Fuzz-PageRanking for Google Search Engine

Dr. Tahseen Jilani[1], Ubaida Fatima[2], Prof. Dr. Mirza Mahmood Baig[3]

[1]Department of Computer Science, University of Karachi, Karachi, Pakistan

[2]Department of Mathematics, NED University of Engineering and Technology

Karachi, Pakistan

[3]Department of Mathematics, NED University of Engineering and Technology

Karachi, Pakistan

Abstract:

In this research we have proposed a novel approach for the computation of PageRank vector of Google web search engine and appellate this approach as Fuzz-PageRank Approach (FUPRA). Practically, the number of web connections is a fuzzy concept and thus can be modeled using fuzzy logic and fuzzy sets. Through fuzzification of the fuzzy transition probability matrix, the proposed approach have accelerated the convergence rate of PageRank and renamed this vector as Fuzz-PageRank. For simplicity, we have assumed a triangular membership function for each element in a Google matrix. We have compared the convergence rate and number of iterations of the standard PageRank algorithms with our proposed method. The results has shown that our proposed approach has clearly outperformed the PageRank techniques.

Keywords:

Fuzz-PageRank, Eigenvalues, PageRank, Power Method, Fuzzy Adaptive Method

1. Introduction:

Web comprises of a complex anatomy with billions of webpages and numbers of links. Interminably changes have been made in web pages, that is new pages are being modify, some of the pages being blot out and at any time remaining web pages are being revised, all this will increase the complications to fully understand the mathematics behind the world wide web. As millions of people creates their different style webpages, so world web contains distinct range of web pages style. Anyhow it was anxious to whether the data is confined by a distinct article or is blowout in a number of articles. [1]

When the surfer visiting webpage 'a' at some particular time t after some time that is t+1 randomly move to another webpage 'b' with probability 1/deg(a) (where deg is the number of outlinks of webpage 'a') this movement is defined by stochastic transition matrix. In form of random walk of surfer Markov chain was induced and the stochastic transition matrix describe the movement of surfer from webpage 'a' to webpage 'b' and demonstrated as P with P_{ab} in Equation (1). [2][19]

$$P_{ab} = \frac{1}{\deg(a)} \tag{1}$$

Markov process was demonstrated by number of web surfer simulations. [11][12][2] Many links in each web page represents the transition probabilities. [4] It was fictive that there was no ambiguity in the considered web pages or in the transition probabilities. This is not the actual situation in reality. Fuzzy sets or fuzzy numbers help to simulate this uncertainty. [11] This forms the ground for the current study, where we have developed the Fuzzy-PageRank approach (FUPRA). Markov chain was converted into fuzz-Markov chain as we had cope with fuzzy numbers.

We introduce the Fuzz-PageRank in this article by making use of fuzzy sets or fuzzy numbers, transition probabilities as define Fuzz-Markov chain. [1] Our study on Google hyperlink data strongly validate the edge of Fuzz-PageRank over PageRank.

Ambiguities of non-probabilistic systems can be simulated easily by Fuzzy numbers. The real numbers set of fuzzy subsets contains fuzzy numbers. Arithmetic interval or the extension principle are the main ground for arithmetic operations on fuzzy numbers. The shape of membership function sturdily indicates how the fuzzy numbers were operated. Analogous figure of membership function can inferred naturally. Present implementations involve Triangular fuzzy numbers or Trapezoidal fuzzy numbers. [1]

Bhamidipati and Pal (2006) proposed the different model for web surfer, in which links between webpages are represented by fuzzy magnitudes. They compared fuzzRank with PageRank by its ranking, rate of convergence and validness. Benefits of fuzzRank over PageRank were studies through comprehensive example. [1]

Yan, Gui, Du and Guo (2011) studied that searching through web had become essential to gain cognizance from titanic web material. They presented a genetic PageRank algorithm (GPRA) based on previously studied PageRank algorithm. They had proved by experiments that genetic PageRank algorithm (GPRA) is efficient as compared to PageRank algorithm. [3]

Dubey and Roy (2011) deliberated the PageRank algorithm. They revised the two methods for ranking of webpages that are Hypertext Induced Topic Search (HITS) and PageRank techniques to evaluate the significance of webpages through the Hyperlink anatomy of the web. They had introduced the novel PageRank algorithm that based on normalization method to increase its convergence rate by decreasing the iterations. [4]

Vajargah and Gharehdaghi (2012) studied fuzzy finite Markov chains and demonstrate their assets ground on possibility theory. Simulation of fuzzy Markov chain was done by them for dissimilar sizes. Fuzzy Markov chain periodicity behavior was upgraded. [5]

Kandiah and Shepelyansky (2012) simulated the PageRank of opinion formation and genuine directed graph of some Universities and sites were examined. PageRanking was done for Google search engine by considering opinion formation of connected constituents is subjective with their PageRank probability. They analyzed that sites network which they had considered had a robust susceptibility to a despotic opinion formation as compared to universities network and also inferred that the Sznajd model for scale free-networks. [6]

Tabrizi, Shakery, Asadpour, Abassi and Tavallaie (2013) discussed graph clustering and its role in numerous applications. They presented the Personalized PageRank Clustering (PPC) algorithm that is top-down algorithm that can expose graph clustering more precisely as compared to bottom-up algorithms that are near-linear approaches. This top-down algorithm consist of linear time and space complexity and been preferred to number of accessible clustering algorithm. [7]

Garcia, Pedroche and Romance (2013) showed some new outcomes that aid the concept of personalized PageRank. They worked on directed graph that may be comprised of dangling nodes. The personalized PageRank for every webpage analytically described that helps to categorize the surfers of Social Networks Websites. Novel ideas were presented in dealing ambitiously complex networks by making use of data associated to directed graph adjacency matrix and its dangling node distribution vector. [8]

Bourchtein and Bourchtein (2013) inferred some properties of the general PageRank algorithm. The showed that usually convergence of PageRank vector is non-uniform. PageRank vector were evaluated in form of actual stochastic matrix and personalization vector. From outcomes of experiment webpages ranking analyzed. [9]

Sharma and Gupta (2013) studied the significance of Fuzzy logic in different areas like Mathematics, Information Technology and Artificial Intelligence. They explained their fundamental concepts and fuzzy sets operations. They discussed that the uncertainties in life can be understand by correlating it with fuzzy set theory and fuzzy logic. As Fuzzy logic is a technique where one can reckon gradation of truth, which can be a number between 0 and 1. [10]

Koumenides and Shadbolt (2014) presented the analysis of semantic web search methods. They concentrated on ranking techniques and supporting methods discovered by present semantic search organizations. They discussed the future work in the field of research and constructing community consensus for better ordered assessment and slow progress for computation of PageRank algorithm. [11]

2. Outline to Fuzzy Logic and Fuzzy Set Theory:

The number of sets in the world in our surroundings are outlined by an un-sharp boundary. [12] Fuzzy logic are not only consist of two options but an entire range of truth values for logical propositions. [10] Fuzzy logic comprises 0 and 1 as the extreme cases of truth but also embraces in between 0 and 1 the number of conditions of truth [12][13] In reality, fuzzy knowledge used as uncertain, probabilistic, imprecise, ambiguous or inexact form. [10][22]

Whereas the classical set theory is constructed on the fundamental concept of which one can be belongs to set or not belongs to set. A crisp set is sharp and definite survives on only two options that "one belongs to" or "one not belongs to" the set in this theory, and there is a very specific and clear boundary to specify that an individual is a member of the set. The fuzzy set theory is an expansion of Fuzzy set theory where entities are matter of degrees. [14]

Membership function determines a grade of sameness of elements in the universe of discourse U to fuzzy sets. [14][15] Each entity of a universe of discourse U maps by a membership function into real numbers between 0 and 1. [14] A more appropriate and summarizing approach to outline a MF is to demonstrate it as a mathematical formula. [16] MF can be demonstrated as

$$\mu_A: U \to [0,1] \quad \textbf{(2)}$$

3. Markov Chains:

A sequence $\{A_n\}_{n \in N}$ is representing a Markov chain of first order of random variables where every random variable, A_i, take values from state space U and satisfies the Equation (3). [1]

$$P(A_{n+1} \backslash A_o, A_1, \dots \dots.., A_n) = P(A_{n+1} \backslash A_n) \quad \textbf{(3)}$$

First order Markov chain is declared as homogenous in the case where $P(A_{n+1} \backslash A_n)$ is independent of n. Suppose p_{ij} demonstrate single move transition probability from state i to state j represented as $P(A_{n+1} = i \backslash A_n = j)$. Presuming a state j in Equation (4) as:

$$P(A_{n+1} = j) = \sum_{i=1}^{N} P(A_{n+1} = j \backslash A_n = i)\, P(A_n = i) \quad \textbf{(4)}$$

$$P(A_{n+1} = j) = \sum_{i=1}^{N} p_{ij}\, P(A_n = i) \quad \textbf{(5)}$$

Similarly for k-step transition probability from state i to state j is given by $p_{ij}^{(k)}$ as:

$$p_{ij}^{(k)} = P(A_{n+k} = i \backslash A_n = j) \quad \textbf{(6)}$$

$$p_{ij}^{(k)} = \sum \prod_{m=1}^{k} P(A_{n+m} = a_{n+m} \backslash A_{n+m-1} = a_{n+m-1}) \quad \textbf{(7)}$$

$$p_{ij}^{(k)} = \sum_{a_{n+1},\dots\dots,a_{n+k-1} \in U} \prod_{m=1}^{k} p_{a_{n+m}, a_{n+m-1}} \quad \textbf{(8)}$$

It was perceived that k-step transition probability is equal to the k^{th}power of P. If every state of U is aperiodic then Markov chain is also declared as aperiodic. In a case when each pair of states in U is connected to each other then it is known as irreducible Markov chain. A regular Markov chain should be finite, aperiodic and irreducible that is $p_{ij}^{(n)} \to \pi_j \forall i, j \in U$. Stationary distribution of Markov chain is given by $\pi = \pi_1, \pi_2, \dots \dots, \pi_N$ which is considered as ergodicity means independent of initial state, $P(A_n = j)$ converges to a distinctive π_j. [1]

4. Fuzzy Markov Chains:

The probabilities mentioned in Markov chains are real numbers and are identified. Practically they are assessed, and some errors are related with the assessment method, which might be assessed under appropriate suppositions. Fuzzy numbers are a better way to model the ambiguity in the transition probabilities. [1]A fuzzy distribution on U is outlined by a mapping μ_a : U → [0, 1], and a vector a is demonstrated as in Equation (9). [1]

$$a = (\mu_a(1), \dots, \mu_a(N)). \quad \textbf{(9)}$$

Fuzzy distribution on the Cartesian product $U \times U$ is demonstrated as fuzzy transition matrix P is of the form $\left((p_{ij})\right)_{i,j \in U}$ is called single step transition probability matrix.[1][17] A fuzzy Markov chain is outlined as a series of random variables, through which movement from state to another state can be evaluated by the fuzzy relation P and follow. [1]

$$\mu_{a(n+1)}(j) = \max_{i \in U} \{\mu_{a(n)}(i) \wedge p_{ij}\}, j \in U \quad \textbf{(10)}$$

Above equation is the fuzzy algebraic equivalent of the transition law of classical Markov chains. The attention-grabbing outcome is that, contrasting the situation of classical Markov chains, whenever the sequence of matrices P^n converges, it does so in finitely many steps to a matrix P^T. If it does not converge, it oscillates with a finite period v starting from some finite power. [1][17]

The fuzzy Markov chain is said to be aperiodic if the powers of related transition matrix P converge to a non-periodic solution P^T that is known as limiting fuzzy transition matrix. When the row of solved matrix P^T are same then a fuzzy Markov chain is named as ergodic. [1]

5. The Fuzzy Matrix of the Web:

In any directed Web graph the number of pages are exhibit by nodes and the edges in between pages of directed web graphs are links. For example the 6-page web directed graph is revealed in figure below. [18][15][16]

The directed graph in the figure below is demonstrated by directed graph matrix (DGM) in Equation (11) as:

$$DGM = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

(11)

Page 1 to page 2 of the web was linked by directed edge and page 2 is not linked to page 1 as there is no edge between these two nodes. The inlinks to a particular pages indicates the importance of that page. Page 3 has zero out degree as it has no directed edge towards any page.

The computation of Fuzz-PageRank is begin by converting directed graph matrix (DGM) to Fuzzy matrix (FM). Based on the strength of relation from i^{th} webpage to j^{th} webpage, we can represent the fuzzy numbers to demonstrate the directed web page. As the number of links from i^{th} webpage to j^{th}webpage is a fuzzy concept and is highly unpredictable, thus this uncertainty can be qualified using fuzzy sets. [1][14]

Heuristically, we assume that the strength of relationship from i^{th} to j^{th}webpage is a linguistic variable for example if there is no link from webpage i to j then we assign a membership value of 0.25. We assume that actually the two websites are not connected but there is some weak independence between i^{th} and j^{th} webpages.

We quantify the web links strength as weak, medium, high and very high. In the diagram demonstrated below, we have presented the concept using fuzzy triangular membership functions. As we assumed the values by judgment, thus selection of initial fuzzy membership values is a handcrafting problem as well and need further research to formally define this assignment.

$$FM = \begin{bmatrix} 0.25 & 0.5 & 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.5 & 0.5 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 & 0.5 & 0.5 \\ 0.25 & 0.25 & 0.25 & 0.25 & 0.25 & 0.5 \\ 0.25 & 0.25 & 0.5 & 0.25 & 0.25 & 0.25 \end{bmatrix}$$

(12)

Now making it row stochastic Fuzzy Adjacency matrix (FA) as shown in Equation (13). [2]

$$FA = \begin{bmatrix} row1/sum(row1) \\ row2/sum(row2) \\ row3/sum(row3) \\ row4/sum(row4) \\ row5/sum(row5) \\ row6/sum(row6) \end{bmatrix}$$

(13)

That is Fuzzy Adjacency matrix becomes:

$$FA = \begin{bmatrix} 0.1429 & 0.2857 & 0.1429 & 0.1429 & 0.1429 & 0.1429 \\ 0.1250 & 0.1250 & 0.2500 & 0.2500 & 0.1250 & 0.2500 \\ 0.1667 & 0.1667 & 0.1667 & 0.1667 & 0.1667 & 0.1667 \\ 0.1250 & 0.1250 & 0.1250 & 0.1250 & 0.2500 & 0.2500 \\ 0.1429 & 0.1429 & 0.1429 & 0.1429 & 0.1429 & 0.2857 \\ 0.1429 & 0.1429 & 0.2857 & 0.1429 & 0.1429 & 0.1429 \end{bmatrix}$$

(14)

So in Fuzzy Adjacency or Fuzzy Hyperlink matrix row 3 represents Fuzzy dangling node that is node with zero out-degree in fuzzy hyperlink matrix, we can say row with similar values. Continuing rows recognizes Fuzzy non-dangling nodes.

Through fuzzy concept we have also treated dangling nodes by given weightage to them known as Fuzzy dangling nodes collectively with non-dangling nodes known as Fuzzy non-dangling nodes. [2][18]

6. Fuzzy Google Matrix:

The stochastic irreducible Fuzzy Google matrix demonstrated in Equation (15). [2][18]

$$G_f = \alpha FA + (1-\alpha)\mathbb{1}v_f$$

(15)

In Equation (12), $(1-\alpha)$ is a probability with which surfer at any page will transport to any web page avoids to jump to outlink (page with zero out degree). The terminus of the random move is selected with respect to the fuzz-personalization vector v_f and $\mathbb{1}$ is a column matrix holding each component equals to one.

Fuzzy Google matrix G_f is defined by Markov matrix as its each row sum equals to 1. [2]

Damping factor $\alpha = 0.85$ is the most preferred choice for reckoning of PageRank vector for Google matrix similarly we had considered $\alpha = 0.85$ for reckoning of fuzz-PageRank vector for fuzzy Google matrix. [18]

$$G_f = 0.85\begin{bmatrix} 0.1429 & 0.2857 & 0.1429 & 0.1429 & 0.1429 & 0.1429 \\ 0.1250 & 0.1250 & 0.2500 & 0.2500 & 0.1250 & 0.2500 \\ 0.1667 & 0.1667 & 0.1667 & 0.1667 & 0.1667 & 0.1667 \\ 0.1250 & 0.1250 & 0.1250 & 0.1250 & 0.2500 & 0.2500 \\ 0.1429 & 0.1429 & 0.1429 & 0.1429 & 0.1429 & 0.2857 \\ 0.1429 & 0.1429 & 0.2857 & 0.1429 & 0.1429 & 0.1429 \end{bmatrix} + (1-0.85)\begin{bmatrix}1\\1\\1\\1\\1\\1\end{bmatrix}\begin{bmatrix}1/6\\1/6\\1/6\\1/6\\1/6\\1/6\end{bmatrix} \quad (16)$$

$$G_f = \begin{bmatrix} 0.1465 & 0.2678 & 0.1465 & 0.1465 & 0.1465 & 0.1465 \\ 0.1313 & 0.1313 & 0.2375 & 0.2375 & 0.1313 & 0.1313 \\ 0.1667 & 0.1667 & 0.1667 & 0.1667 & 0.1667 & 0.1667 \\ 0.1313 & 0.1313 & 0.1313 & 0.1313 & 0.2375 & 0.2375 \\ 0.1465 & 0.1465 & 0.1465 & 0.1465 & 0.1465 & 0.2678 \\ 0.1465 & 0.1465 & 0.2678 & 0.1465 & 0.1465 & 0.1465 \end{bmatrix} \quad (17)$$

7. Computation of PageRank vector and Fuzz-PageRank vector by Numerical Inference:

7.1 Fuzz-PageRank by Standard Power Method:

The most preferred choice for the computation of stationary distribution Fuzz-PageRank vector $\pi_f{}^{(k+1)}$ is standard power method as for the estimation of PageRank studied by Wills. Eigenvalues are approximated by standard power method. For estimation of Fuzz-PageRank we uses $\pi_f{}^{(k+1)} = G_f\pi_f{}^{(k)}$, where $\pi_f^{(k+1)}$ converges to a desired tolerance value. See Algorithm 1 and Algorithm 2. [2][18][19]

We have calculated the PageRank vector and Fuzz-PageRank vector for the directed graph for figure 2.

Algorithm: 1 PageRank by Power Method

function PageRank(G,$\boldsymbol{\pi}^{0}$)
repeat
$\boldsymbol{\pi}^{(k+1)}$**= G**$\boldsymbol{\pi}^{(k)}$ **;**
r = |$\boldsymbol{\pi}^{(k+1)}$**| –**$\boldsymbol{\pi}^{(k)}$**|| ;**
until r < ε
return $\boldsymbol{\pi}^{(k+1)}$**;**

Algorithm: 2 Fuzz-PageRank by Power Method

function PageRank(G,$\boldsymbol{\pi}_f^{0}$)
repeat
$\boldsymbol{\pi}_f^{(k+1)}$**= G**$\boldsymbol{\pi}_f^{(k)}$ **;**
r = ||$\boldsymbol{\pi}_f^{(k+1)}$ **–**$\boldsymbol{\pi}_f^{(k)}$**|| ;**
until r < ε
return $\boldsymbol{\pi}_f^{(k+1)}$ **;**

7.2 Fuzz-Adaptive Method:

Kamvar et.al researched the methods to improve the rate of convergence of PageRank vector. In this research work I am considering the methods to progress the rate of convergence of Fuzz-PageRank. Fuzz-Adaptive method converges rapidly. By power method we perform redundant computation for pages that converges rapidly with the webpages that have slow rate of convergence. [2]

Rending the Fuzz-PageRank vector $\pi_f^{(k)}$ into C, converged $n-m$ webpages and into N, non-converged m webpages. $\pi_f^{(k)}$ for the current iteration may be demonstrated as in Equation (18):

$$\pi_f^{(k)} = \begin{bmatrix}\pi_{fN}^{(k)}\\ \pi_{fC}^{(k)}\end{bmatrix} \quad \textbf{(18)}$$

The Fuzzy Google Matrix can also be render into two matrices. The pages that have not converged are demonstrated by G_{fN} with dimensions $m \times n$ and similarly G_{Cf} is demonstrating the pages that are converged with dimensions$(n-m)\times n$. The standard power method may be represented as in Equation (19). [2]

$$\begin{bmatrix}\pi_{fN}^{(k+1)}\\ \pi_{fC}^{(k)}\end{bmatrix} = \begin{bmatrix}G_{fN}\\ G_{fC}\end{bmatrix}\begin{bmatrix}\pi_{fN}^{(k)}\\ \pi_{fC}^{(k)}\end{bmatrix} \quad \textbf{(19)}$$

As $\pi_{fC}^{(k)}$ is the Fuzz-PageRank vector of the nodes that have already been converged, so no more computation is desirable for them. The proceeding iteration may be demonstrated in Equations (20) and (21). [2]

$$\pi_{fN}^{(k+1)} = G_{fN}\pi_f^{(k)} \quad \textbf{(20)}$$

$$\pi_{fC}^{(k+1)} = \pi_{fC}^{(k)} \quad \textbf{(21)}$$

We can accelerates the rate of converge by extracting non-converged webpages in Fuzz-Adaptive method for further computation. [2] See algorithm 3 and algorithm 4.

Algorithm:3 PageRank by Adaptive Method

Function AdaptivePageRank(G,π^0)
repeat
$\pi_N^{(k+1)}$ = G$_N\pi_N^{(k)}$;
$\pi_C^{(k+1)}$ = $\pi_C^{(k)}$;
[N,C] = detectConverged($\pi^{(k)}$, $\pi^{(k+1)}$, ε);
r = ||G$\pi^{(k)}$ –$\pi^{(k)}$||$_1$;
until r < ε;
return $\pi^{(k+1)}$;

Algorithm:4 Fuzz-PageRank by Adaptive Method

Function AdaptivePageRank(G,π_f^0)
repeat
$\pi_{fN}^{(k+1)}$= G$_N\pi_f^{(k)}$;
$\pi_{fC}^{(k+1)}$=$\pi_{fC}^{(k)}$;
[N,C] = detectConverged($\pi_f^{(k)}$, $\pi_f^{(k+1)}$, ε);
r = ||G$\pi_f^{(k)}$ –$\pi_f^{(k)}$ ||$_1$;
until r < ε;
return $\pi_f^{(k+1)}$;

7.3 Fuzz-PageRank by Aitken's Method:

Aitken's method is used for speed up the convergence rate of iterative methods. [19][21] Aitken's method has been applied to computer PageRank vector with robust acceleration rate and named this vector as Fuzz-Aitken PageRank vector.
Considering $\pi_f^{(k)}$ is achieved after k iterations then approximating it with the fist two π_f , u_2 :

$$\pi_f^{(k)} = \pi_f u_2; \quad \textbf{(22)}$$

Hence:

$$\pi_f^{(k+1)} = G\pi_f^{(k)} = \pi_f + \alpha\lambda_2 u_2; \quad \textbf{(23)}$$

$$\pi_f^{(k+2)} = G\pi_f^{(k+1)} = \pi_f + \alpha\lambda^2 u_2. \quad \textbf{(24)}$$

Signifying g , h as:

$$g_i = \left(\pi_{fi}^{(k+1)} - \pi_{fi}^{(k)}\right)^2; \quad \textbf{(25)}$$

$$h_i = \pi_{fi}^{(k+2)} - 2\pi_{fi}^{(k+1)} + \pi_{fi}^{(k)}. \quad \textbf{(26)}$$

Acheiveing:

$$g_i = \alpha^2(\lambda_2 - 1)^2\left(u_2(i)\right)^2, \quad \textbf{(27)}$$

$$h_i = \alpha(\lambda_2 - 1)^2\left(u_2(i)\right). \quad \textbf{(28)}$$

If $h_i \neq 0$, then express vector f

$$f_i = \frac{g_i}{h_i} = \alpha u_2(i), \quad \textbf{(29)}$$

That is,

$$f = \alpha u_2, \quad \textbf{(30)}$$

Dominant eigenvector is represented as:

$$\pi_f = \pi_f^{(k)} - f. \quad \textbf{(31)}$$

Combination of the preceding iterations are used to acquire a forward approximation of true eigenvector. [19][20]

Algorithm:5 PageRank by Aitken's Method

function y = Aitkenpagerank($\pi^{(k)}$,$\pi^{(k+1)}$,$\pi^{(k+2)}$)
g= ($\pi^{(k+1)}$ – $\pi^{(k)}$)^2;
h=$\pi^{(k+2)}$+2$\pi^{(k+1)}$+$\pi^{(k)}$;
f= g .h ;
y= $\pi^{(k+2)}$–f;
renormalize y;
return

Algorithm:6 Fuzz-PageRank by Aitken's Method

function y = Aitkenpagerank($\pi_f^{(k)}$, $\pi_f^{(k+1)}$, $\pi_f^{(k+2)}$)

g= $(\pi_f^{(k+1)} - \pi_f^{(k)})^2$;
h=$\pi_f^{(k+2)} + 2\pi_f^{(k+1)} + \pi_f^{(k)}$;
f= g .h ;
y=$\pi_f^{(k+2)}$ **–f;**
renormalize y;
return

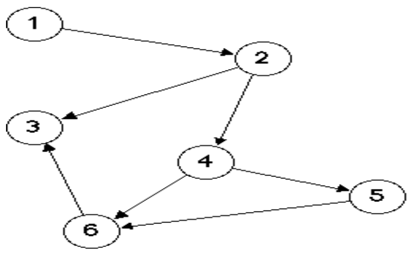

Fig 1 : Six webpage(node) Directed Graph

8. PageRank vs Fuzz-PageRank for the 6-node Directed Graph in Figure 1 by Standard Power Method:

Table 1 At $\alpha = 0.85$ with tolerance level of 0.001.

Fuzz-PageRank vector π_f computed in 5 iterations and 0.000373 seconds as represented in Table 1(a) whereas PageRank was computed in 15 iterations and 0.006129 seconds as represented in Table 1(b).

(a)

Page	Initial value	Iteration 1	Iteration 2	Iteration 3	Iteration 4	Iteration 5
N	$\pi_f^{(0)}$	$\pi_f^{(1)}$	$\pi_f^{(2)}$	$\pi_f^{(3)}$	$\pi_f^{(4)}$	$\pi_f^{(5)}$
1	0.1667	0.1447	0.1452	0.1453	0.1452	0.1452
2	0.1667	0.1650	0.1623	0.1630	0.1628	0.1629
3	0.1667	0.1827	0.1853	0.1844	0.1847	0.1846
4	0.1667	0.1625	0.1627	0.1624	0.1626	0.1625
5	0.1667	0.1625	0.1624	0.1626	0.1625	0.1625
6	0.1667	0.1827	0.1820	0.1823	0.1822	0.1822

(b)

Page	Initial value	Iteration 1	Iteration 2	…	Iteration 14	Iteration 15
N	$\pi^{(0)}$	$\pi^{(1)}$	$\pi^{(2)}$	…	$\pi^{(14)}$	$\pi^{(15)}$
1	0.1667	0.0486	0.0647	…	0.0705	0.0704
2	0.1667	0.1903	0.0859	…	0.1303	0.1303
3	0.1667	0.2611	0.3855	…	0.3204	0.3206
4	0.1667	0.1194	0.1475	…	0.1258	0.1258
5	0.1667	0.1194	0.1114	…	0.1239	0.1238
6	0.1667	0.2611	0.2049	…	0.2292	0.2292

9. Adaptive PageRank vs Fuzz-Adaptive PageRank for the 6-node Directed Graph by Adaptive Method:

Table 2 At $\alpha = 0.85$ with tolerance level of 0.001.

Fuzz-Adaptive PageRank vectorπ_f computed in 5 iterations and 0.000686 seconds whereas Adaptive PageRank was computed in 15 iterations and 0.001058 seconds.

Nodes	PageRank 'π'	Fuzz-PageRank 'π_f'
1	0.0704	0.1452
2	0.1303	0.1692
3	0.3206	0.1846
4	0.1258	0.1625
5	0.1238	0.1625
6	0.2292	0.1822

10. Fuzz-PageRank vs PageRank for the Google matrix of 100 x 100 by Standard Power Method:

Table 3 Google matrix of 100 x 100 for $\alpha = 0.85$ with tolerance value of 0.001.

(a)

Page	Initial value	Iteration 1
N	$\pi_f^{(0)}$	$\pi_f^{(1)}$
1	0.0100	0.0100
2	0.0100	0.0100
3	0.0100	0.0100
4	0.0100	0.0100
5	0.0100	0.0100
.	.	.
.	.	.
.	.	.
96	0.0100	0.0100
97	0.0100	0.0100
98	0.0100	0.0100
99	0.0100	0.0101
100	0.0100	0.0100

(b)

Page	Initial value	Iteration 1	Iteration 2	..	Iteration 6	Iteration 7
N	$\pi^{(0)}$	$\pi^{(1)}$	$\pi^{(2)}$	..	$\pi^{(6)}$	$\pi^{(7)}$
1	0.0100	0.0076	0.0070	..	0.0074	0.0074
2	0.0100	0.0076	0.0070	..	0.0074	0.0074
3	0.0100	0.0076	0.0070	..	0.0074	0.0074
4	0.0100	0.0076	0.0070	..	0.0074	0.0074
5	0.0100	0.0076	0.0070	..	0.0074	0.0074
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
96	0.0100	0.0076	0.0070	..	0.0074	0.0074
97	0.0100	0.0076	0.0070	..	0.0074	0.0074
98	0.0100	0.0076	0.0070	..	0.0074	0.0074
99	0.0100	0.0061	0.0135	..	0.0137	0.0137
100	0.0100	0.0076	0.0070	..	0.0074	0.0074

11. PageRank vs Fuzz-PageRank for the Google matrix of 100 x 100 by Adaptive Method:

Table 4 Google matrix of 100 x 100 for $\alpha = 0.85$ with tolerance value of 0.001.

(a)

Page N	Initial value $\pi_f^{(0)}$	Iteration 1 $\pi_f^{(1)}$
1	0.0100	0.0100
2	0.0100	0.0100
3	0.0100	0.0099
4	0.0100	0.0100
5	0.0100	0.0101
.	.	.
.	.	.
.	.	.
96	0.0100	0.0100
97	0.0100	0.0100
98	0.0100	0.0100
99	0.0100	0.0100
100	0.0100	0.0100

(b)

Page	Initial value	Iteration 1	Iteration 2	Iteration 3	Iteration 4
N	$\pi^{(0)}$	$\pi^{(1)}$	$\pi^{(2)}$	$\pi^{(3)}$	$\pi^{(4)}$
1	0.0100	0.0088	0.0090	0.0089	0.0089
2	0.0100	0.0088	0.0090	0.0089	0.0089
3	0.0100	0.0517	0.0463	0.0470	0.00469
4	0.0100	0.0088	0.0090	0.0089	0.0089
5	0.0100	0.0431	0.0388	0.0394	0.0393
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
96	0.0100	0.0088	0.0090	0.0089	0.0089
97	0.0100	0.0088	0.0090	0.0089	0.0089
98	0.0100	0.0088	0.0090	0.0089	0.0089
99	0.0100	0.0088	0.0090	0.0089	0.0089
100	0.0100	0.0088	0.0090	0.0089	0.0089

12. PageRank vs Fuzz-PageRank for the Google matrix of 100 x 100 by Aitken's Method:

Table 5 Google matrix of 100 x 100 $\alpha = 0.85$ with tolerance value of 0.001.

(a) Fuzz-PageRank vector coverge in 1st iteration by Aitken's method in 0.00098 seconds.

Page N	Initial value $\pi_f^{(0)}$	Iteration 1 $\pi_f^{(1)}$
1	0.0100	0.0100
2	0.0100	0.0100
3	0.0100	0.0099
4	0.0100	0.0100
5	0.0100	0.0101
.	.	.
.	.	.
.	.	.
96	0.0100	0.0100
97	0.0100	0.0100
98	0.0100	0.0100
99	0.0100	0.0100
100	0.0100	0.0100

(b) Whereas PageRank vector converges by Aitken's Method in 65 iterations in 0.0632 seconds.

Page	Initial value	Iteration 1	Iteration 2	…	Iteration 64	Iteration 65
N	$\pi^{(0)}$	$\pi^{(1)}$	$\pi^{(2)}$	…	$\pi^{(64)}$	$\pi^{(65)}$
1	0.0100	0.0097	0.0105	…	0.0201	0.0201
2	0.0100	0.0097	0.0105	…	0.0201	0.0201
3	0.0100	0.0522	0.0520	…	0.1052	0.1052
4	0.0100	0.0097	0.0105	…	0.0201	0.0201
5	0.0100	0.0437	0.0437	…	0.0882	0.0882
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
96	0.0100	0.0097	0.0105	…	0.0201	0.0201
97	0.0100	0.0097	0.0105	…	0.0201	0.0201
98	0.0100	0.0097	0.0105	…	0.0201	0.0201
99	0.0100	0.0097	0.0105	…	0.0201	0.0201
100	0.0100	0.0097	0.0105	…	0.0201	0.0201

Page	Initial value	Iteration 1	Iteration 2	…………	Iteration 64	Iteration 65
N	$\pi^{(0)}$	$\pi^{(1)}$	$\pi^{(2)}$	…………	$\pi^{(64)}$	$\pi^{(65)}$
1	0.0100	0.0097	0.0105	………..	0.0201	0.0201
2	0.0100	0.0097	0.0105	………..	0.0201	0.0201
3	0.0100	0.0522	0.0520	………..	0.1052	0.1052
4	0.0100	0.0097	0.0105	…………	0.0201	0.0201
5	0.0100	0.0437	0.0437	…………	0.0882	0.0882
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
96	0.0100	0.0097	0.0105	…………	0.0201	0.0201
97	0.0100	0.0097	0.0105	…………	0.0201	0.0201
98	0.0100	0.0097	0.0105	…………	0.0201	0.0201
99	0.0100	0.0097	0.0105	…………	0.0201	0.0201
100	0.0100	0.0097	0.0105	…………	0.0201	0.0201

13. Contrasting iterations(Elapsed time) for PageRank and Fuzz-PageRank by Standard Power Method:

Table 6 Google matrix of 100 x 100 for different values of α with tolerance value of 0.001.

α	PageRank	Fuzz-PageRank
0.7	6(0.000694)	1(0.000312)
0.85	7(0.000659)	1(0.000449)
0.9	8(0.000673)	1(0.000305)
0.99	9(0.000683)	1(0.000400)

14. Contrasting iterations(Elapsed time) for Adaptive PageRank and Fuzz-Adaptive PageRank by Adaptive method:

Table 7 Google matrix of 100 x 100 for different values of α with tolerance value of 0.001.

α	Adaptive PageRank	Fuzz-Adaptive PageRank
0.7	6(0.004634)	1(0.000580)
0.85	7(0.001029)	1(0.000551)
0.9	8(0.001104)	1(0.000513)
0.99	25(0.002613)	1(0.000461)

15. Contrasting iterations(Elapsed time) for Aitken PageRank and Fuzz-Aitken PageRank by Aitken's method:

Table 8 Google matrix of 100 x 100 for different values of α with tolerance value of 0.001.

α	Aitken PageRank	Fuzz-Aitken PageRank
0.7	18(0.00900)	1(0.00096)
0.85	65(0.06320)	1(0.00098)
0.9	98(0.09160)	1(0.00097)
0.99	NAN	1(0.00099)

16. Conclusion:

Throughout the world, the titanic of web comprises of numerous webpages that are hooked and deliver data material to surfer visiting webpage. With the passage of time, the webpages are continuously undergoing augmentation and amplification that is leading to complications on part of understanding ability of the surfer. Therefore in order to achieve highly pertinent results to our desired search data, the significance of mathematical mysteries behind the Google search engine cannot be overlooked.

The new technique for the computation of Google PageRank is Fuzz-PageRank approach (FUPRA). In this novel approach, comparison is done between this new technique and the previously used techniques for the analysis of Google PageRank vector. It is observed that there is massive particularity in terms of converging acceleration and computational time in between PageRank vector and Fuzz-PageRank vector. Hence **fuzz**-PageRank vector is superior to PageRank vector as computed by Standard Power method, Adaptive method and Aitken's method. Through fuzzy concept, we have also treated dangling nodes by giving weightage to them known as **Fuzz-dangling nodes** collectively with non-dangling nodes known as **Fuzz-non-dangling nodes.**

17. References:

1. Bhamidipati and Pal, Fuzzy Web Surfer Models: Theory and Experiments; WImBI 2006, LNAI 4845, pp. 324–340, 2007.9
2. Kamvar, Haveliwala and Golub, Adaptive methods for the computation of PageRank, Linear Algebra and its applications,(2004),pp 51-65.
3. Yan, Gui, Du and Guo, An Improved PageRank Method based on Genetic Algorithm for Web Search, Procedia Engineering 15 (2011), pp 2983-2987
4. Dubey and Roy, An Improved Page Rank Algorithm based on Optimized Normalization Technique, (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (5) (2011), 2183-2188
5. Vajargah and Gharehdaghi, Ergodicity of Fuzzy Markov Chains Based on Simulation Using Halton Sequences, The Journal of Mathematics and Computer Science Vol. 4 No.3 (2012) 380 – 385
6. Kandiah and Shepelyansky, PageRank model of opinion formation on social networks, Physica A 391 (2012), pp 5779-5793
7. Tabrizi, Shakery, Asadpour, Abassi and Tavallaie, Personalized PageRank Clustering: A graph clustering algorithm based on random walks, Physica A 392 (2013), pp 5772-5785
8. Garcia, Pedroche and Romance, On the localization of the personalized PageRank of complex networks, Linear Algebra and its Applications 439 (2013), pp 640-652
9. Bourchtein and Bourchtein, On some analytical properties of a general PageRank algorithm, Mathematical and Computer Modelling 57 (2013), pp 2248-2256
10. Sharma and Gupta, A Comprehensive Study of Fuzzy Logic, International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 3 (2) ,(2013), pp 398-403
11. Koumenides and Shadbolt, Ranking Methods for Entity-Oriented Semantic Web Search, Journal of the Association for Information Science and Technology 65(6), (2014), pp 1091-1106
12. Zadeh, Fuzzy sets, Information and Control 8 (1965),pp 338–353
13. http://whatis.techtarget.com/definition/fuzzy-logic
14. Chen and Pham, Introduction to Fuzzy Sets, Fuzzy logic, and Fuzzy Control Systems, (2001)

15. http://www.myreaders.info/07_Fuzzy_Systems.pdf
16. http://www.bindichen.co.uk/post/AI/fuzzy-inference-membership-function.html
17. Avrachenkov and Sanchez, Fuzzy markov chains and decision-making, Fuzzy Optimization and Decision Making 1(2), (2002), pp 143-159
18. Willis S.R, Google's PageRank: The Math Behind the Search Engine (2006), pp 1-15
19. Kamvar, Haveliwala, Manning and Golub, Extrapolation Methods for Accelerating PageRank Computations, In Proceedings of the Twelfth International World Wide Web Conference (2003)
20. Sun and Wei, A note on the PageRank algorithm, Applied Mathematics and Computation Vol. 179 (2), (2006), pp 799-806.
21. https://mat.iitm.ac.in/home/sryedida/public_html/caimna/transcendental/iteration%20methods/accelerating%20the%20convergence/aitken.html
22. Fodor and Bede, Arithmetics with Fuzzy Numbers: a Comparative Overview,