

COMPARISON BETWEEN GOOGLE'S PAGERANK ALGORITHM AND PAGE RANKING BASED ON LINK VISITS (PRLV)

Amar Singh¹, Gyanendra Kumar², Pushpneel Verma³

*^{1,2}M.tech Scholar, ³Associate Professor, Department of Computer Science,
Bhagwant Institute of Technology, Uttar Pradesh Technical University, Lucknow, (India)*

ABSTRACT

This paper analyzes the PageRank algorithm in brief and puts forward Page Ranking based on Link Visits (PRLV) algorithm. The PRLV algorithm takes into account the probability of user visits on a page with the motive to determine the importance and relevance score of web pages. The main aim of the paper is to compare the PageRank algorithm with the PRLV and show how PRLV could be the successor of PageRank.

Keywords: Search Engine, PageRank algorithm, PRLV (Page Ranking based on Link Visits), Rank of pages, Web Structure Mining, Web Usage Mining.

1. INTRODUCTION

The web creates new challenges for information retrieval. The amount of information on the web is growing rapidly, as well as the number of new users inexperienced in the art of web research.

The problem in locating the correct information on the internet led to the creation of search technology, known as the internet search engine. Search engines are the foundation of the Internet. Most users will turn to a search engine as the quickest way of finding the information, or product that they want.

A search engine receives user query, it processes the query, it search into its index for relevant documents (documents that are likely related to their query and supposed to be interesting to them), then the search engine ranks the documents found relevant and it shows them the results.

The problems with this approach are:

- The present approaches use text matching as a solution to define the relevance between documents. However, these approaches have many disadvantages. For example, if a person searches for information about C, how would the engine know whether to search about the alphabet C or the programming language "C"?
- The second problem is that the methods and algorithms to rank the results by search engine are not best and can be further optimized. The best technique to optimize is to know and include information about the user

For example, if we know that the user search for information related to the computer science field we can infer that with “C”, he is probably referring to the programming language.

Each search engine does the same functions of collecting, organizing, indexing and serving results in its own unique ways, thus employing various algorithms and techniques, which are their trade secrets. The functions of a search engine can be categorized into the following:

1. Crawling the internet for web content.
2. Indexing the web content.
3. Storing the website contents.
4. Search algorithms and results.

The amount of information on the internet on any topic is very large; a large number of pages related to any query are available. But the user usually requires a small amount of the information. Hence, the search engines incorporate ranking algorithms in order to sort the results, so that the user will have the most important results first.

II PAGERANK ALGORITHM

Within the past few years, Google[1] has become the far most utilized search engine worldwide. The Google search engine has two important features that help it produce high precision results.

- It makes use of the link structure of the Web to calculate a quality ranking for each Web page.
- Google utilizes links to improve search results.

PageRank[2] is an algorithm used by Google Search to rank websites in their search engine results. Sergey Brin[3] and Larry Page[4] developed PageRank, named after Larry Page (cofounder of Google search engine), that uses the link structure of the web to determine the importance of web pages.

According to Google:

“PageRank works by counting the number and quality of links to a page to determine a rough estimate of how important the website is. The underlying assumption is that the more important websites are likely to receive more links from other websites -Facts about Google and Competition.”

The idea behind PageRank is that pages with many backlinks are more important than pages with only a few backlinks. The concept is really simple; let us think of each link from page q to page p as a vote or recommendation. It means that highly linked pages are more “important” (voted or recommended).

Imagine a person creates a new important page. So, he/she wants it to be visible in the top results. The only thing to be done is to create many other pages that point to the particular page or if the page has only two backlinks, but they are Netscape and Yahoo!, the page will have a greater rank. We can state the PageRank as “a page has greater rank if sum of ranks of backlinks is high”, i.e. if a page has important incoming links, then outgoing links to other pages also become important.

Therefore, the main disadvantage with PageRank is that it favors old pages, because a new page, even an important one will not have many links unless it is a part of an existing multi-connected site. A simplified version of page rank is defined as follows:

$$PR(p) = (1 - c) \sum_{q \in I(p)} \frac{PR(q)}{O(q)} + cE(p) \quad (2.1)$$

A factor c is used for normalization. {Note that $0 < c < 1$.}

III PAGE RANKING BASED ON LINK VISITS (PRLV)

PRLV (Page Ranking based on Link Visits) is based on Web Structure Mining and Usage Mining. It takes into account the user visits with the motive to determine the importance and relevance score of web pages. The various subtasks to be performed from gathering the usage characterization till final rank is determined are:

- Saving user's access information (hits) on an outgoing link of a page in related server log files.
- Fetching the pages and access information by the targeted web crawler.
- Compute weights based on the probabilities of the link being visited by the users for each page link.
- Final rank computation of pages based on the weights of their incoming links.
- Retrieval of ranked pages corresponding to user queries.

The weights are determined for out-links of pages (inner-document structure mining) and ranks are computed by taking back-links into account (inter-document structure mining).

3.1 Features of PRLV

- As PRLV method uses link structure of pages and their browsing information, the top returned pages in the result list are supposed to be highly relevant to the user information needs. In the PageRank method, the rank value of any page will be same regardless of whether the user sees it or not as it is totally dependent upon link structure of Web graph whereas the ordering of pages using PRLV is more target-oriented.
- In PRLV, the rank depends on the probability and not number of visits. Therefore, a user can not deliberately increase the rank by visiting the page multiple times.
- There is an extra request to store the number of hits of a particular page. Hence, congestion over network will increase.

3.2 Calculation of rank based on PRLV

Outbound link

Suppose we have a page p , with n hyperlinks embedded. The outbound link set is:

$$O(p) = \{o_1, o_2, \dots, o_n \mid \text{each } o_i \text{ is a URL (page) which can be accessed from page } p\}. \quad (3.1)$$

Inbound Link

A page p is said to have a set B of m inbound links:

$$B = \{b_1, b_2, \dots, b_m \mid \text{each } b_i \text{ is a URL (page) from which } p \text{ can be accessed}\}. \quad (3.2)$$

Probability-Weight of Link

If p is a page with outbound-link set $O(p)$ and each outbound link is associated with a numerical integer indicating visit-count (VC), then the weight of each outbound link connecting p and o is calculated by:

$$Weight_{link}(p, o) = \frac{VC(p, o)}{\sum_{o' \in O(p)} VC(p, o')} \quad (3.3)$$

Page Rank based on Link Visits (PRLV)

If p is a page having inbound-linked pages in set $B(p)$, then the rank (PRLV) is given by:

$$PRLV(p) = (1-d) + d \left(\sum_{b \in B(p)} PRLV(b) \cdot Weight_{link}(b, p) \right) \quad (3.4)$$

Where d is the damping factor as is used in PageRank; $Weight_{link}(b, p)$ is the weight of the link calculated as above..

3.3 Retrieval of Ranked Pages by PRLV

The crawler passes the downloaded pages and their links visit data to the indexer and PRLV Calculator. The indexing of pages is done and their rank score is calculated. The index repository maintained by the search engine holds the necessary information about the crawled pages.

When a user submits query to the search engine interface, query processor matches the keywords and then finds the set of pages with the rank values. These pages along with their respective rank values are shown to the user in list-form. As soon as the user clicks on a page, the same process is repeated.

IV COMPARISON OF PRLV WITH PAGERANK [5]**4.1 Definition and Basic Principles**

The idea behind PageRank is that pages with many backlinks are more important than pages with only a few backlinks. PageRank computes scores at indexing time. Results are sorted according to importance of pages.

PRLV takes into account the user visits with the motive to determine the importance and relevance score of web pages. It computes scores at indexing time. Pages are sorted according to importance and relevance.

4.2 Mining Techniques Used

PageRank uses Web Structure Mining (WSM) to generate structural summary about the Web sites and Web pages. The structure of a typical Web graph consists of Web pages as nodes and hyperlinks as edges connecting two related pages. In technical terms, WSM aims to find the link structure of the hyperlinks at the inter-document level.

PRLV uses both Web Structure Mining (WSM) and Web Usage Mining (WUM). Web Structure Mining is used to find the link structure of the hyperlinks and Web Usage Mining (WUM) tries to discover user navigation patterns from web data and the useful information from the secondary data derived from the interactions of the users while surfing on the Web. It focuses on the techniques that could predict user behavior while the user interacts with Web.

4.3 Rank Score of Pages and Its Nature

In PageRank, all the outgoing links from the page have same ranks, irrespective of the page they are leading to. Also the rank values are less dynamic as the rank changes with link structure.

In PRLV, the rank of each page is decided by the probability of visit to that page. So, ranks are unequally distributed among outgoing links according to their probabilities of visit. Thus the rank values are more dynamic as the rank changes with visit counts & structure of links.

4.4 Parameters to be considered

Only the inbound links of pages, i.e. the pages linked to a particular page are the deciding parameters for the rank calculation of the page in PageRank.

In PRLV, inbound links i.e. links leading to a particular page, outbound links i.e. links from that page and Visit Counts of links i.e. the probability of the visits to the link, all are used in rank calculation.

4.5 Relevancy & Importance of Pages

PageRank focuses only on the incoming links into a page to calculate rank. Thus the relevance of pages is lost. But PRLV considers the user feedback, i.e. the probability of a user using a particular page. Hence, the pages displayed have more relevance in a particular concept.

Though the pages in PRLV are more relevant, the PageRank and PRLV, both display important pages.

4.6 Complexity

The complexity of PageRank is $O(\log n)$ and that of PRLV is $> O(\log n)$.

4.7 Advantages

Advantages of Page Rank algorithm

- Computation of ranks with minimum effort
- Fewer complexes as compared to PRLV.

Advantages of PRLV

- Pages returned are of high quality and relevancy as user feedbacks are taken into account.
- Search space can be very much pruned as pages are sorted according to users' information needs.

4.8 Limitations

The limitations of PageRank are

- No relevancy of pages is considered in rank computation.
- All links are considered equally important.

The limitations of PRLV are

- Extra effort on crawlers to fetch the visit counts of pages from web servers.
- Extra calculations to find the weights of links.

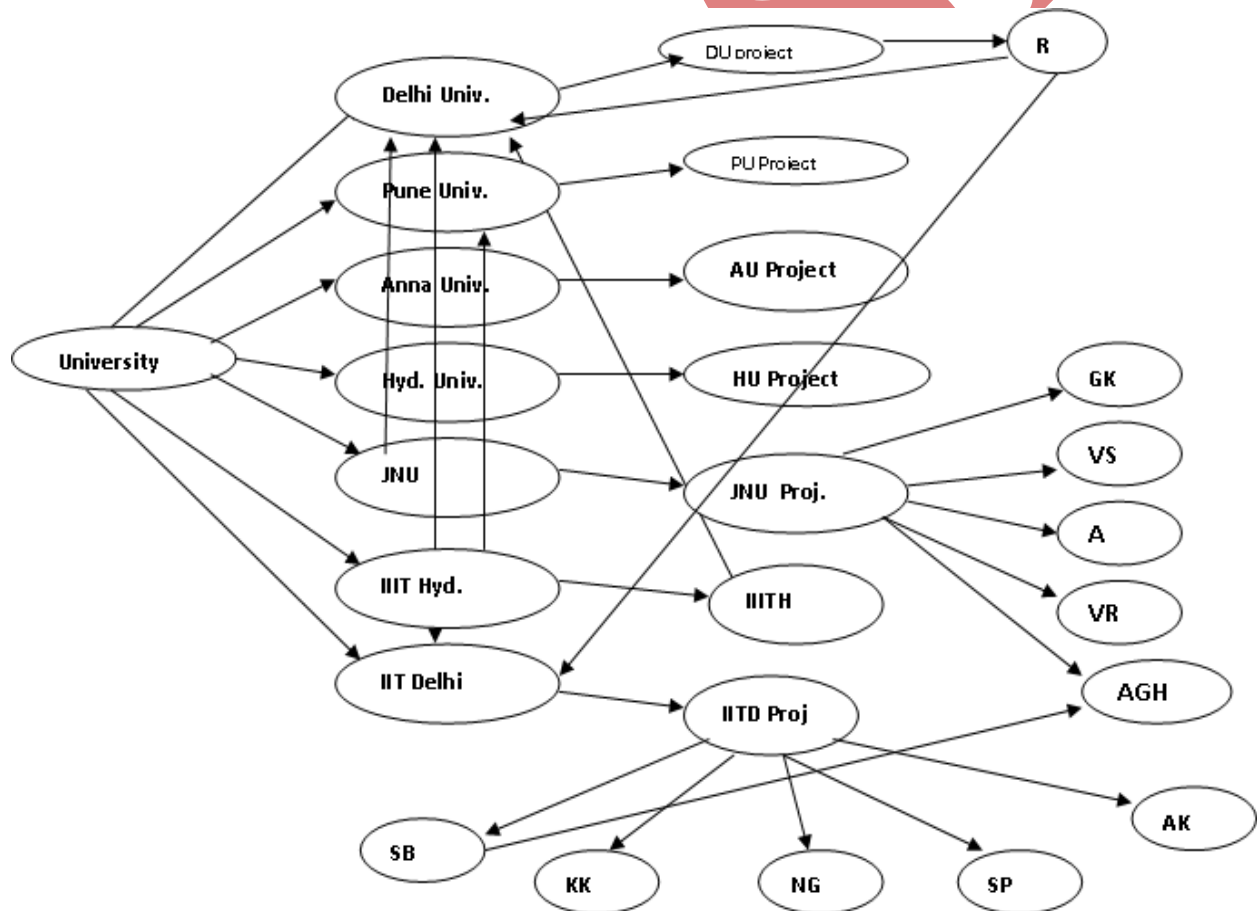


Fig: 1

V COMPARISON WITH EXAMPLE

A web graph as shown above in Fig 1: and the link structure in the graph was considered. Firstly, the original page rank algorithm, PageRank was applied on this graph to calculate rank value of each webpage and then PRLV was applied to calculate rank scores. In this graph each node is abbreviated with name of page.

5.1 Rank Score of pages based on Page Rank

PageRank algorithm is based on the link structure of the web; rank value of a page will not change until the changes in web link structure. The online Page Rank Calculator tools were used to calculate rank value of pages in this graph.

Value of damping factor $d=0.85$ is taken.

Values of different pages are given in Table 5.1 and this value will not change until we change the graph.

Table 5.1 Rank Score based on PageRank Algorithm

Web Page	Rank Score
anna university project.html	0.2929
anna university.html	0.1682
Anshul Kumar IIT Delhi.html	0.2499
Arun K. Attri jnu.html	0.1876
Du project.html	0.6431
IIIT Hyderabad.html	0.1682
IIT DELHI.html	0.5184
Indian Council of Agricultural Research.html	0.4066
gnu.html	0.1682
K.G. Saxena JNU.html	0.1876
Naveen Garg IIT DELHI.html	0.2499
Prof. K. K. Biswas.html	0.2499
Project in JNU.html	0.2214
R & D of IIIT Hyderabad.html	0.1857
rakesh kumar.html	0.7011
Research Areas IIT Delhi.html	0.5824
Sanjiva Prasad IIT Delhi.html	0.2499
Sorav Bansal IIT Delhi.html	0.2499
university of delhi.html	0.5898
university of hyderabad.html	0.1682
university of pune.html	0.3618
university.html	0.15
V. Rajamani jnu.html	0.1876

5.2 Rank Score of pages based on PRLV

PRLV is based on number of visits of links and this link visit information is so dynamic that is why we calculate the rank score for three different state of link visit information.

Figures 5.1, 5.2 and 5.3 show the state1, state2 and state3 of link visits information. Table 5.2 shows the PRLV rank scores for all the three states, viz. state1, state2 and state3.

Figure 5.1: A Cut of State1 of Link Visits Information

Page	Url	Hits
http://localhost/webgraph/university.html	http://localhost/webgraph/jnu.html	35
http://localhost/webgraph/Research Areas IIT Delhi...	http://localhost/webgraph/Prof. K. K. Biswas.html	10
http://localhost/webgraph/university.html	http://localhost/webgraph/university of hyderabad....	20
http://localhost/webgraph/university.html	http://localhost/webgraph/IIT DELHI.html	47
http://localhost/webgraph/university.html	http://localhost/webgraph/university of delhi.html	17
http://localhost/webgraph/university.html	http://localhost/webgraph/university of pune.html	7
http://localhost/webgraph/university.html	http://localhost/webgraph/anna university.html	14
http://localhost/webgraph/university.html	http://localhost/webgraph/IIIT Hyderabad.html	12
http://localhost/webgraph/Research Areas IIT Delhi...	http://localhost/webgraph/Sanjiva Prasad IIT Delhi...	18
http://localhost/webgraph/Research Areas IIT Delhi...	http://localhost/webgraph/Sorav Bansal IIT Delhi.h...	32
http://localhost/webgraph/jnu.html	http://localhost/webgraph/IIT DELHI.html	13
http://localhost/webgraph/Du project.html	http://localhost/webgraph/Rakesh Kumar.html	23
http://localhost/webgraph/IIT DELHI.html	http://localhost/webgraph/Research Areas IIT Delhi...	18
http://localhost/webgraph/jnu.html	http://localhost/webgraph/Project in JNU.html	1
http://localhost/webgraph/Sanjiva Prasad IIT Delhi...	http://localhost/webgraph/Indian Council of Agricu...	10
http://localhost/webgraph/jnu.html	http://localhost/webgraph/university of delhi.html	3
http://localhost/webgraph/Research Areas IIT Delhi...	http://localhost/webgraph/Anshul Kumar IIT Delhi.h...	5
http://localhost/webgraph/Research Areas IIT Delhi...	http://localhost/webgraph/Naveen Garg IIT DELHI.ht...	1
http://localhost/webgraph/jnu.html	http://localhost/webgraph/Research Areas IIT Delhi...	5
http://localhost/webgraph/anna university.html	http://localhost/webgraph/anna university project...	12
http://localhost/webgraph/university of hyderabad....	http://localhost/webgraph/university of pune.html	21
http://localhost/webgraph/university of pune.html	http://localhost/webgraph/IIIT Hyderabad.html	60
http://localhost/webgraph/university of pune.html	http://localhost/webgraph/Vaibhav V. Kaware pune.h...	26
http://localhost/webgraph/university of delhi.html	http://localhost/webgraph/Du project.html	12
http://localhost/webgraph/IIIT Hyderabad.html	http://localhost/webgraph/R & D of IIIT Hyderabad....	12
http://localhost/webgraph/Project in JNU.html	http://localhost/webgraph/Arun K. Attri jnu.html	20
http://localhost/webgraph/Project in JNU.html	http://localhost/webgraph/V. Rajamani jnu.html	3
http://localhost/webgraph/Project in JNU.html	http://localhost/webgraph/Indian Council of Agricu...	6
http://localhost/webgraph/Project in JNU.html	http://localhost/webgraph/V. Subramanian jnu.html	2
http://localhost/webgraph/V. Rajamani jnu.html		0

Figure 5.2: A Cut of State2 of Link visit information

Page	Url	Hits
http://localhost/webgraph/university.html	http://localhost/webgraph/jnu.html	45
http://localhost/webgraph/Research Areas IIT Delhi...	http://localhost/webgraph/Prof. K. K. Biswas.html	15
http://localhost/webgraph/university.html	http://localhost/webgraph/university of hydarabad....	20
http://localhost/webgraph/university.html	http://localhost/webgraph/IIT DELHI.html	47
http://localhost/webgraph/university.html	http://localhost/webgraph/university of delhi.html	27
http://localhost/webgraph/university.html	http://localhost/webgraph/university of pune.html	7
http://localhost/webgraph/university.html	http://localhost/webgraph/anna university.html	14
http://localhost/webgraph/university.html	http://localhost/webgraph/IIIT Hyderabad.html	12
http://localhost/webgraph/Research Areas IIT Delhi...	http://localhost/webgraph/Sanjiva Prasad IIT Delhi...	18
http://localhost/webgraph/Research Areas IIT Delhi...	http://localhost/webgraph/Sorav Bansal IIT Delhi.h...	32
http://localhost/webgraph/jnu.html	http://localhost/webgraph/IIT DELHI.html	13
http://localhost/webgraph/Du project.html	http://localhost/webgraph/Rakesh Kumar.html	23
http://localhost/webgraph/IIT DELHI.html	http://localhost/webgraph/Research Areas IIT Delhi...	18
http://localhost/webgraph/jnu.html	http://localhost/webgraph/Project in JNU.html	1
http://localhost/webgraph/Sanjiva Prasad IIT Delhi...	http://localhost/webgraph/Indian Council of Agricu...	10
http://localhost/webgraph/jnu.html	http://localhost/webgraph/university of delhi.html	3
http://localhost/webgraph/Research Areas IIT Delhi...	http://localhost/webgraph/Anshul Kumar IIT Delhi.h...	5
http://localhost/webgraph/Research Areas IIT Delhi...	http://localhost/webgraph/Naveen Garg IIT DELHI.ht...	1
http://localhost/webgraph/jnu.html	http://localhost/webgraph/Research Areas IIT Delhi...	5
http://localhost/webgraph/anna university.html	http://localhost/webgraph/anna university project...	12
http://localhost/webgraph/university of hydarabad....	http://localhost/webgraph/university of pune.html	21
http://localhost/webgraph/university of pune.html	http://localhost/webgraph/IIIT Hyderabad.html	60
http://localhost/webgraph/university of pune.html	http://localhost/webgraph/Vaibhav V. Kaware pune.h...	26
http://localhost/webgraph/university of delhi.html	http://localhost/webgraph/Du project.html	12
http://localhost/webgraph/IIIT Hyderabad.html	http://localhost/webgraph/R & D of IIIT Hyderabad....	22
http://localhost/webgraph/Project in JNU.html	http://localhost/webgraph/Arun K. Attri jnu.html	20
http://localhost/webgraph/Project in JNU.html	http://localhost/webgraph/V. Rajamani jnu.html	3
http://localhost/webgraph/Project in JNU.html	http://localhost/webgraph/Indian Council of Agricu...	6
http://localhost/webgraph/Project in JNU.html	http://localhost/webgraph/V. Subramanian jnu.html	2
http://localhost/webgraph/V. Rajamani jnu.html		0

Figure 5.3: A Cut of State3 of Link visit information

Page	URL	Hits
http://localhost/webgraph/university.html	http://localhost/webgraph/jnu.html	45
http://localhost/webgraph/Research Areas IIT Delhi...	http://localhost/webgraph/Prof. K. K. Biswas.html	27
http://localhost/webgraph/university.html	http://localhost/webgraph/university of hydarabad....	44
http://localhost/webgraph/university.html	http://localhost/webgraph/IIT DELHI.html	47
http://localhost/webgraph/university.html	http://localhost/webgraph/university of delhi.html	27
http://localhost/webgraph/university.html	http://localhost/webgraph/university of pune.html	7
http://localhost/webgraph/university.html	http://localhost/webgraph/anna university.html	23
http://localhost/webgraph/university.html	http://localhost/webgraph/IIIT Hyderabad.html	12

http://localhost/webgraph/Research Areas IIT Delhi...	http://localhost/webgraph/Sanjiva Prasad IIT Delhi...	38
http://localhost/webgraph/Research Areas IIT Delhi...	http://localhost/webgraph/Sorav Bansal IIT Delhi.h...	32
http://localhost/webgraph/jnu.html	http://localhost/webgraph/IIT DELHI.html	13
http://localhost/webgraph/Du project.html	http://localhost/webgraph/Rakesh Kumar.html	23
http://localhost/webgraph/IIT DELHI.html	http://localhost/webgraph/Research Areas IIT Delhi...	18
http://localhost/webgraph/jnu.html	http://localhost/webgraph/Project in JNU.html	4
http://localhost/webgraph/Sanjiva Prasad IIT Delhi...	http://localhost/webgraph/Indian Council of Agricu...	10
http://localhost/webgraph/jnu.html	http://localhost/webgraph/university of delhi.html	30
http://localhost/webgraph/Research Areas IIT Delhi...	http://localhost/webgraph/Anshul Kumar IIT Delhi.h...	5
http://localhost/webgraph/Research Areas IIT Delhi...	http://localhost/webgraph/Naveen Garg IIT DELHI.ht...	21
http://localhost/webgraph/jnu.html	http://localhost/webgraph/Research Areas IIT Delhi...	23
http://localhost/webgraph/anna university.html	http://localhost/webgraph/anna university project....	15
http://localhost/webgraph/university of hyderabad...	http://localhost/webgraph/university of pune.html	21
http://localhost/webgraph/university of pune.html	http://localhost/webgraph/IIIT Hyderabad.html	60
http://localhost/webgraph/university of pune.html	http://localhost/webgraph/Vaibhav V. Kaware pune.h...	26
http://localhost/webgraph/university of delhi.html	http://localhost/webgraph/Du project.html	12
http://localhost/webgraph/IIIT Hyderabad.html	http://localhost/webgraph/R & D of IIIT Hyderabad...	22
http://localhost/webgraph/Project in JNU.html	http://localhost/webgraph/Arun K. Attri jnu.html	20
http://localhost/webgraph/Project in JNU.html	http://localhost/webgraph/V. Rajamani jnu.html	3
http://localhost/webgraph/Project in JNU.html	http://localhost/webgraph/Indian Council of Agricu...	60
http://localhost/webgraph/Project in JNU.html	http://localhost/webgraph/V. Subramanian jnu.html	12
http://localhost/webgraph/V. Rajamani jnu.html		0

Table 5.2 Rank Score based on PRLV for state1, state2 and state3

Web Page	PRLV St1	PRLV St2	PRLV St3
anna university project.html	0.15	0.286	0.290
anna university.html	0.1617	0.160	0.164
Anshul Kumar IIT Delhi.html	0.1771	0.193	0.174
Arun K. Attri jnu.html	0.2567	0.236	0.178
Du project.html	0.3072	0.313	0.347
IIIT Hyderabad.html	0.3365	0.334	0.338
IIT DELHI.html	0.2795	0.630	0.585
Indian Council of Agricultural Research.html	0.3606	0.435	0.518
jnu.html	0.1793	0.183	0.178

K.G. Saxena JNU.html	0.15	0.150	0.150
Naveen Garg IIT DELHI.html	0.1554	0.159	0.251
Prof. K. K. Biswas.html	0.2043	0.280	0.280
Project in JNU.html	0.1569	0.157	0.159
R & D of IIIT Hyderabad.html	0.436	0.434	0.438
rakesh kumar.html	0.416	0.416	0.445
Research Areas IIT Delhi.html	0.4222	0.721	0.697
Sanjiva Prasad IIT Delhi.html	0.2478	0.305	0.333
Sorav Bansal IIT Delhi.html	0.324	0.426	0.304
university of delhi.html	0.185	0.191	0.232
university of hydarabad.html	0.1667	0.165	0.177
university of pune.html	0.2976	0.295	0.305
university.html	0.15	0.150	0.150
V. Rajamani jnu.html	0.166	0.163	0.154
V. Subramanian jnu.html	0.1606	0.159	0.167
Vaibhav V. Kaware pune.html	0.2264	0.226	0.228

Table 5.3 PRLV VS. Page Rank of State1,State2 and State3

Web Page	PRLV St1	PRLV St2	PRLV St3	Page Rank
anna university project.html	0.15	0.286	0.290	0.2929
anna university.html	0.1617	0.160	0.164	0.1682
Anshul Kumar IIT Delhi.html	0.1771	0.193	0.174	0.2499
Arun K. Attri jnu.html	0.2567	0.236	0.178	0.1876
Du project.html	0.3072	0.313	0.347	0.6431
IIIT Hyderabad.html	0.3365	0.334	0.338	0.1682
IIT DELHI.html	0.2795	0.630	0.585	0.5184
Indian Council of Agricultural Research.html	0.3606	0.435	0.518	0.4066
jnu.html	0.1793	0.183	0.178	0.1682
K.G. Saxena JNU.html	0.15	0.150	0.150	0.1876
Naveen Garg IIT DELHI.html	0.1554	0.159	0.251	0.2499
Prof. K. K. Biswas.html	0.2043	0.280	0.280	0.2499
Project in JNU.html	0.1569	0.157	0.159	0.2214
R & D of IIIT Hyderabad.html	0.436	0.434	0.438	0.1857
rakesh kumar.html	0.416	0.416	0.445	0.7011
Research Areas IIT Delhi.html	0.4222	0.721	0.697	0.5824
Sanjiva Prasad IIT Delhi.html	0.2478	0.305	0.333	0.2499
Sorav Bansal IIT Delhi.html	0.324	0.426	0.304	0.2499
university of delhi.html	0.185	0.191	0.232	0.5898
university of hydarabad.html	0.1667	0.165	0.177	0.1682
university of pune.html	0.2976	0.295	0.305	0.3618
university.html	0.15	0.150	0.150	0.15
V. Rajamani jnu.html	0.166	0.163	0.154	0.1876
V. Subramanian jnu.html	0.1606	0.159	0.167	0.1876
Vaibhav V. Kaware pune.html	0.2264	0.226	0.228	0.1876

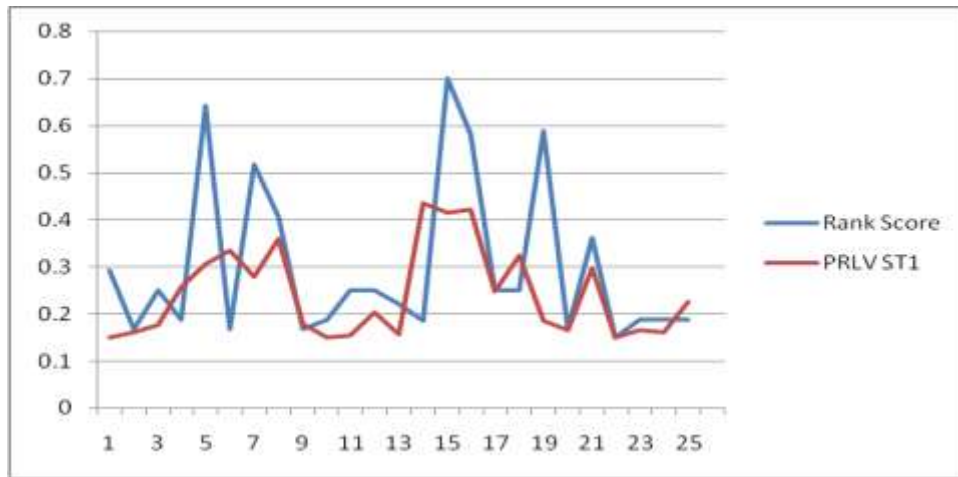


Figure 5.4: Variation of PRLV and PageRank for State1

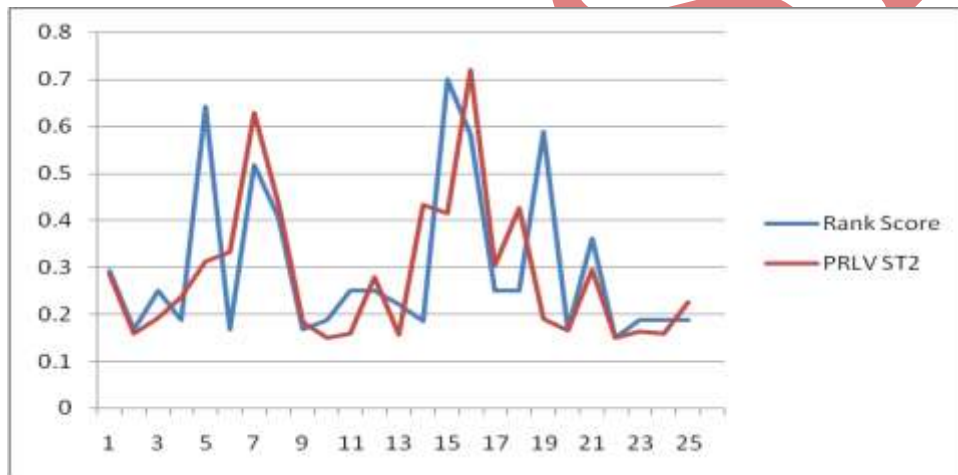


Figure 5.5: Variation of PRLV and PageRank for State2

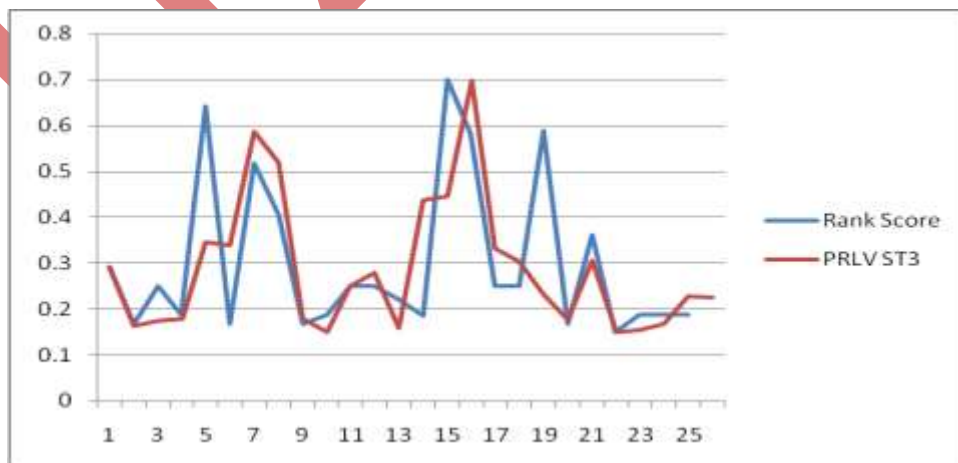


Figure 5.6: Variation of PRLV and PageRank for State3

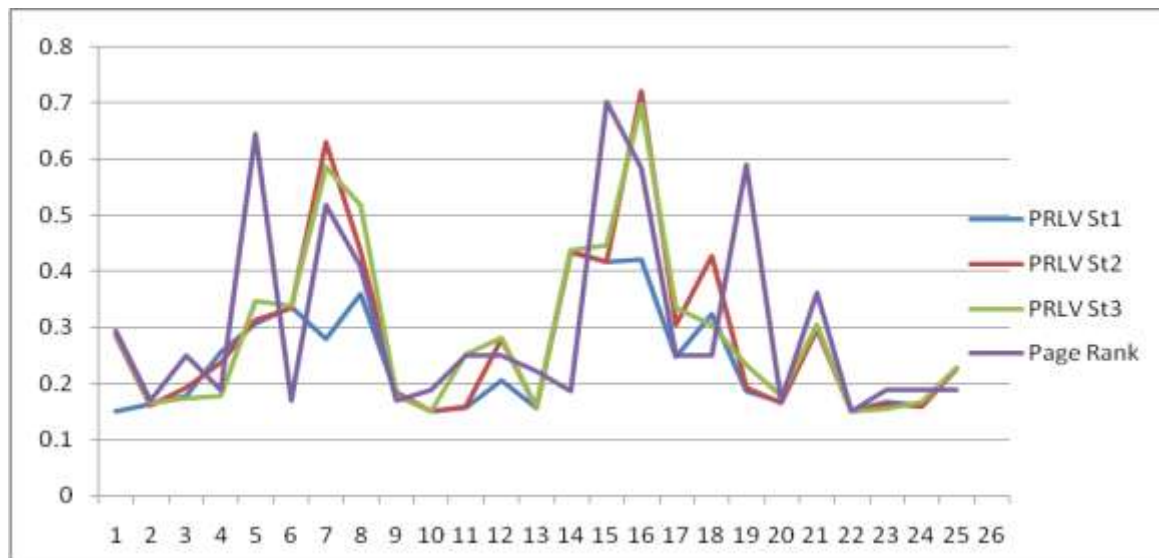


Figure 5.7: Variation of PRLV and PageRank for State1, State2 and State3

VI CONCLUSIONS

The users today spend of lot of time in browsing through the search results to find the appropriate pages. The paper presented a novel page ranking algorithm called PRLV that provides more relevant results than original PageRank and compares them both. PRLV calculates rank value of a web page based on the user visits on incoming links of that page. So, the results frequently used by the user are shown first. The ordering of pages in this way increases the relevancy of pages and thereof provides the user with quality search results.

Therefore, the user finds the desired output in the few top pages and search space is diminished to a very large scale.

REFERENCES

- [1] Google, Inside Search. <http://www.google.co.in/insidesearch/>
- [2] Sergey Brin and Lawrence Page. "The Anatomy of a Large-Scale Hypertextual Web Search Engine." Computer Science Department, Stanford University, Stanford, CA. <http://infolab.stanford.edu/~backrub/google.html>
- [3] Sergey Brin and Larry Page. Google search engine. <http://google.stanford.edu>
- [4] Page, Lawrence and Brin, Sergey and Motwani, Rajeev and Winograd, Terry (1999) . "The PageRank Citation Ranking: Bringing Order to the Web." Technical Report. Stanford InfoLab. <http://ilpubs.stanford.edu:8090/422/>