dampening factor increases, the page rank decreases. The comparison of results is shown in Table which shows the values of page rank using WPR, WPR_{VOL} and $EWPR_{VOL}$ at different d values of 0.35, 0.50 and 0.85.

Table 3: Comparison of page ranks using different algorithms

d		0.35	0.50	0.85
WPR	A	1.00535	0.97677	0.58335
	B	0.70865	0.58140	0.23335
	C	1.01532	0.95351	0.51505
WPR(VOL)	A	1.01736	1	0.64037
	B	0.68956	0.55556	0.21495
	C	1.04960	1	0.57463
EWPR(VOL)	A	1.01406	0.99527	0.63531
	B	0.70915	0.58140	0.23335
	C	1.04017	0.99052	0.57001

The values of page rank using WPR, WPR_{VOL} and $EWPR_{VOL}$ have been compared using a bar chart. The values retrieved by $EWPR_{VOL}$ are better than original WPR and WPR_{VOL}. The WPR uses only web structure mining to calculate the value of page rank, WPR_{VOL} uses both web structure mining and web usage mining to calculate value of page rank but it uses popularity only from the number of inlinks not from the number of outlinks. The proposed algorithm $EWPR_{VOL}$ method uses number of visits of inlinks and outlinks to calculate values of page rank and gives more rank to important pages. Fig. 3 compares the page ranks of A, B and C using WPR, WPR_{VOL} and $EWPR_{VOL}$ for d=0.35.

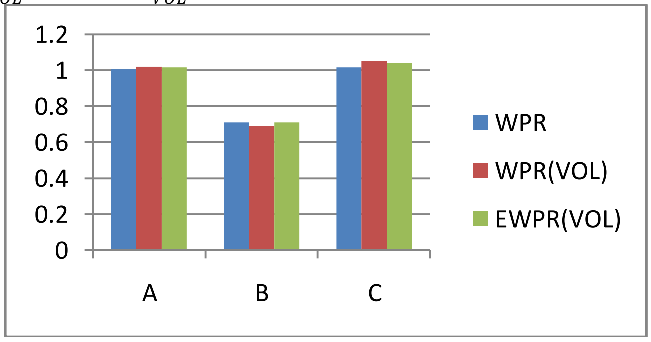

Fig. 3: Comparison of page ranks at d=0.35

Fig. 4 and Fig. 5 compares the page ranks of A, B and C using WPR, WPR_{VOL} and $EWPR_{VOL}$ for d=0.50.

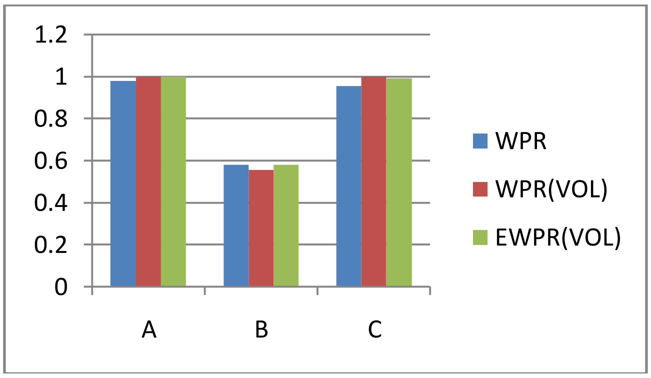

Fig. 4: Comparison of page ranks at d=0.50

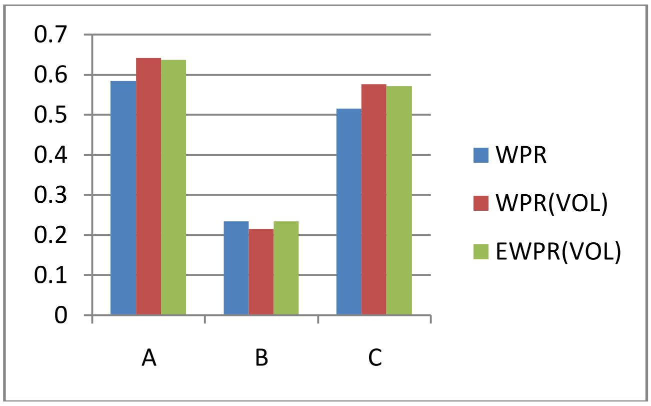

Fig. 5: Comparison of page ranks at d=0.85

VI. Conclusions and Future Work

Due to enormous amount of information present on the web, the users have to spend lot of time to get pages relevant to them. So the proposed algorithm $EWPR_{VOL}$ makes use of number of visits of links (VOL) to calculate the values of page rank so that more relevant results are retrieved first. In this way, it may help users to get the relevant information quickly. Some of the future works for the proposed algorithm are:

- The values of page rank have been calculated on a small web graph only. A web graph with large number of websites and hyperlinks should be used to check the accuracy and importance of method.
- We need some other measures like most recent use of link, information about the user and time spent on web page corresponding to a link. So the future work includes deriving a formula for page rank using these parameters also.

References

[1] Sergey Brin and Lawrence Page, "The Anatomy of a Large-Scale Hypertextual Web Search Engine", *Computer Science Department, Stanford University, Stanford, CA 94305.*

[2] Wenpu Xing and Ali Ghorbani, "Weighted PageRank Algorithm", *Faculty of Computer Science, University of New Brunswick,Fredericton, NB, E3B 5A3, Canada.*

[3] Gyanendra Kumar, Neelam Duhan, A. K. Sharma, "Page Ranking Based on Number of Visits of Links of Web Page*", Department of Computer Engineering, YMCA University of Science & Technology, Faridabad, India*.

[4] Naresh Barsagade, "Web Usage Mining And Pattern Discovery: A Survey Paper", CSE 8331, Dec.8, 2003.

[5] Dell Zhang, Yisheng Dong, "A novel Web usage mining approach for search engines", Computer Networks 39 (2002) 303–310

[6] Neelam Duhan, A. K. Sharma, Komal Kumar Bhatia, "Page Ranking Algorithms: A Survey" Advance Computing Conference, 2009. IACC 2009 IEEE International.

[7] R.Cooley, B.Mobasher and J.Srivastava, "Web Mining: Information and Pattern Discovery on the World Wide Web". In Proceedings of the 9th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'97), 1997.

[8] Companion slides for the text by Dr. M. H. Dunham, "Data Mining: Introductory and Advanced Topics", Prentice Hall, 2002.

[9] Neelam tyagi, Simple Sharma, "Weighted Page Rank Algorithm Based on Number of Visits of Links of Web Page", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-3, July 2012