Journal of Information Hiding and Multimedia Signal Processing ©2015 ISSN 2073-4212
Ubiquitous International Volume **6**, Number **4**, July **2015**

A Matrix-Geometric Method for Web Page Ranking Systems

Thi Thi Zin[1], Pyke Tin[1,2], Hiromitsu Hama[3], and Takashi Toriu [2]

[1]Faculty of Engineering, University of Miyazaki, Miyazaki, Japan
thithi@cc.miyazaki-u.ac.jp
[2]Graduate School of Engineering, Osaka City University, Osaka, Japan
[3]R&D Center of 3G Search Engine, Incubator, Osaka City University, Osaka, Japan

Received December, 2014; revised January, 2015

Abstract. *An algorithmic approach which is known as Matrix-Geometric method is an efficient and popular in solving various types of stochastic models especially for queueing and water storage models. This paper proposes a modified Matrix-Geometric method for calculating random web page ranks in World Wide Web (WWW) by considering web page ranking system as a stochastic model in random environments. Since the Matrix-Geometric approach is very powerful in variety of stochastic models, it may develop new promising directions in Web Page Ranking Systems and related research areas. In order to do so as first step, a web page ranking system is modeled in the frameworks of queueing models. Then an attempt is made to reduce the obstacles occurred in the problems of web page ranking systems. Similarly, in the second step, a Web page ranking system is modeled as a framework of stochastic water storage theory to derive a list of Web page rankings by using Matrix-Geometric method. Some comparison results are presented to confirm the efficiency of the proposed methods. The experimental results shows the proposed approach is promising for establishing a new research area which can improve the current situations and difficulties occurred in search engines and their ranking systems in particular and some problems in WWW as a whole.*
Keywords: Matrix-Geometric, Web page ranking, queueing system, water storage, random environment

1. **Introduction.** In these days, we are living in the world of information where the information are stored, retrieved and utilized for variety of purposes and by millions of users. In this aspect, WWW is the biggest collection of information and it is growing continuously. Since the web has dynamic nature, the web pages and documents are frequently added and deleted so that the importance of information is up and down all the times. So the web becomes the major source of meaningful information related to query made by the user. It is also known as easily accessible and searchable consisting billions of interconnected web pages developed by millions of people. Figure 1 shows the structure of a classic Web graph. Structure of a classic Web graph can be seen as web pages as nodes, and hyperlinks as edges which connects two related nodes. The web serves as the key resource of meaningful information depending on users request. In this context, searching key information based on their degree of importance with respect to relevancy, reliability and popularity plays a major role for today search engines.

Most of existing search engines applies different methods for analyzing web page links to retrieve the most relevant pages which are presented at the top of the result list. To make the search results easy accessible to users, ranking methods are applied on the search results. A tremendous amount of ranking algorithms have been proposed in the literature by different researchers. Among them the most popular ones are PageRank (PR) [1],

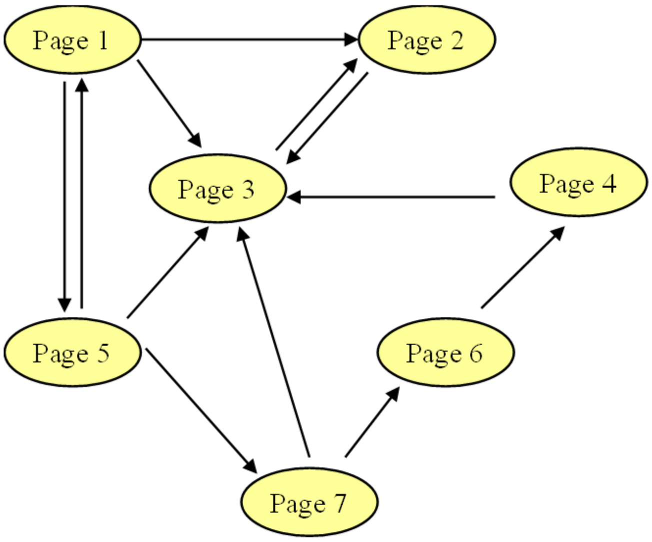

FIGURE 1. The wiring diagram of S3C6410 and DM9000

Weighted PageRank (WPR) [2], Hyperlink-Induced amount of Topic Search (HITS) [3-5], Page Ranking Algorithm Based On Number Of Visits Of Links Of Web Page [6-7], An Improved Page Ranking Algorithm Based On Optimized Normalization Technique [8], Improved method for computation of Page Rank [9]. Most of existing web ranking algorithms are based on hyperlink analysis on Web page graphs. In particular, the key idea is to consider hyperlinks from a source page to sink pages as an influential factor on ranking sink pages. Based on this concept a lot of excellent research woks by various researchers have been proposed to improve the existing web page ranking systems apart from the ones mentioned in the above [10-13].

But different methods have utilized different concepts to analyze Web link graphs [14-15]. For example Googles Page Rank model assumes that the hyperlinks between the Web pages can be approximated as the transition probabilities of a Markov Chain of lag 1 [16]. On the other hand, HITS algorithm observes a Web graph as bipartite forms. Even more some researchers have introduced community ranking systems by using social network platforms based on popularity, reliability and relevancy measures [17-18]. Thus, it is learnt that various approaches have been appeared to measure the Web Graph by using different aspects of quality measures such as relevancy, reliability and popularity concepts. The success of Google has made the role of Page Rank significance as a metric measure for Web pages. This has also led to a variety of modifications and improvisations of the basic PageRank metric. They have either focused on changing the underlying model or on reducing the computation cost.

On the other hand the rise of giant social networks such as Facebook, Twitter and YouTube and so on has influenced on the ways how users are seeking information. In this aspect, some researchers have introduced a concept of social community ranks and combined them with contents of web pages to form an integrated ranking system [19-20]. But they have said that even though their approach is new, it is only at infant stage so that it remains more research has to be done in this direction. However, the original Page Rank algorithm is quite popular and has led many excellent research works, the authors of this paper feel that it is worthwhile to look into to the Page Rank algorithm from new perspectives. Therefore, this paper explores a new look into the original Page Rank algorithm from the perspectives of Matrix-Geometric method in queuing theory and Dam theory which had been introduced by P.A P Moran [21] later extended by J. Gani [22] and R. M. Phatarfod [23].

2. **Search Engine and Web Page Ranking System.** The simplest way of explaining a search engine is to be elaborate as a kind of system which gives response in relation with information which represented by web pages for each users request. Generally speaking, Google, Yahoo, MSN and other search engines keep their billions of web pages indexed and stored on their servers across the world to serve billions of users. When a user shows up a query in the form text or image the search engine identifies several web pages having related content that satisfies the keywords of users queries. Search engines are evaluated on the basis of two quantities such as the number of web pages that are indexed and the usefulness of the ranking process that determines which pages are returned. For those purposes the search engines use different algorithms to perform link analysis and the results are presented according to their rank scores.

In the mechanism of basic search engine (see Fig. 2) output results are listed according to

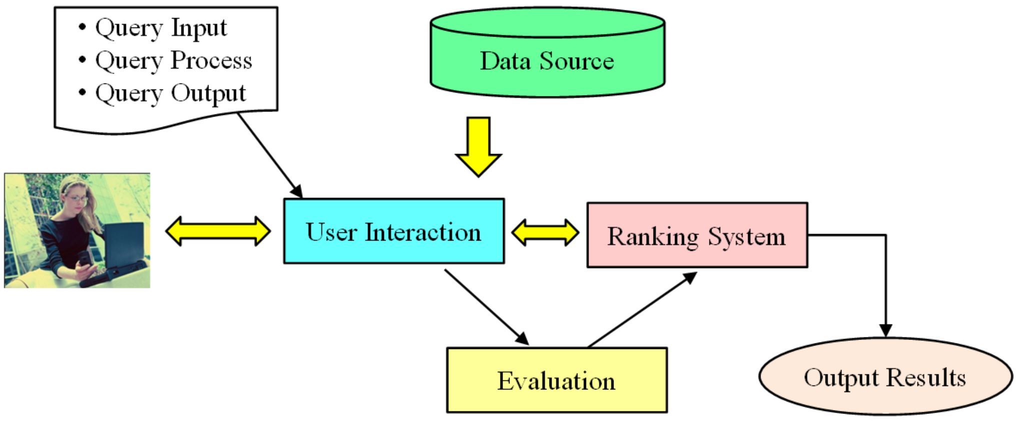

FIGURE 2. Mechanism of Basic Search Engine

their ranking scores. It has also been well recognized that ranking mechanism is governed by a probability distribution [24-25]. Specifically, we first define the probability that a web page w is relevant for a given query q, Pr (w| q). By using Bayes rule we have

$$Pr(w|q) \ = Pr(w)Pr(q|w)/Pr\,(q) \tag{1}$$

From web page ranking perspective views, since the term Pr(q) does not depend on the web page w, the term Pr(q| w) becomes one of the central roles for the information search engines to evaluate the importance of a web page in the web graph. Also this simple formula had led to the most popular web page ranking system, Page Rank algorithm and variety of other ranking systems for information search systems. Thus the simplest version of Page Rank can be considered as the stationary distribution of a random walk in the web graph. Also the random walk model for Page Rank can be formulated as follows. Suppose that Ri denotes a random variable representing the rank score of web page i=1, 2, 3, , N where N is the total number of pages in the web graph.

From web page ranking perspective views, since the term Pr(q) does not depend on the web page w, the term Pr(q| w) becomes one of the central roles for the information search engines to evaluate the importance of a web page in the web graph. Also this simple formula had led to the most popular web page ranking system, Page Rank algorithm and variety of other ranking systems for information search systems. Thus the simplest

version of Page Rank can be considered as the stationary distribution of a random walk in the web graph. Also the random walk model for Page Rank can be formulated as follows. Suppose that R_i denotes a random variable representing the rank score of web page i=1, 2, 3, , N where N is the total number of pages in the web graph.

Let the first step transition probability that the random walk jumps from page i to page j be p_{ij} and $\Pr(R_i = k) = q_{ik}$ for i, k =1,2,3, ... This formulation gives us

$$q_{ik} = (1-\alpha)\sum_{i\in w_j} p_{ij}q_{ik} + \alpha \sum_{i\in w} \frac{q_{ik}}{N} \tag{2}$$

Traditionally, the value of α is chosen as 0.15, and it appears that this value provides reasonable ranking for the web pages. Let us write equation (2) in matrix form. Define

$$Q = \begin{bmatrix} q_{11} & q_{12} & q_{13} & \cdots & q_{1N} \\ q_{21} & q_{22} & q_{23} & \cdots & q_{2N} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ q_{N1} & q_{N2} & q_{N3} & \cdots & q_{NN} \end{bmatrix}$$

$$H = (1-\alpha)\begin{bmatrix} q_{11} & q_{12} & q_{13} & \cdots & q_{1N} \\ q_{21} & q_{22} & q_{23} & \cdots & q_{2N} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ q_{N1} & q_{N2} & q_{N3} & \cdots & q_{NN} \end{bmatrix} + \alpha/N \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}$$

Then the equation (2) can be written as

$$\mathbf{Q}{=}\mathbf{HQ} \tag{3}$$

The matrix **H** is known as Google matrix. This matrix is stochastic (each row sums to 1), irreducible (all pages are connected due to the teleportation jump), aperiodic ($\mathbf{H}_{ii}$>0), and primitive $\mathbf{H}^r > 0$), which implies that a unique positive **Q** exists and Matrix-Geometric method guarantees to converge to this vector. Given some initial distribution $\mathbf{Q}(0)$, e.g.

$$\mathbf{Q}(0) = [e_{ij}] \text{ where } e_{ij} = 1 \text{ for all i, j}$$

then the MatrixGeometric method may lead an iteration procedure:

$$\mathbf{Q}(\mathrm{r}) = \mathbf{H}\ \mathbf{Q}(\mathrm{r1})\ ,\ \mathrm{r} \geq 1$$

Note that the limiting distribution does not depend on the initial distribution Q(0). Then, the number of iterations that is needed to achieve ϵ- accuracy of the order t = log(ϵ-)/log(c) independent of the underlying graph structure where c is a constant.

Since we consider the Page Rank of a random web page at time t as the random variable R_i(t) for t = 1, 2, 3,... Then q_{ik} can be interpreted as the probability of $R_i(t) = k$ for all i. Then the equation (2) becomes a stochastic equation which has been occurred in queues, dams and branching process theories.

One of the objectives of this paper is that to provide better understanding on what the terms Pr(d), Pr(q|d) in equation (1) and the terms q_{ik} for i, j=1,2,, N look like, and how they related to theory of queues and theory of water storage along with Matrix-Geometric method for opening up a new research in web search engine frameworks. To achieve this goal, the importance of a web page will be defined for a search engine based on secret combination of rankings according to different criteria, such as relevancy, reliability and popularity of the pages in the web graph.

In particular, the remaining part of this paper is arranged as follows. In section 3, a Page Rank in the framework of queues and dams is presented. Then, Matrix-Geometric solutions for Web Page Ranks are derived in section 4. Some simulation works are shown in section 5 followed by concluding remarks in section 6.

3. **Page Rank in Queue and Dams Framework.** The matrix-geometric method introduced by [26] has been successfully applied in analyzing many queueing systems and water storage systems (dams). It is well recognized that both queuing and dam models are fallen into a class of Random Walk type models in which following recurrence relationship is fundamental and plays an important role.

$$W_{t+1} = \max[W_{t+1} - X_t, 0] \quad \text{for } t = 1, 2, 3, \ldots \tag{4}$$

There are a variety of ways to interpret and make assumptions for the processes $\{W_t\}$ and $\{X_t\}$. For example in theory of queues, X_t represents the arrival patterns process for new customers coming into a queue along with other processes such as service time distribution and service mechanism. On the other hand, from the dam theory aspect, the random variable X_t is the amount of water inflows into a dam during a time interval $(t, t+1)$ for $t = 0, 1, 2, ...$ Keeping these concepts in mind, the random variable X_t shall be considered as the amount of incremental ranks added to a random web page due to in-degrees. In all cases, W_t will be the number of customers who are waiting for service including the one being served. On the other hand, W_t will represent the depletion from the maximum level, say *K-1* where K is the capacity of a finite dam. In terms of web rank, K can be considered as the maximum rank score for a random web page can reach in the web graph which has N web pages in total. In this case, W_t will represent the level of page ranks at time t. It is this analogy to be exploited in analyzing Web Page Ranking Systems by using matrix-geometric method. Since the queueing and dam theory have a variety of analytical results for W_t by taking independent and dependent structure of input process into account, this paper will explore and examine a new way of formulating Web Page Ranking by using the matrix-geometric method. The overview of proposed system is described in Fig. 3.

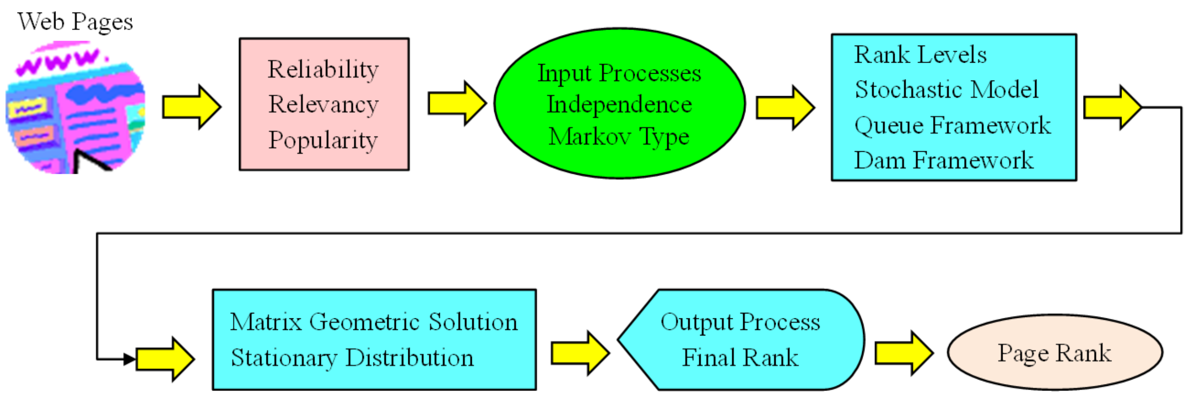

FIGURE 3. Overview of proposed system

4. **Matrix-Geometric Model for Web Page Rank.** In this section, we will introduce and analyze the Matrix-Geometric model Web Page Ranking system. The Matrix-Geometric Method (MGM) contains two parts: the initial part and the repetitive part. The initial part has a non-regular structure and each component in it is to be represented in detail. On the other hand, the repetitive part has a regular structure which can be represented as a composition of several components in stochastic process algebra. In the Web Page Ranking System, the two parts can be interpreted as follows. Typically, we

can assume that there exists an integer k such that from k onwards, the behavior of the system that is the level of web page ranks, is the same as the behavior of the system at k but it is not necessary to be the same as that of *0,1,2,..., k-1*.Then, the initial part is represented by the system at *0,1,2,..., k-1* and the repetitive part is represented by *k, k+1, k+2,...* After that the two parts form a generator matrix which can be decomposed into sub-matrices. Each sub-matrix represents the transition rates in the particular area within the given part or between the parts. Thus the size of the state space is reasonably small compared to the size of a generator matrix of a Markov process. Through solving the system of sub-matrices by using MGM we obtain the steady state probability distribution of Web Page Ranking system which can further compute the performance measure of the proposed model.

A.Independent Type In-Degree Case. In this case, $\{X_t\}$ forms a sequence of independent and identically distributed random variables in queues and dams while it will represent rank increments due to in-degrees pointing the page considered. It is also assumed that the probability generating function the random variable X_t is

$$g(s) = \sum_j \Pr(X_t = j)s^j$$

Generally speaking, a problem of interest to hydrologists, queue controllers and web search engines optimizers is finding the limiting distribution of the process $\{W_t\}$.

Let $Pr(W_t = j) = \pi_j$ when $t \to \infty$. In queues and dams, when $\mu = g'(0) > 0$ the limiting distribution π_j is given by

$$\pi_j = (1-\theta)\theta^j \tag{5}$$

where θ is the smallest positive root of $g(s) = s$. This leads to the results

$$\pi_j = \frac{(1-\theta)\theta^{K-1-j}}{1-\theta^K} \quad \text{for } j = 0, 1, \ldots, K-1 \tag{6}$$

where k is the capacity of a dam or size of waiting space for customers.

The probability in (6) is generally used to estimate the size of a dam for reliability factor P as shown in (7).

$$K \cong \frac{1+\log P}{\log\theta} \tag{7}$$

By using the analogy between Dams and Page Rank, the Page Rank of a random page can be approximated as

$$PR(j) \cong \frac{(1-\theta)\theta^{N-1-j}}{1-\theta^N} \quad \text{for } j = 0, 1, \ldots, N-1 \tag{8}$$

where the number of pages in the web graph is estimated for reliability factor P as

$$N \cong \frac{(1-\theta)\theta^{N-1-j}}{1-\theta^N}. \tag{9}$$

B. Markov Type In-Degree Case. In queueing theory framework, it has been shown that the queue length process embedded in a general arrival process such as moving average [20] and Gamma Markov [21] has a modified geometric distribution similar to those described in (6) and (7).

Let the sequence X_t a Markov chain with transition probability r_{ij} for i,j=0,1,...N-1

Let $Z_{ij} = s^i r_{ij}$ and $\left[Z_{ij}^n\right] = [Z_{ij}]^n$ Then it can be proved that $\lambda(s) = \lim \left(Z_{jj}^n\right)^{1/n}$ when $n \to \infty$ exists and independent of j.

By assuming that $0 < \theta < 1$ is a solution of $\lambda(s) = s$, the limiting distribution of $\{Wt\}$ Moreover for the case of special type of Markov Chain such that

$$\sum r_{ij} s^j = B(s)[A(s)]^i \tag{10}$$

where $B(s) = [1 + a(1-\rho)a(1-\rho)s]^{-1}$ and $A(s) = \frac{[(1+a)(1-\rho)-(a-(1-\rho)-\rho)s]}{1+a\,(1-\rho)-a\,(1-\rho)s}, 0 < \rho < 1$ is the correlation coefficient of lag 1 and a is the mean value of $\{X_t\}$. Pakes And Phatarfod [22] had proved that the limiting distribution of $\{W_t\}$ is as shown in equation (11)

$$\pi_0 = \lim_{t\to\infty} \Pr(W_t = 0) = 1 - \frac{1+q}{1+a}, \tag{11}$$

$$\pi_j = \lim_{t\to\infty} \Pr(W_t = j) = 1 - \frac{1-q^2}{1+a} q^{j-1} \quad \text{for } j \geq 1. \tag{12}$$

With $q = (2\rho)^{-1}[-a(1-\rho) + (a(1-\rho^2)^2 + 4\rho)^{1/2})]$. By using equations (11) and (12), the page rank distribution is given by

$PR(0) = \frac{a-q}{(1+a)-(1+q)q^{N-1}}$,

$PR(j) = \frac{(1-q^2)q^{N-j-1}}{(1+a)-(1+q)q^{N-1}}$ for $j = 1, 2, \ldots, N-1$.

5. **Numerical Results.** In order to analyze the variations of in-degree distributions effects on the Page-Ranking system we consider three types of distributions in independent case and three in Markov Chain case. In independent case, three distributions are geometric, binomial and Erlangian. For Markov Chain case we have considered geometric, binomial and negative binomial by varying the parameters involved. For Erlangian case, we have utilized the following parameter settings. The probability density of Erlangian type is given by the following equation:

$$f(x) = \frac{\alpha^\gamma e^{-\alpha x} x^{\gamma-1}}{\Gamma_\gamma} \text{for } x > 0,\ \alpha > 0,\ \gamma = \text{ integer.} \tag{13}$$

For $\gamma = 2, \alpha = 1.4$, the discrete probabilities are given as shown in the following Table 1

Table 1. The discrete probabilities (for $\gamma = 2, \alpha = 1.4$)

g0	g1	g3	g4	g5	g6	g7	g8	g9	>9
0.1558	0.4646	0.2437	0.030	0.0095	0.0028	0.0008	0.0002	0.0001	0.4082

In this case the stationary distribution of Page Rank is given by as shown in Table 2.

Table 2. Page Rank (Erlangian Case)

i	0	1	2	3	4	5	6
N large	0.4059	0.7870	0.9297	0.9726	0.9902	0.9902	0,9987
N small	0.4067	0.7886	0.9255	0.9746	0.9921	0.9984	1

It can be seen that, by varying parameter setting there are significant effects on the distribution of Page Ranks especially in Markov Chain case. Thus we would like to point out that the distribution of in-degrees should be carefully taken into account. Further this finding can give why the original Page Ranking algorithm and its extensions fail in heavily correlated or dense in-degree web graphs. More investigations will be needed for revealing the true insight of Page Rank algorithms and their extensions.

Table 3. Page Rank (Two Binomial Cases)
$\Pr[X = j] = (\gamma + i - 1)C_i a^\gamma b^i \quad \text{for } 0 < a < 1,\ b = 1 - a.$

Page Rank Pr {W= i}		0	1	2	3	4	5	6	7	8	9
$\gamma = 2$,	N large	0.3229	0.541	0.695	0.789	0.857	0.904	0.935	0.956	0.97	0.98
a = 0.6	N small	0.325	0.55	0.7	0.802	0.871	0.918	0.95	0.971	0.985	1
$\gamma = 1.5$,	N large	0.389	0.627	0.772	0.861	0.915	0.948	0.968	0.98	0.988	0.992
a = 0.5	N small	0.391	0.63	0.776	0.865	0.92	0.952	0.973	0.986	0.993	1

Table 4. Page Rank (Markov Chain Case)
$\Sigma r_{ij} x^j = h(x)[f(x)]^I, h(x) = [1 + a(1-\rho)a(1-\rho)x]^{-\gamma}$
$f(x) = \frac{(1+a)(1-\rho)\{a(1-\rho)-\rho\}x}{1+a\,(1-q)-a\,(1-\rho)x}$

x		0	1	2	3	4	5	6	7	8	9
Geometric g =1, a = 1.5, ρ =0.3		0.288	0.446	0.568	0.664	0.738	0.796	0.841	0.876	0.903	0.925
Binomial g=2,a =0.7, ρ=0.5	N Large	0.72	0.836	0.904	0.944	0.967	0.98	0.988	0.993	0.996	0.998
	N small	0.721	0.837	0.906	0.945	0.969	0.982	0.99	0.995	0.997	1

6. **Conclusions.** In todays fast growing world, to get accurate information is more important and necessary especially in terms of ranking web pages with respect to in-degree distributions. In particular, this paper concludes that Page Raking needed more powerful sophisticated techniques which include maximum relevancy and correlated in nature for in-degree assumptions from users perspectives. More numerical as well as analytical analysis should go in harmony.

Acknowledgments. This work is partially supported by KAKENHI 25330133 Grant-in-Aid for Scientific Research(C).

REFERENCES

[1] S. Brin and L. Page, *The anatomy of a large-scale Hyper-textual Web Search Engine*, Computer Networks, vol. 30, no.1-7 , pp. 107117, 1998.
[2] W.Xing and A.Gorbani, *Weighted PageRank Algorithm*, 2nd Conf. on Communication Networks and Services Research, pp. 305-314, 2004.
[3] M. Kleinberg, *Authoritative Sources in Hyperlinked Environment*, 9th Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 668-667, 1998.
[4] C. Ding, X. He, H. Zha, P.Husbands and H. Simon, *Link Analysi of Hubs and Authorities on the World Wide Web*,SIAM review, vol. 46, no. 2, pp. 256-268, 2004.
[5] J. M. Kleinberg, *Authoritative sources in a hyperlinked environment*, Journal of the ACM, vol. 46, no. 5, pp. 604.632, 1999.
[6] G.Kumar, N. Duhan and A.K. Sharma, *Page Ranking Based on Number of Visits of Web Pages*, Intl. Conf. on Computer and Communication Technology (ICCCT), pp. 11-14, 2011.

[7] N. Tyagi and S. Sharma, *Weighted Page Rank Algorithm Based on Number of Visits of Links of Web Page*, Intl. Journal of Soft Computing and Engineering (IJSCE), ISSN, pp. 2231-2307, 2012.
[8] H. Dubey and Prof. B.N. Roy, *An Improved Page Rank Algorithm Based on Normalization Techniques*, IJCSIT, pp. 2183-2188, 2011.
[9] W. Huang, B. Li, *An Improved Method for the Computation of PageRank*, Intl. Conf. on Mechatronic Science, Electric Engineering and Computer, pp. 19-22, 2011.
[10] P. Desikan, N. Pathak and J. Srivastava, *Incremental Page Rank Computation on Evolving Graphs*, Special interest tracks and posters of the 14th intl. conf. on World Wide Web, ACM, 2005.
[11] J. OMadadhain and P. Smyth, *EventRank: A framework for ranking time-varying networks*, LinkKDD, ACM, pp. 916, , 2005.
[12] F. David, G. Ryan and A. Rossi, *A Dynamical System for PageRank with Time-Dependent Teleportation*, Purdue University, vol. 20, 2012.
[13] N. Eiron, K. McCurley and J. Tomlin, *Ranking the Web Frontier*, Proc. Conf. 13th on World Wide Web, ACM, pp. 309-318, 2004.
[14] M. Richardson and P. Domingos, *The intelligent surfer: Probabilistic combination of link and content information in pagerank*, Advances in Neural Information Processing Systems, 2002.
[15] T. Haveliwala, *Topic-Sensitive PageRank*, Proc. of Conf. Intl. the 11th World Wide Web, 2002.
[16] L. Page, S. Brin, R. Motwani and T. Winograd, *The PageRank Citation Ranking: Bringing Order to the Web*, Bringing Order to the Web.Technical Report.Stanford Digital Library Technologies, 1996.
[17] P. Tin, T. T. Zin, T. Toriu, H. Hama, *A Cluster Based Ranking Framework for Multi-Typed Information Networks*, Proc. of 10th Intl. Conf. on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP2014), Kita Kyusyu, Japan, pp. 27-29, 2014.
[18] T. T. Zin, P. Tin, T. Toriu, H. Hama, *A Stochastic Model for Measuring Popularity and Reliability in Social Network Systems*, Proc. of IEEE Intl. Conf. on Systems, Man, and Cybernetics (SMC 2013), Manchester, UK, pp. 462-467, 2013.
[19] P. Tin, T. T. Zin, T. Toriu, H. Hama, *An Integrated Framework for Disaster Event Analysis in Big Data Environments*, Proc. of 9th Intl. Conf. on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP2013), Beijing, China, pp. 255-258, 2013.
[20] H. Hama, T. T. Zin and P. Tin, *A Hybrid Ranking of Link and Popularity for Novel Search Engine*, Intl. Journal of Innovative Computing, Information and Control: Special Issue on Multimedia Processing and Network Technologies, Vol. 5, No. 11(B), pp. 4041-4049, 2009.
[21] Moran, P.A.P., *A probability theory of dams with a continuous release*, Quart. J. of Math., vol. 7, no. 2, pp. 130-137, 1956.
[22] J. Gani, *Problems in the probability theory of storage systems*, J. Roy. Statist. Soc. B 14, pp. 181-207, 1957.
[23] P. Tin and R. M. Phatarfod, *On Infinite Dams with Inputs Forming A Stationary Process*, J. Appl. Prob. 11, pp. 553-561, 1974.
[24] D. M. Pennock, G. W. Flake, S. Lawrence, E. J. Glover and C. L. Giles, *Winners dont take all: Characterizing the competition for links on the Web*, Proc. of the National Academy of Sciences, vol. 99, no.8, pp. 5207, 2002.
[25] W. Kraaij, T. Westerveld and D. Hiemstra, *The importance of prior probabilities for entry page search*, Proc. of the 25th Annual Intl. ACM SIGIR Conf. on Research and development in information retrieval (SIGIR 02:), pp. 27-34, ACM, New York, USA, 2002.
[26] M. F. Neuts, *Matrix-geometric solutions in stochastic models*, Johns Hopkins University Press, Baltimore, MD, 1981.