SCIENCE FORUM (JOURNAL OF PURE AND APPLIED SCIENCES) 21 (2021) 601 – 607

http://dx.doi.org/10.5455/sf.99567

Science Forum (Journal of Pure and Applied Sciences)

journal homepage: www.atbuscienceforum.com

An alternative design of the PageRank algorithm for web page optimization using an improved Max–Min ant system

Fatima Umar Zambuk[1], Abdulsalam Yau Gital[1], Souley Boukari[1], Haruna Chiroma[2], Abubakar Umar[1]

[1]*Department of Mathematical Science, Abubakar Tafawa Balewa University, Bauchi, Nigeria*

[2]*Department of Computer Science, Federal College of Education Gombe, Gombe, Nigeria*

ABSTRACT

The rapid growth and the dynamic nature of the web have continued to bring challenges to the PageRank algorithm. The PageRank algorithm remains the best ranking criteria so far. Many researchers have worked to accelerate PageRank computation, but these two challenges still remain an issue. We propose a model architecture that optimizes the dynamic nature of ant colony optimization to prune the cones of graph nodes and retrieve relevant nodes for users' display. The potential benefit of our proposed design indicates an improvement over the traditional PageRank in case of its computation cost and efficiency.

ARTICLE INFO

Article history:
Received 20 July 2021
Received in revised form 30 July 2021
Accepted 01 August 2021
Published 12 October 2021
Available online 12 October 2021

KEYWORDS

PageRank
Ant colony optimization
Max–Min ant system
Improved Max–Min ant system
JGraph framework
Dynamic PageRank model

1. Introduction

The PageRank algorithm has become a very important means of retrieving information efficiently from the massive web data. PageRank has become an important ranking method due to its ability to fight spam. A page is important if the pages pointing to it are important. It is not easy for web page owners to add in links to his/her page from other important pages. Thus, it is not easy to influence PageRank. It is a global measure and is query independent, i.e., the PageRank values of all the pages on the web are computed and saved offline rather than at the query time (Li et al., 2015; Lu et al., 2004; Wang and DeWitt, 2004). Since the explosive growth of web data, the traditional PageRank model has become inefficient to catch up with the growing step of computing PageRank globally on time. The performance of PageRank was looked upon in order to improve it. Various accelerating techniques were used to better improve PageRank performance, for example, aggregation, extrapolation, numerical iterative methods, etc. (Wu et al., 2013).

With the web's continued rapid growth in the amount of data, the traditional PageRank accelerated methods are increasingly insufficient to retrieve relevant information that people need. On the one hand, accelerated PageRank algorithms processing capacity is limited. Thus, PageRank has found a wide range of applications in a variety of domains within computer

*** Corresponding author** Fatima Umar Zambuk ✉ fzambuk2001@yahoo.com ✉ Department of Mathematical Sciences, Abubakar Tafawa Balewa University, Bauchi.

science such as distributed networks, data mining, web algorithms and distributed computing (Bianchini et al., 2005; Brin and Page, 1998; Chien et al., 2004; Fujiwara et al., 2013). Since PageRank vector or PageRank is essentially the steady-state distribution or the top eigenvector of the Laplacian corresponding to a slightly modified random walk process, it is an easily defined quantity. Recently in search of finding efficient ways to compute PageRank, researchers has resolved to genetic programming (Buzzanca et al., 2018) and software computing (Stuzle and Hoos, 2000) in order to catch up with the rapid growth

In this paper, we propose a model design of an efficient PageRank computing algorithm for web optimization using an improved Max–Min ant system on a map-reduced framework that can efficiently retrieve relevant web pages within a stipulated time to cut computational cost and improve the performance of the PageRank algorithm.

The rest of the paper is design as follows. Section 2 discusses the literature review. Section 3 gives the background of the study. Section 4 talks about the Max–Min ant system with it algorithm and the improved Max–Min ant system algorithm. Section 5 discusses a map-reduced framework and in Section 6 we explain the proposed model of our work. Section 7 are discusses about the benefit of the proposed work and Section 8 gives the conclusion of the work.

2. Literature Review

PageRank has emerged as a dominant link analysis model partially due to its query independence, its virtual immunity to spamming and Google's huge business success (Lu et al., 2004). The PageRank form is a probability distribution over web pages; it is calculated using a simple iterative algorithm, and corresponds to the principal eigenvector of the normalized link matrix of the web. Solving this will resolve into a Surfers model, well understood as the Markov process. The size of the adjacency matrix tends to increase as the World Wide Web continues to grow enormously.

As the size of web increases by the day, search engines are challenged to devise a way of retrieving relevant information for a query, if not all type of insignificant information and spam will be retrieved as a result for a query. Some of the literature studies were reviewed (Bahmani et al., 2010; Barroso et al., 2003; Bianchini et al., 2005; Raghavan and Garcia-Molina, 2003) which provided details of storage schemes of the Markov matrix. Raghavan and Garcia-Molina (2001) provided possible implementation of the power method to the adjacency matrix. Brin and Page (1998), Kamvar et al. (2003), Raghavan and Garcia-Molina (2003), and Stewart (1999) suggested techniques for compressing data. (Ding et al. (2001), Haveliwala (2012), and Wu and Louiqa (2009) proposed an I/O-efficient implementation of PageRank. The above-mentioned researchers tried to look at the storage scheme of the gigantic matrix that the web has.

Other researchers tried to improve the rate of convergence of PageRank. Gambardella and Dorigo (1996) discuss the uniqueness, existence, and convergence for several link analysis algorithms. Another group of researchers suggested different accelerating techniques in order to make PageRank rise up to the challenges of massive data in the web (Bahmani, 2010; Buzzanca et al., 2018; Chen et al., 2002; Corne et al., 1999; Dorigo and Gambardella, 2000; Gang et al., 2013; Haveliwala , 2002; Ivkovic et al., 2019; Kreger, 2003; McSherry, 2012; Moler, 2004; Pandurangan et al., 2002; Raghavan and Garcia-Molina, 2001; Wu and Louiqa, 2009).

Wenlei et al. (2015) proposed two strategies to improve the efficiency of the Arnoldi-type algorithm. Anuradha et al. (2014) proposed a framework of an exact solution and an approximate solution for computing ranking on a subgraph. Singh and Sharma (2016) proposed a dynamic system that captures changes to the network centrality of nodes as external interest in those nodes varies. Anuradha et al. (2014) proposed the first fast algorithm for computing PageRank on general graphs when the edge weights are personalized.

57 and 58 proposed the ant colony system in modeling PageRank to obtain relevant web pages from the massive structure of the web.

3. Background of the Study

The PageRank algorithm was invented by Chien et al. (2004), students studying at Stanford University; it is a model based on the hyperlink structure of the web. The PageRank algorithm became a success but later they resigned and formed a company which became a famous search engine company, Google. The PageRank algorithm is the most frequently used algorithm for ranking billions of web pages. Functioning of the PageRank algorithm depends upon link structure of the web pages. The PageRank algorithm is based on the concepts that if a page surrounds important links towards it, then the links of this page near the other page are also to be believed to be imperative pages. The PageRank imitates on the back link in deciding the rank score. Thus, a page gets hold of a high rank if the addition of the ranks of its back links is high.

A simplified version of PageRank is given in Equation 1 as follows:

$$\mathrm{PR}(u) = c\sum\nolimits_{v\in B(u)} \frac{\mathrm{PF}(v)}{N_o} \tag{1}$$

where u represents a web page, $B(u)$ is the set of pages that point to u, PR(u) and PR(v) are rank scores of page u and v, respectively, N_o indicates the number of outgoing links of page v, and c is a factor applied for normalization.

As Equation (1) does not suite well with the structure of the web later, the PageRank was customized. To include

$$\mathrm{PR}\,(u) = (1-d) + d\sum\nolimits_{v\in B(u)} \frac{\mathrm{PF}(v)}{N_o} \tag{2}$$

where d is a dampening factor or teleportation factor that is frequently set to 0.85 by Google. d can be thought of as the prospect of users following the direct links and (1–d) as the PageRank distribution from non-directly linked pages.

4. Max–Min ant system

Max–Min ant system was designed to improve the quality of performance of the ant colony system as the problem space became large (1997). Stuzles and Hoos (1998) observed that choosing the appropriate trial limit depends on the problem size and alpha parameter. Therefore, Max–Min has intervals which limit the search space and avoid stagnation. To exploit the best solutions found during iteration or during the run of the algorithm, after iteration, only one single ant adds pheromone which is the best ant. Initially, m ant is placed on m random cities. Each node i and j has a distance of d_{ij} Each ant k moves on the probability.

$$P_{ij}^{k}(t) = \frac{[\tau_{ij}(t)]^{\alpha} * [\eta_{ij}]^{\beta}}{\sum_{l\in N_i^k} [\tau_{ij}(t)]^{\alpha} * [\eta_{ij}]^{\beta}} \tag{3}$$

where m = number of ant;

τ_{ij}= pheromone trial;

η_{ij}= heuristics information which is 1/d$_{ij}$ (Hadoop);

α, β = parameters determining the relative importance of the pheromone trial and heuristic Information, respectively. Max–Min allows only the best ant to updates the pheromone trails and that the value of the pheromone is bound. The pheromone update is implemented as follows:

$$\tau_{ij} \leftarrow \left[(1-\rho) * \tau_{ij} + \Delta_{ij}^{\mathrm{best}}\right]_{\tau\min}^{\tau\max} \tag{4}$$

where

ρ is the evaporation rate;

$\tau_{\max}$ and τ_{main} are the upper and lower bounds imposed on the pheromone.

$$\Delta\tau_{ij}^{\mathrm{best}} \qquad \Delta t_{ij}^{\mathrm{best}} = \begin{cases} 1/L_{\mathrm{best}} \text{ if } (i,j) \text{ belong to the best tour} \\ 0 \text{ otherwise} \end{cases} \tag{5}$$

The basis of Max–Min ant system is to provide dynamically evolving bounds on the pheromone trail intensities such that the pheromone intensity on all paths is always within a specified limit of the path with the greatest pheromone intensity. As a result, all paths will always have a non-trivial probability of being selected and thus wider exploration of the search space is encouraged (Sargolzaei and Soleymani, 2010).

4.1. Max–Min ant system algorithm

Procedure for the Max–Min ant system

(1) Initialize the pheromone trails and the parameters

(2) While (termination condition not met) do

For k:=1 to m do

Construct a solution for k

End for;

Select ants to improve their solution by local search

Apply local search to selected ants

Update the pheromone trail, $\tau ij_{\min_{\max}}$

(3) Return best solution found

End MAX-MIN

According to Hadoop and Sargolzaei and Soleymani (2010), in order to ensure pheromone intensities, upper and lower bounds are given a range, i.e., $\tau\mathrm{ij}_{\max_{\min}}$. The upper bound $\tau_{\max}$ is given as follows:

$$\tau \frac{1}{1-\rho f\left(S^{gb}(t-1)\right)}_{\max} \tag{6}$$

where

ρ =

fs^{gb} = is the global solution value for a given problem.

And the lower bound is given as follows:

$$\tau \frac{\tau \sqrt[n]{P_{\mathrm{best}}}_{\max}}{(\mathrm{avg}-1)\sqrt[n]{P_{\mathrm{best}}}}_{\min} \tag{7}$$

where p_{best} is the probability of the current global-best path, $S_{gb(t)}$, that will be selected given that all non-global best edges have a pheromone level of $\tau_{\min(t)}$ and all global best edges have a pheromone level of $\tau_{\max(t)}$, n

is the number of decision points, and is the average number of edges at each decision point.

The Max–Min ant system was improved using a minimizer and the improved Max–Min ant system is used to prune nodes in the graph.

4.2. Algorithm for improved Max–Min ant system

Procedure for the improved MAX–MIN ant system

(1) Initialize the pheromone trails and the parameters
(2)While (termination condition not met) do
For k = 1
Use Minimizer to construct a solution
end for
Return best solution found
For k = 2 to m do
Construct a solution for k
end for;
Select ants to improve their solution by local search
Apply local search to selected ants
Update the pheromone trail,
(3) Return best solution found

End Improved MAX-MIN

5. Design of the Dynamic PageRank Algorithm for a Web Page Using the Improved Max–Min Ant System

The improve Max–Min ant colony system uses the dynamic nature of ant colony optimization (ACO) in an optimized form to prone the massive database of a search engine and produce a subgraph for web page ranking and the propose method is going to be implemented on the map-reduced framework.

Fig. 1 shows the proposed model which comprises seven components that interact to achieve the efficient retrieval of relevant information and ranking them in a stipulated time. These components are WWW, improved Max–Min ant system, subgraph, PageRank, display result, and user, which are explained as follows.

i. WWW
This is the World Wide Web, where all information is available. It is the interconnection of computers/devices for the purpose of sharing information.

ii. Massive database
A massive database is a data bank of a search engine where all websites are indexed and stored to be recalled later when needed for a search query. A search engine has a crawler that always crawls the Internet

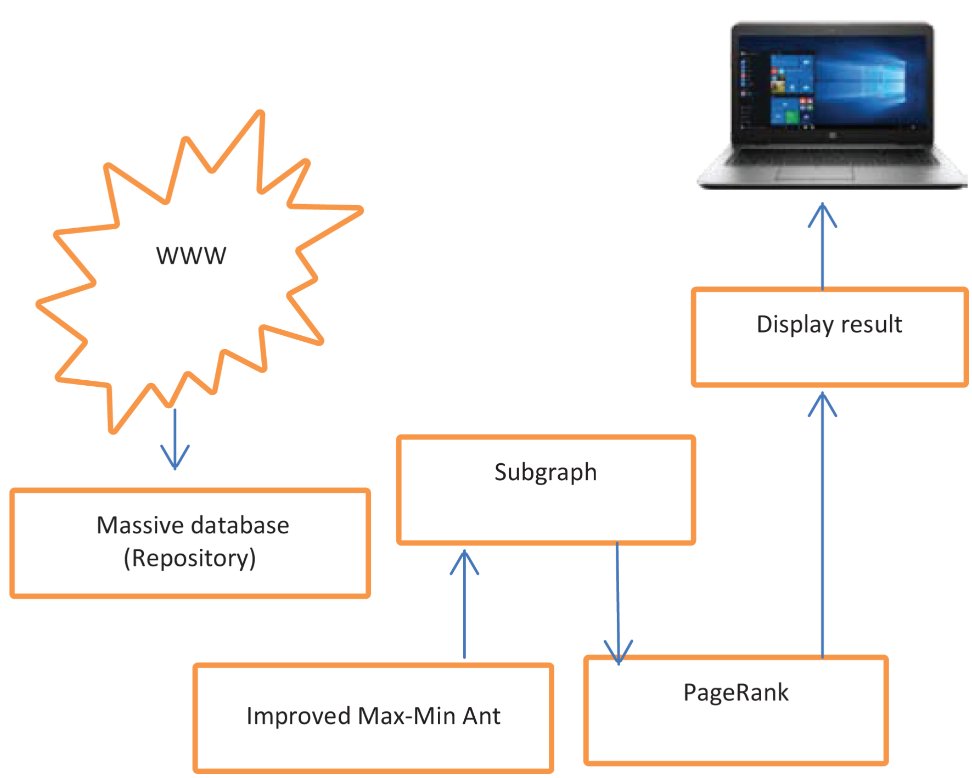

Figure 1. Proposed model of the system.

and indexes and stores information in its data storage unit to accomplish the monumental task of holding billions of pages that can be accessed in a fraction of a second. For instance, the Google search engine has constructed datacenters all over the world.

iii. Improved Max–Min
It is an algorithm, an improved version of the Max–Min ant colony that transverse the data store and create a subgraph that constitutes nodes and vertices as described below.

iv. Subgraph
A graph whose vertices and edges are a subset of another graph, formally can be defined as a graph G = (V', E'), which is a subgraph of another graph G = (V, E) if and only if $V' \subseteq V$ and $E' \subseteq E \wedge ((v1, v2) \in E' \rightarrow v1, v2 \in V')$. A graph G' is a subgraph of another graph. A subgraph is a small portion of graph that is created based on certain criteria. In this case, an ant colony algorithm was used to create the sub-graph.

v. PageRank
This is where ranking of the relevant web pages is carried out before it is being displayed to the user. From the subgraph created, the PageRank ranks relevant pages accordingly before releasing it to the user in an order of relevance.

vi. Sort/display result
This is where all the relevant web pages are then sorting accordingly and display the result to the user.

vii. User
A user is a person who initiates the query in this system. A user may not have the technical expertise required to fully understand the system. All users wants is a result of his/her query displayed within seconds. In order for user to be satisfied, this result must be displayed within seconds.

6. Benefit of the Proposed Work

The map-reduced framework is among the prominent techniques introduced by Google to made an indelible contribution to the development of cloud computing technology. The proposed dynamic PageRank algorithm-based architecture has the potential ability for pruning nodes efficiently using the improve Max–Min with lesser computational costs. The JGraph framework used simulates the actual real-world scenario to predict the outcome of the output. All this in turn improves the performance of the PageRank algorithm.

7. Conclusion and Further Research

The challenges that the traditional PageRank of Page and Brin faces are the sheer size of the Internet and the dynamic nature of the web. These challenges can be resolve by incorporating a dynamic mechanism in the PageRank computation. The injection of an optimized dynamic mechanism in the traditional algorithm will resolve these challenges and efficiently improve the performance of the PageRank. To avoid the drawback of the traditional method, this paper propose a model of computing PageRank algorithm that uses an improve Max–Min ACO to prune the graph for nodes JGraph for implementing it. The potential advantage of this model is to improve the performance of PageRank.

Computation costs make it more efficient. We intend to forge ahead to implement this propose model in our future research.

References

Anuradha G, Devi LG. Artificial Bee Colony (ABC) Approach for Ranking Web Pages International Journal of Computer Applications 2014; 99(1):0975–8887.

Anuradha G, Devi LG, Prasad Babu MS. AntRank: an Ant Colony algorithm for web page. Int J Emerg Trends Technol Comput Sci 2014; 3(2).

Arasu A, Cho J, Garcia-Molina H, Paepcke A, Raghavan S. Searching the web. ACM Trans Internet Technol 2001; 1(1):2–43.

Bahmani B, Chowdhury A, Goel A. Fast incremental and personalized PageRank. In Proceedings of the VLDB Endowment, 2010; 4(3):VLDB Endowment 2150-8097/10/12.Barroso LA, Dean J, Holzle U. Web search for a planet: the Google cluster architecture. IEEE Micro 2003; 23(2):22–28.

Bharat K, Henzinger MR. Improved algorithms for topic distillation in hyperlinked environments. In 21st International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), ACM Press, New York, NY, pp 104–11, 1998.

Brin S, Page L. The anatomy of a large-scale hypertextual web search engine. Comput Networks ISDN Syst 1998; 30:107–17.

Broder A, Ronny L, Farzin M, Jan P. Efficient Pagerank approximation via graph aggregation. In Proceedings of the 13th International World Wide Web conference. Alternate Track Papers & Posters. ACM Press, New York, –, 2004.

Buzzanca M, Carchiolo V, Longheu A, Malgeri M, Mangioni G. Black hole metric: overcoming the PageRank normalization problem. Inf Sci 2018; 238:58–72.

Chien S, Dwork C, Kumar R, Simon DR, Sivakumar D. Link evolution: analysis and algorithms. Internet Math 2004; 1(3):277–304.

Chen YY, Gan Q, Suel T. I/O-Efficient Techniques for computing PageRank. In Proceedings of the Eleventh International Conference on Information and Knowledge Management (CIKM'02), ACM Press, New York, NY, –, 2002.

Chun Liu and Yuqiang Li (2015) A parallel PageRank algorithm with power iteration acceleration. Int J Grid Distrib Comput 2015; 8(2):273–84; doi:10.14257/ijgdc.2015.8.2.24.

Corne D, Dorigo M, Glover F, Dasgupta D, Moscato P, Poli R, Price KV, editors. New ideas in optimization. McGraw Hill, London, UK, pp 63–76, 1999.

Ding C, He X, Husbands P, Zha H, Simon H. PageRank, HITS and a unified framework for link analysis. In Proceedings of the 25th ACM SIGIR Conference, ACM Press, New York, NY, pp 353–4, 2003.

Ding CH, Zha H, He X, Husbands P, Simon HD. Link analysis: hubs and authorities on the World Wide Web. Lawrence Berkeley National Laboratory, Berkeley, CA, 2001. Technical Report 47847.

Dorigo M, Gambardella LM. Ant Colony System hybridized with a new local search for the sequential ordering problem. INFORMS J Comput 2000; 12(3):237–55.

Liu C, Li Y. A Parallel PageRank algorithm with power iteration acceleration. Int J Grid Distrib Comput 2015; 8(2):273–84; doi:10.14257/ijgdc.2015.8.2.2.

Farahat A, Lofaro T, Miller JC, Rae G, Ward LA. Existence and uniqueness of ranking vectors for linear link analysis. SIAM J Sci Comput 2004; 27:1181–201.

Fujiwara Y, Nakatsuji M, Shiokawa H, Mishima T, Onizuka M. Fast and exact top-k algorithm for PageRank. In Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence, California, 2013.

Gambardella LM, Dorigo M. Solving symmetric and asymmetric TSPs by ant colonies. In T. Baeck et al. (Eds.). Proceedings of IEEE International Conference on Evolutionary Computation, IEEE Press, Piscataway, NJ, Nagoya, Japan, pp 622–7, 1996.

Gang Wu, Ying Zhang & Yimin Wei (2013) Accelerating the Arnoldi-Type Algorithm for the PageRank Problem and the ProteinRank Problem. Journal of Scientific Computing ISSN 0885-7474 Volume 57

Gleich D, Zhukov L, Berkhin P. Fast parallel PageRank: a linear system approach. Research of mining the mutative knowledge with extension data mining[J]. Eng Sci 2007; 11(8):70–8.

Gleich DF, Rossi RA. A dynamical system for PageRank with time-dependent teleportation. Internet Math 2014; 10:188–217; doi:10.1080/15427951.2013.814092

Gori M, Scarselli F. Inside PageRank. ACM Trans Internet Technol 2005; 5:92–128.

Hadoop. Hadoop – MapReduce. Available via http://www.tutorialspoint.com/hadoop/hadoop_mapreduce.htm

Haveliwala T. Efficient computation of PageRank. Technical Report. Computer Science Department, Stanford University, Stanford, CA, 1999.

Haveliwala T. Efficient encodings for document ranking vectors. Technical Report. Computer Science Department, Stanford University, Stanford, CA, 2002.

Haveliwala T. Topic-sensitive PageRank. In Proceedings of the Eleventh International Conference on World Wide Web, ACM Press, New York, NY, pp 517–26, 2002.

Ivkovic N, Golob M, Malekovic M. A pheromone trail model for MAX-MIN Ant System. In Conference: Proceedings of the10th Biennal International Conference on Artificial Evolution, EA2011, Anger, France, 2019.

Kamvar SD, Haveliwala TH, Christopher DM, Golub GH. Exploiting the block structure of the web for computing PageRank. Technical Report. Stanford University, Stanford, CA, 2003.

Kamvar S, Haveliwala T, Golub G. Adaptive methods for the computation of PageRank. Technical Report. Stanford University, Stanford, CA, 2003.

Kamvar SD, Haveliwala TH, Manning CD, Golub GH. Extrapolation methods for accelerating PageRank computations. In Proceedings of the Twelfth International Conference on World Wide Web, ACM Press, Budapest, Hungary, pp 261–70, 2003.

Kreger H. Fulfilling the web services promise. Commun ACM 2003; 46(6):29–34.

Lam H, Baudisch P. Summary thumbnails: readable overviews for small screen Web browsers. In Proceedings of the SIGCHI conference on Human factors in computing systems, 2005.

Langville AN, Mayer CD. A reordering for the PageRank problem N. Carolina State University, Raleigh, NC, Vancourver, pp 27695-8205, 2004.

Langville AN, Meyer CD. Information retrieval and web search. Handbook of Linear Algebra. 2006

Langville AN, Meyer CD. Updating Markov chains with an eye on Google's PageRank. SIAM J Matrix Anal Appl 2006; 27:968–87.

Li L, Qiuli Z, Yin K, Yiming D. Improved subgraph estimation PageRank algorithm for web PageRank. Int J Secur Appl 2015; 9(11):231–48; doi:10.14257/ijsia.2015.9.11.22.

Lili Y, Gui Z, Du W, Guoa Q. An Improved PageRank Method based on Genetic Algorithm for Web Search. Procedia Eng 2011; 15:2983–7.

Lu Y, Benyu Z, Wensi X, Zheng C, Yi L, Michael RL, et al. The power rank web link analysis algorithm. In Proceedings of the 13th International World Wide

Web Conference. Alternate Track Papers & Posters, ACM Press, New York, pp 254–5, 2004.

McSherry F. A uniform approach to accelerated PageRank computation [J]. Princeton University Press, Princeton, NJ, vol. 35, no. 8, pp 17–20, 2012.

Moler CB. Numerical computing with MATLAB. SIAM, Philadelphia, PA, 2004.

Pandurangan G, Raghavan P, Upfal E. Using PageRank to characterize web structure. In 8th Annual International Computing and Combinatorics Conference (COCOON), ACM Press, New York, pp 330–9, 2002.

Raghavan S, Garcia-Molina H. Compressing the graph structure of the web. In Proceedings of the IEEE Conference on Data Compression, IEEE Press, Los Alamitos, CA, pp 213–22, 2001.

Raghavan S, Garcia-Molina H. Towards compressing web graphs. In Proceedings of the IEEE Conference on Data Compression, IEEE Press, Los Alamitos, CA, pp 203–12, 2001.

Raghavan S, Garcia-Molina H. Representing web graphs. In Proceedings of the 19th IEEE Conference on Data Engineering, IEEE Press, Los Alamitos, CA, pp 405–16, 2003.

Sargolzaei P, Soleymani F. Pagerank problem, survey and future research directions. In International Mathematical Forum, pp 937–56, 2010.

Singh J, Sharma P. PageRank algorithm based on the dynamic approach. Int J Technol Comput 2016; 2(7).

Stewart WJ. Numerical methods for computing stationary distribution of finite irreducible Markov chains. In: Grassmann W (ed.). Advances in Computational Probability, chapter 4, Kluwer Academic Publishers, Springer, Boston, pp 81–111, 1999.

Stutzle T, Dorigo M. ACO algorithms for the quadratic assignment problem. In: Corne D, Dorigo M, Glover F (eds.). New Ideas in Optimization, McGraw Hill, London, UK, pp 33–50, 1999.

Stuzles T, Hoos T. Improvement on Ant-System: introducing Max-Min Ant System. Technical University of Darmstadt, Department of Computer Science, Intellectics Group, 1998.

Stuzle T, Hoos HH. MAX–MIN Ant System. Future Gener Comput Syst 2000; 16(8):889–914.

Wang Y, DeWitt DJ. Computing PageRank in a distributed Internet search system. In Proceedings of the 30th VLDB Conference, Toronto, Canada, 2004.

Wenlei X, David B, Alan D, Johannes G. Edge-Weighted personalized PageRank: breaking a decade-old performance barrier KDD'15, August 10-13, 2015, Sydney, Australia, ACM, 2015; doi:10.1145/2783258.2783278

Wu G, Ying Z, Wei Y. Accelerating the Arnoldi-Type Algorithm for the PageRank Problem and the ProteinRank Problem. J Sci Comput 2013; 57.

Wu Y Louiqa R. ApproxRank: estimating rank for a subgraph. In 25th IEEE International Conference on Data Engineering, IEEE, Shanghai, China, 2009; doi:10.1109/ICDE.2009.108