



OPTIMIZATION OF PAGE RANKING WITH SWARM INTELLIGENCE ALGORITHMS FOR RE-RANKING OF WEB SEARCH RESULTS

R. Deepa¹, ¹Research Scholar, NGM College Pollachi, Coimbatore, India, deepaharini2015@gmail.com

Dr. R ManickaChezian², ²Associate Professor, NGM College Pollachi, Coimbatore, India, Chezian_r@yahoo.co.in

ABSTRACT- Search engines use the Page Ranking Algorithms to deliver the search results by taking the relevance, importance and content score into consideration. Web mining techniques are brought to use by the search engines for the extraction of relevant documents from the web database documents and supply the necessary and required information to the users. If the search results are not displayed according to the user's interest then it will lead to the search engine losing its popularity. PageRank is the famous web page sorting algorithm used commonly by Google.com. The web page whose content is often updated by the owner with the most relevant data obviously gains the user casually getting attracted towards that web page. This makes the web page to get more interests than its competitors. On the other hand, this is not possible with Page Rank algorithm as the referential concept only provides the URL irrespective of the content. This paper optimizes the Page Rank Algorithm with Cuckoo Search Optimization Algorithm (OCSPPA) and compares

with OptimizedBee Colony Page Ranking Algorithm (OBPRA).

Keywords: Search Engine, Bee Colony, Cuckoo Search, Optimization.

1. INTRODUCTION

The Page Rank algorithm presented by (Page et al, 1999) is one of the factors utilized by Google for computing the relative importance of the web pages. Page Ranks render a probability distribution over web pages, such that the sum of all web pages' Page Ranks will be one. Page Rank PR(A) is given as follows.

$$PR(A) = (1 - d) + d (PR(T1)/C(T1) + \dots + PR (Tn) / C(Tn))$$

(1)The rank score of a web page is divided evenly over the pages to which it links. The bee colony optimization algorithm and cuckoo search optimization is employed along with page rank to form Optimized PageRank algorithms.

2. LITERATURE REVIEW

The Page Rank considers the back link in deciding the rank score. If the addition of the all the ranks of the back links is large then the page then it is provided a large rank (Page et al, 1999; Ridlings et al, 2002). Recently (Tzikas et al, 2006) introduced another machine learning technique The process of the search of honeybees for nectar in flowers can be taken as an optimization (Ashita 2012) process. (Tereshko 2000) developed a model of scrounging behavior of a honeybee colony based model. This model which reveals the collective intelligence of honeybee swarms consists of three essential components: food sources, employed foragers, and unemployed foragers, Tereshko explains the main components of his model as below (Tereshko 2000; Tereshko & Lee 2002). Cuckoo search is an optimization algorithm developed by (Yang & Deb 2009). It was inspired by the obligate brood parasitism of some cuckoo species by laying their eggs in the nests of other host birds (of other species).

3. BEE PAGE RANK ALGORITHM

The Optimized bee based Algorithm OBPA; the performance of waggle dance by an employee bee provides his fellow bees the information about the Direction of availability of nectar or pollen, Distance of food zone from the hive and Quality of the nectar available. The optimal solution relates to the users who are most nearest to the current user.

4. CUCKOO SEARCH OPTIMIZED PAGERANK ALGORITHM

Algorithm Optimized Cuckoo Search Page Rank Algorithm

1. Initially assume PAGE RANK of all web pages to be any value, let it be 1.
2. Calculate page ranks of all pages by following formula
$$PR(A) = 0.15 + 0.85 (PR(T1)/C(T1) + PR(T2)/C(T2) + \dots + PR(Tn)/C(Tn))$$
Where T1 through Tn are pages providing incoming links to Page A, PR(T1) is the Page Rank of T1, PR(Tn) is the Page Rank of Tn, C(Tn) is total number of outgoing links on Tn.
3. Calculate mean value of all page ranks by following formula:-
Summation of page ranks of all web pages/number of web pages
4. Then optimize page rank of each page using CS.
Optimize PR (A) = PR (A)/mean value
Where norm PR (A) is Normalized Page Rank of page A and PR (A) is page rank of page A
5. Assign PR(A)= Optimize PR (A)
6. Repeat step 2 to step 4 until page rank values of two consecutive iterations are same. The pages which have the highest page rank are more significant pages.

Sri Vasavi College, Erode Self-Finance Wing

3rd February 2017

National Conference on Computer and Communication **NCCC'17**

<http://www.srivasavi.ac.in/>

nccc2017@gmail.com

Fig1. Optimized Cuckoo Search Page Rank Algorithm

Cuckoo search algorithm improves the ranking results more than the bee colony optimization. The reason for replacing BCO is that it is good at exploration but poor at exploitation and the convergence speed is also an area of serious concern. Hence the Cuckoo search optimization is employed in optimizing the web pages.

5. PERFORMANCE EVALUATION

The performance metrics namely accuracy, precision, recall, f-measure, iterations and time are used for comparison. The comparisons are made in two categories of topics, Academic and research domains based on the user test queries which extracted the related web pages. The performance of the PRA, OBPRA, OCSRA method in terms of performance metrics are given in the following table.

<u>PerEval</u> Algorithms	Accuracy (%)	Recall (%)	Precision (%)	F-Measure (%)	Time (ns)	Memory Usage (kb)
<u>Optimised Ranking</u>						
PRA	86.4	83.2	82.6	82.3	36	2345
OBPRA	90.5	86.4	91.5	88.4	33	2196
OCSRA	94.1	90.5	94.1	93.5	32.6	2036

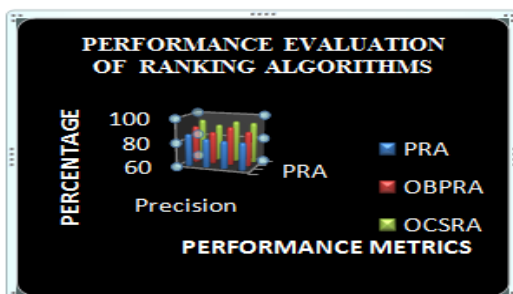


Fig 1. Performance Evaluation of Ranking Algorithms

6. CONCLUSION

This paper provides a detailed explanation about the optimized PageRank algorithm. The Bee colony optimization and Cuckoo search optimization are incorporated with the famous PageRank algorithm to form the proposed



Sri Vasavi College, Erode Self-Finance Wing

3rd February 2017

National Conference on Computer and Communication **NCCC'17**

<http://www.srivasavi.ac.in/>

nccc2017@gmail.com

approaches. The evaluation results prove that the proposed Cuckoo Search PageRank achieves higher performance than the Bee PageRank and the existing PageRank and it is better suited for the personalized web page recommendation systems. The OCSPPRA obtains the Accuracy, Recall, Precision, F-Measure as, 94.1%, 90.5%, 94.1%, and 93.5% respectively which deviates as high of an average of 4.6% from the existing Page Rank and Optimized BPRA.

REFERENCES

- [1] Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank citation ranking: bringing order to the web. Technical report, Stanford Digital Libraries SIDL-WP-1999-0120
- [2] Ridings, C., & Shishigin, M. (2002). PageRank uncovered. Technical Paper for the Search Engine Optimization Online Community.
- [3] Teodorovic, D., Lucic, P., Markovic, G., & Dell'Orco, M. (2006). Bee colony optimization: principles and applications. In 2006 8th Seminar on Neural Network Applications in Electrical Engineering (pp. 151-156). IEEE.
- [4] Bonabeau, E., Dorigo, M., & Theraulaz, G. (1999). Swarm intelligence: from natural to artificial systems (No. 1). Oxford university press.
- [5] Karaboga, D., & Akay, B. (2009). Artificial bee colony (ABC), harmony search and bees algorithms on numerical optimization. In Innovative production machines and systems virtual conference.
- [6] Ashita, S. B. (2012). Artificial Bee Colony (ABC) Algorithm for Vehicle Routing Optimization Problem. International Journal of Soft Computing and Engineering (IJSC), 2231-2307.
- [7] Tereshko, V. (2000). Reaction-diffusion model of a honeybee colony's foraging behaviour. In International Conference on Parallel Problem Solving from Nature (pp. 807-816). Springer Berlin Heidelberg.
- [8] Tereshko, V., & Lee, T. (2002). How information-mapping patterns determine foraging behaviour of a honey bee colony. Open Systems & Information Dynamics, 9(02), 181-193.
- [9] Shalya, N., Shukla, S., & Arora, D. (2012). An Effective Content Based Web Page Ranking Approach. International Journal of Engineering Science and Technology (IJEST), 4(08).
- [10] Yang, X. S., & Deb, S. (2009). Cuckoo search via Lévy flights. In Nature & Biologically Inspired Computing, 2009. NaBIC 2009. World Congress on (pp. 210-214). IEEE.