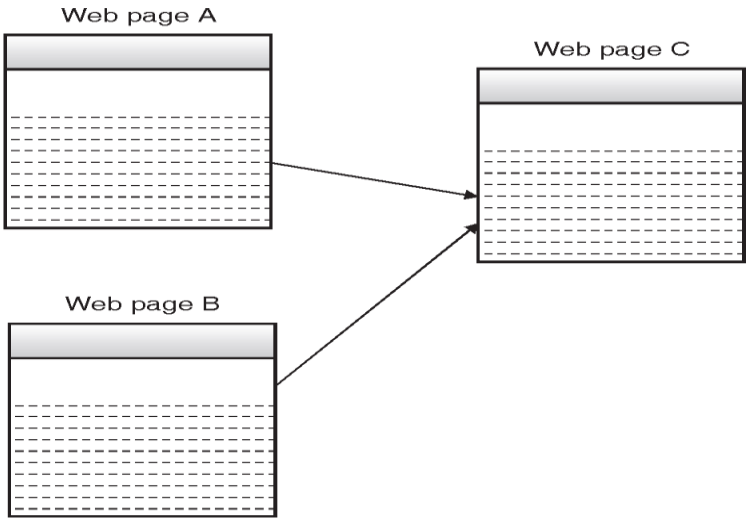
Page Ranking : The term page Ranking can be defined as, "A classical method used to arrange the web pages according to its objective and the usage of terms involved in it on the world wide web by using any link data structure."

The page Ranking mechanism was developed by Larry page and sergey Brin. This page-Ranking mechanism was a part of their research project which was started in 1995 and result into a functional prototype in 1998.

Links in Page Ranking

Suppose we have 3 pages A, B, C in a given domain of web sites. They have interconnection links between them as shown in Fig.



The number of links exists between two or more web pages can be categorize as follows :
1. Back links
2. Forward links

1. Back links
With reference to Fig. A and B are the Back links of web page 'C' i.e. Back link indicates given web page is referred by how many number of other web pages.

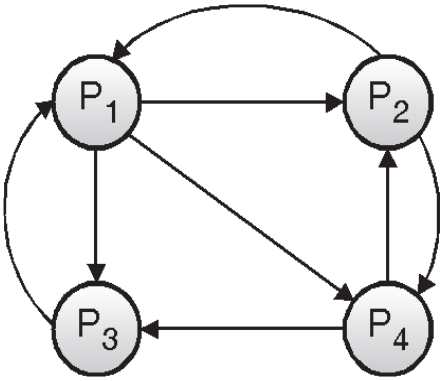
2. Forward link
– Forward link represents the fact that, how many web pages will be referred by a given web pages.
– Clearly, out of these two types of links back links are very important from Ranking of documents perspective.
– A web page which contains number of back links is said to be important web page and will get upper position in Ranking.

– A page Ranking in mathematical format can be represented as

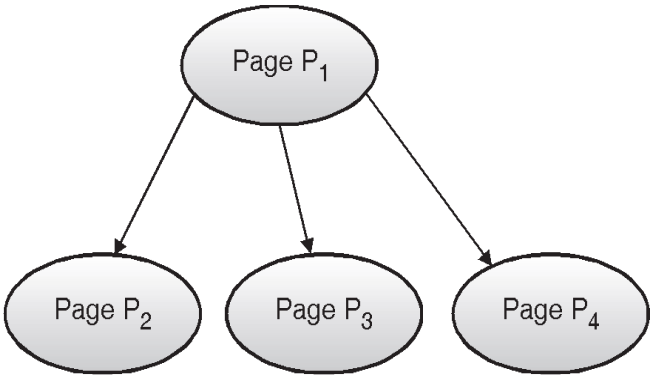
$$R(u) = c \sum_{V \in B_u} \frac{R(V)}{N_V}$$

Where,
μ : Represents the web page Bμ
N_v : It represents number of forward links of page v.
C : It represents the Normalizations factor to make
$| R |_{L1} = 1 (| R |_{L1} = |R_1 + R_2 + \dots R_n|)$
– A world wide web can be considered as the 'Di-graph' i.e. Directed graph Any graph 'G' is composed of two fundamental components vertices and Edges.

Here, vertices or Nodes can be mapped to pages.
If we consider a small part of world wide web containing 4 web pages named as P_1, P_2, P_3, P_4.
Every page i has Back links and forward links to other pages.
Fig. shows the above mentioned structure.



In Fig. page P1 has Forward links to page P2, P3 and P4 respectively.
Page P2 has links to page 1 and page P3 .
Page 3 has link to page 1 and
Page 4 has links to page 2 and page 3.
If a user starts surfing with page P1 in above web page P1 has links to page P2 , P3 and P4



Probability that user will be at page P_2 / P_3 / P_4 is equal to 1/3.
Probability that user will be at page 1 itself is '0'.
Suppose user has chosen page P_2 then
Probability that user will be at page P_1 is 1/2,
Probability that user will be at page P_4 is 1/2.
Probability that user will be at page P_2 or P_3 is 'D'.
– These possibilities of web surfing by a given user can be represented using special structure known as "Transition Matrix".
– In general, the transition Matrix 'M' is composed of 'n' pages 'n' rows and 'n' columns. Two pointer c and j will be to represent the current row and columns respectively.

Any given element can be represent as m_{ij}.
$m_{ij} = 1/k$ if and only if the page at j_{th} column has k forward links.
– Additionally one of the forward links to same page itself.
∴ The transition Matrix for above web can be represented as

$$M = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 1/2 & 1 & 0 \\ 1/3 & 0 & 0 & 1/2 \\ 1/3 & 0 & 0 & 1/2 \\ 1/3 & 1/2 & 0 & 0 \end{bmatrix} \end{array}$$

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
(ARTIFICIAL INTELLIGENCE & MACHINE LEARNING)