IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 14, No. 1, February 2025, pp. 73~82
ISSN: 2252-8938, DOI: 10.11591/ijai.v14.i1.pp73-82

A page rank-based analytical design of effective search engine optimization

Vinutha Mysore Srinivas[1], Padma Muthalambikasheta Halli Cheluvae Gowda[2]
[1]PET Research Centre, Department of Computer Science and Engineering, University of Mysore, Mandya, India
[2]Department of Computer Science and Engineering, PES College of Engineering Mandya, Mandya, India

Article Info

Article history:

Received Feb 12, 2024
Revised Jul 2, 2024
Accepted Jul 26, 2024

Keywords:

PageRank
Search engine
Search engine optimization
Web page ranking
Web pages
Web structure mining

ABSTRACT

Search engine optimization (SEO) is an important internet marketing strategy and process that facilitates maximizing an intended website's visibility with search engine results. It is widely employed nowadays to improve traffic volume or quality from search engines to a particular website. Even though a significant number of publications imply the essential aspects of SEO, only a few provide generalized ideas to deal with the complex structure of the web. Also, the critical issues of content quality, site popularity, keyword density, and publicity factors were not much considered in the traditional ranking algorithms during SEO processes. This has negatively influenced the retrieval rate in the existing SEO techniques, and consequently, inadequate search results were obtained through search engines. Hence, the study considers web page ranking as a theoretical basis for the research and addresses these limitations in the existing system. It further improves SEO performance by introducing a unique web-page ranking strategic design to gain higher page rank results. The results of the investigational study show that the proposed system effectively contributes towards SEO with an improved page ranking strategy and also provides higher accuracy in calculating the importance score of web pages which is comparable with popular ranking algorithms such as hyperlink-induced topic search (HITS) and PageRank.

Corresponding Author:

Vinutha Mysore Srinivas
PET Research Centre, Department of Computer Science and Engineering, University of Mysore
Mandya, India
Email: vinupraj2014@gmail.com

1. INTRODUCTION

The concept of a search engine is not new; 'Archie' was the first to be released in the early 1990s, specifically to search file transfer protocol (FTP) data. Conversely, 'Veronica' was believed to be the first text-based search engine ever created [1]. The fundamental motive of businesses in the current digital marketing era is to communicate appropriate information about their products and services to the right customers through websites with minimal effort, which has led to an increase in the number of websites on the 'world wide web' [2], [3]. Search engine theory develops from the perspective of giving appropriate website pages to the targeted consumers, and in this way, it is consistent with the fundamental ideas of internet marketing. Search engines navigate through the billions of pages available on the internet and sort the web pages based on their relevance to the user-generated query (e.g., keywords and phrases) [4]–[6]. Currently, the popular internet search engines are Google, Yahoo, and Bing. As a result, whenever a user enters a specific phrase or keyword in a search instead of the whole website URL for a company, the search engine uses that term to find the appropriate web pages [7]–[10]. Further, a list with the most pertinent page at the top is displayed. This approach helps organizations reach their potential consumers by appearing at the top of the search results. Many of the traditional search engine optimization (SEO) design strategies are intended to attain higher page rank results where the popularity of page ranking algorithms arises [11], [12]. Various research studies are being carried out on designing effective SEO strategies to gain higher page rank results from the perspective of performance improvement. However, scattered challenges arise when it comes to the practical implementation of research-based studies in SEO. The research-based studies on SEO have reported various pitfalls associated with the existing ranking algorithms [13]. Most existing page rank algorithms only employ a few relevant keywords to retrieve top-k web pages. Also, the resulting web pages may not meet the intended search query. Also, in the existing page rank algorithms, the need for cost-effective analytical processing with reliable and higher efficient page ranking is highly envisioned to ensure the adequate performance of SEO [14]–[16].

The research on web structure mining evolves with the purpose of discovering information from the web. It also searches for information pertaining to relevant scores for web pages and hyperlinks to determine the quality of the search results. It basically focuses on organizing the hyperlink structure of the web. The study by Dubey and Roy [17] talks about the significant factors of page ranking for measuring the importance and behaviour of web pages. Jayaraman *et al*. [18] also highlights that the performance of SEO could be increased if the process of website ranking is designed in an optimized flow of execution for particular search terms in search engines. It also shares the conventional ideas related to SEO on-page ranking while managing the incoming links and website characteristic features. There is no denial of the fact that the performance of traditional page rank algorithms could be improved by using web mining techniques, viz., web structure, web content, and web usage as highlighted in [19].

Alghamdi and Alhaidari [19] also talk about the core web page ranking algorithms and explores their idea for enhancing the performance of SEO. The idea of page ranking evolved from the most popular baseline of the PageRank algorithm, which Google employed. The design and operational factors of the traditional PageRank algorithm utilize the web structure mining concept to compute the page rank values. However, the core design model of this algorithm computes the rank score of the page at indexing time and evaluates the page score considering the in-links, which could mislead the search results in the post-ranking phase, as claimed by Suri *et al*. [3]. The computational approach of representing the working flow of hypertext-induced topic selection (HITS) considers a directed graph structure where vertices represent the web pages and a set of edges depicts the links [20]. Zhang *et al*. [20] claimed that the performance of this algorithm could be improved to enhance the scope of SEO operations for various user queries. Chowdhary and Kumar [21], in their study, talk about the sub-variant of the main PageRank algorithm, which is referred to as the weighted page rank algorithm (WPR). The authors claim that this baseline approach has a broader scope for improvement owing to its advantages of computing both in-links and out-links towards measuring the importance of a web page's score.

Kelotra *et al*. [22] designed an improved method of page ranking considering the baseline properties of the traditional PageRank algorithm. Here the method operates on page rank based on the duration a user spends on the web page and its link structure. The authors claim this approach could effectively offer a better retrieval rate for web search engines. Hao *et al*. [23] also enhanced the traditional PageRank algorithm for web content search. The study by Usha and Nagadeepa [24] also introduces a hybrid page rank algorithm where the algorithm utilizes web structure, web content, and web usage mining techniques to compute the in-links of the web pages. Tuteja *et al*. [25] perform modifications on the design features of traditional WPR considering the frequency of visits of in-links and out-links, which is further combined with the original mathematical formulation of the WPR algorithm strategy. Singh and Sharma [26] also introduced another form of page rank algorithm, which considers both web structure and web usage mining techniques. The algorithm design computes both in-links and out-links weights and the frequency of visits of in-links on web pages. The study offers better results with its strategy for improving search engine results. Another approach to page ranking considering content weight is presented by Joshi and Gupta [27]; here, the study considers content weight parameters corresponding to web pages for respective query terms to calculate the page ranks. In the study by Jaganathan and Desikan [28], the page rank algorithm also computes the in-link and out-link weights associated with the web pages. However, it also constructs its weight matrix, which facilitates retaining the page ranks.

An agent weighted page ranking algorithm (AWPR) is designed as an enhancement of the traditional WPR algorithm, which is subjected to perform web structure mining while computing the weight of in-links and out-links presented by Nagappan and Elango [29]. Another similar study by Gupta and Singh [30] also represents a user preference-based page ranking algorithm where it addresses the topic drift problem of the AWPR algorithm considering the advantageous factors of web-usage mining in calculating visits of in-link. Mahajan *et al*. [31] introduced a extended weighted page rank based on visits of links (EWPR VOL) algorithm to improve the performance of WPR. Alghamdi and Alhaidari [19] further improvised the performance of EWPR VOL algorithm to enhance search engine performance for appropriate retrieval of search results considering web structure, web content and web usage mining approaches. Alhaidari *et al*. [32] have also introduced a decision-making trial and evaluation laboratory (DEMATEL) model to improve the performance of websites towards fulfilling the user's requirements. A linear programming-based statistical modelling was introduced to enhance the ranked list of web search engines by Amin and Emrouznejad [33]. The experimental outcome shows that optimizing the search engine results takes much longer than usual. Another ranking technique is designed in the study of Bozkir and Sezer [34], where the approach considers computing visual similarities among webpages. However, the retrieval scores of a search engine are affected by higher false positive scores [34]. A similar ranking scheme is also introduced Ahmad *et al*., [35], where an enumerative feature, subset–based ranking, modelling was developed to improve the search engine results. However, the retrieval rate of the scheme was found to be very poor. There are SEO techniques as depicted in the studies in [36], [37], where the prime emphasis was laid on increasing a website's visibility. However, the time to mine the top-ranked website pages was not minimized. Also, the computational complexity is much higher in the approach of [37]. Ozdemiray and Altingovde [38] introduced a ranking aggregation technique. The primary motive of this study was to minimize the computational complexity of web page optimizations. However, the experimental results show that the approach takes comparatively longer to retrieve the top-ranked results by the search engines. Banaei and Honarvar [14] have encouraged using machine learning-based approaches in SEO for determining the website's rank. However, the approach was found satisfactory with the test data but the retrieval time for extracting top-k web page results was higher due to its iterative operations.

However, the challenge arises to balance the trade-off between computing effort and efficient page ranking for retrieval of targeted pages on the first page of search results. The analysis of existing approaches to SEO exhibits that there have been extensive research efforts towards improving the analytical operations in page ranking strategies. However, it is essential to outline both the strength factors and limitations in the existing designs of page rank strategies so that readers will have a clearer idea about the scope of improvement for the future line of research. The analysis of the conventional strategies for SEO-based page ranking algorithms generates two general research questions (RQs), which are as follows:

- RQ1: What contributes to search engine rankings?
- RQ2: What can web content creators and admins do to make their content and sites easier to find by audiences using search engines?

The proposed study also reviewed some essential page ranking strategies in SEO operations. However, despite having, popularity, the existing page ranking designs in SEO suffers from various shortcomings, which could be noted as follows:

- Most existing page ranking strategies do not cover multiple parameters during SEO operations. This leads to critical issues regarding content quality, site popularity, keyword density, and publicity factors [17]–[21].
- Even though page ranking analytical design modelings are highly encouraged for effective search engine results towards retrieving top-k targeted pages, the retrieval rate in most of the existing SEO approaches could be better.
- The existing page rank algorithms for SEO need to meet the requirements for intended search as it evaluates the page score by considering only links. Also, most of the existing SEO designs are affected by the problem of topic drift [22], [24].
- In many cases, challenges arise when dealing with links connected with two or more similar sites. In contrast, some links could be created unrealistically to encourage the appearance of spam pages by the search engine (SE) to be in the top-ranked search results. This also misleads the results to the users [29], [30].
- It is also observed that the traditional SEO approaches execution workflow models are computationally complex and affects the retrieval time performance, even though, in many cases, inadequate SE results are also obtained [35].
- The existing SEO approaches need to ensure a proper balance between the retrieval rate of the top-ranked search results with the computational complexity aspects which influences the time of retrieval [14], [36]–[38].

All the research mentioned above problems are identified to have concrete solutions; hence, the proposed scheme deploys a novel page rank strategy to address these open-end research problems in SEO.

The study addresses the practical implementation constraints associated with SEO from the cost point of view and aims to improve the performance of SEO with an optimized page rank strategy. The proposed study, in this regard, introduces a framework that analytically develops an optimized page rank algorithm to obtain the significant ranking of web pages so that intended search results could be published with lesser computational effort. The proposed strategy of page ranking is modeled to facilitate gaining higher page rank results with an optimized flow of execution and also enhances the retrieval performance while covering multiple parameters. The uniqueness of the proposed SEO approach lies in the fact that it incorporates novel analytical strategic execution to contribute towards best search results through improving the performance of the page rank algorithm and also ensuring adequate convergence value for retrieval of top-ranked targeted pages within considerable iterations.

The entire manuscript is structured as follows. In section 2, the extensive idea corresponding to the research methodology is discussed, along with elaborated discussion on the system design, and the algorithm description. Section 3 highlights the acquired results and justifies the proposed study's effectiveness. Finally, section 4 also provides conclusive remarks about the overall work and highlights its novel contributions.

2. METHOD

The proposed study introduces a simplified and unique design approach to formulate a page rank strategy for effective search engine operations. It explores the ideas behind the baseline PageRank models and further attempts to optimize its flow of execution to normalize ranking a set of websites which also contributes towards effective search engine results. Here the prime motive of this algorithm is to make the targeted search results appear on the top of the front page with the pages having the highest relevancy scores. The conventional SEO designs are based on the PageRank algorithm variants, which aim to retain high page rank scores at the top of the front page considering a specific user's provided search (query).

The study applies a power method that updates the weighted reference counts generated by the hyperlinks between pages. Further, the system applies a connectivity vector and evaluates the hyperlinks between the pages to measure the in-degree and out-degree for the respective pages. The computation further also explores the probabilistic factors of user behaviour and constructs a transitional probabilistic matrix. Further, the strategy applies Perron–Frobenius theorem and assesses the scaling factor to obtain the page rank score for respective pages. The study also applies a normalization technique to make the ranking of the web pages more reliable within considerable amount of iteration for execution schema. The page scores are further normalized considering the total count of out-going links of the source nodes. The strategy also shares the idea of normalizing page rank of each page considering a mean value operation and further assesses the iterations to retain highest page rank scores for more significant pages during SEO operations.

The proposed approach's prime motive is to design and develop a novel analytical operation for page ranking in SEO. An explicit system design and modeling approach is constructed to realize the system operations for effective search engine results from both retrieval performance and cost of computation point of view. The system design modeling considers graph theory to explore the link structure of world wide web (www). Here the system design and modeling are also inspired from the conventional PageRank algorithm towards ranking the search engine results. The study also explores the idea through which PageRank algorithm ranks a collection of websites.

2.1. Modelling of markov process

The study considers the theory of the Markov chain or Markov process [39] to design the analytical operations of page ranking in SEO. The underlying idea of Markov chain refers to a stochastic process in which states change for the transition probabilities. Here the transition probabilities are determined by the steps of the previous time step. This theory is crucial to model large-scale systems with random behaviour where the area of search engine operation also arises. In web surfing, a user navigates from one page to another by randomly choosing the outgoing links. This can lead to the dead end of web pages with no outgoing links. Alternatively, it can also happen that the user cycles around interconnected pages. So, it is evident that a user tends to choose a random page from the web in a certain fraction of time. This scenario is often called a random walk and theoretically can be described with Markov process.

2.2. Analytical operation of pagerank strategy

A probabilistic strategic evaluation is related to estimating page rank score. Here, the proposed system considers a limited probabilistic score (P_l), which denotes the likelihood that a random web surfer will visit any website. It also goes by the name PageRank. Let W is a set of web pages which can be represented as $W = \{W_i\}_{i=1}^{n}$. Here n represents the number of web pages. The formulated page ranking strategy basically operates on the count of incoming (I_{link}) and outgoing links (O_{link}) to a page. The strategy also evaluates the quality of links to a page for effective SEO. This idea helps determining and generalizing the degree of importance (DOI) of a particular website. For example, a graph-based representation can be modeled to depict three web pages such as $W = \{W_i\}_{i=1}^{3}$ in the form of three vertices. The webpages can be reached through the formulation of hyperlinks (h_l) which begins at any of the root page. The proposed ranking strategy initially constructs the connectivity vector and further enables an explicit functional module $\mathfrak{f}_1(x)$ to compute the row and column summation from the $V_{n\times n}$. The system also computes a perceived importance factor (I_F) for a particular website or web-page through the proposed page rank strategy for effective SEO. The proposed page rank strategy is implemented over SEO for effective search engine results considering graph-based modeling. The connectivity vector V_{ij} represents connection or hyperlink between page i to page j. Also, the strategy further estimates a probability factor of P_r , which implies the possibility of an Internet user to randomly select and follows a link of a current page. Another probability measure of δ implies the possibility of choosing a specific random page which can be computed as (1):

$$\delta = {(1 - \mathrm{P_r})}/{\mathrm{n}} \quad (1)$$

Here the $V_{n\times n}$ represents a connectivity matrix which corresponds to a portion of web structure. The proposed formulation of page rank strategy also estimates the quantities of row_j and col_j which indicates the in-degree and out-degree measure of jth page. The strategic solution further also constructs a matrix T which is also of the dimension $(n \times n)$. The elements of T_{ij} can be represented with the (2).

$$T_{ij} = \begin{cases} \frac{P_r V_{ij}}{col_j} + \delta & : col_j \neq 0 \\ \frac{1}{n} & : col_j = 0 \end{cases} \tag{2}$$

Here the computation of T_{ij} takes place by scaling the connectivity matrix V_{ij} with respect to its column sums. Here the jth column in T_{ij} indicates the possibility of an user jumping from one page to another pages in the web. If it is found that the jth page is dead end then it has not out-going links to be associated. The strategy applies a uniform probability factor of $1/n$ in all the elements of the column vector. It can be seen that most of the elements in the matrix T_{ij} belongs to δ that indicates the possibility of jumping from one page to another without following a link. Here the transition probability matrix T_{ij} is computed considering the theory of Markov chain. The characteristics of this T_{ij} is that their element lies between $\{0,1\}$ and its columns sum is computed as 1. The study further also applies another explicit functional modeling $f_2(x)$ to compute T_{ij}. The proposed study further also employs a methodology of Perron–Frobenius theorem [40] to the matrix which is retained. The study here applies another functional module f_3(x) to compute the non-zero solution. The study also explores the best way to optimize the computation of page rank strategy by exploring advantage of particular structure of Markov matrix. Here the approach attempts to preserve the sparsity factor associated with V_{ij}. The formulation of the computation of transition matrix can be formed as (3):

$$T = \sum(P * V * Diag, \varepsilon) \tag{3}$$

Here $Diag$ represents a diagonal matrix(d_{jj}) considering the reciprocals of the out-degrees. Also ε implies a rank one matrix which is accounted for the random choices of web pages that do not follow the links. Here the page rank strategy can be optimized with the (4).

$$(I - P * V * Diag)x = \kappa \tag{4}$$

Here κ is the n vector of all ones and correspond to ε. The progressive computation of page ranking can be further updated with respect to the following normalized expression:

$$r = \sum \frac{(1-P_r)}{n}, P_r \times \left(M' \times \left(\frac{r.}{d}\right) + \frac{s}{n}\right) \tag{5}$$

Here r denotes a vector consisting of page rank scores whereas P_r implies a scalar dumping factor and its value is considered to be 0.85. This probability factor indicates the possibility of a user to click on a link on a current page rather continuing to another random page. Here M' represents an adjacency matrix of the web graph structure. Also, the vector d indicates the out-degree measure of a node in the graph structure. The value of d is considered to be 1, if there exist nodes with no outgoing links. n represents the scalar number of nodes in the graph. Here in expression (5), s represents sum of the page rank scores for the pages having no links.

2.3. Algorithm design for page ranking in search engine optimization

The following analytical algorithm exhibits the workflow model of the proposed algorithm design strategy of page ranking for SEO. Here study formulates simplified work flow modeling of the design of page ranking strategy for effective SEO. The study also incorporates a set of explicit functionalies to model the design strategy of SEO where a set of baseline strategies are also reffered for optimized execution. The steps associated with the proposed ranking strategy are illustrated in Algorithm 1.

Algorithm 1: For page ranking for effective SEO
Input: *source(s), target (t)*
Output: rank r
Start
1. Init s,t
2. $\{W_i\}_{i=1}^{n}$ for s,t
3. For i=1:s

4. For j = 1:t
5. $V_{n\times n}$, I_{link}, O_{link}, digraph
6. Execute $f_1(x)$
7. $row_i, col_j P_r$,$1 - P_r$, δ
8. $f_2(x)$ to compute T_{ij} (*Markov Chain*)
9. *Perron–Frobenius theorem* [40]
10. Optimize transition matrix T
11. $(I - P * V * Diag)x = \kappa$ (5)
12. Compute Page Rank r for SEO using (5)
13. Normalization of each r
14. Retain Significant Page Rank Score
15. End
End

The above analytical operations involved in the proposed page ranking strategy are applied over SEO for effective SE results for target and relevant top page retrieval. The study also applies formulation of directed graph structure to visualize the web structure model to illustrate how each node representing web page confers its specific rank score to other nodes or web pages. Unlike existing page rank strategies (HITS and PageRank), the proposed idea of the simplified and light-weight analytical framework of SEO considering the optimized page rank strategy not only enhance the retrieval performance, also the reduced iterations for optimal processing of the work flow model ensures effective retrieval time.

The novelty of the proposed page rank algorithm of SEO is as follows:
- Unlike existing page rank strategies, the proposed page rank algorithm contributes towards enhancing the retrieval efficiency of SE for target top-k pages.
- The design idea is simplified for light-weight analytical operations which also ensures cost effective computation and shorter retrieval time for SEO
- Unlike existing system, the proposed algorithm handles the problem of topic drift and set high rank values to more popular pages for effective search engine results on the top of the first page.

A closer look into the entire algorithm implementation shows that the proposed scheme offers a novel and sophisticated SEO operations with balanced performance between retrieval efficiency and retrieval time. The next section further illustrates the experimental outcome obtained from a strategic implementation of the formulated optimized page rank algorithm of SEO.

3. RESULT ANALYSIS

This section illustrates the numerical outcome obtained after simulating the proposed algorithm over six different websites W. The prime reason behind adopting analytical strategy for numerical simulation is – it provides better representation of the outcome considering different metrics through which the effectiveness of the proposed strategy could be validated to a greater extent. The result analysis also covers the simulation assessment strategy along with experimental outcome and analysis to conclude the effectiveness of the study model.

3.1. Simulation assesment strategy

The study considers MATLAB to construct the framework modeling for proposed optimized page ranking of SEO. It considers a regular 64-bit Windows machine with i5 processing capability. The algorithm is strategically modeled and scripted considering analytical schema to realise the objective of the proposed research study. The numerical analysis is considered to compute the values and to visualize the outcome as obtained from the proposed algorithm. For the purpose of performance assessment of the proposed strategic schema of page ranking the study not only relies on evaluating the page rank score for web pages but it also considers computational time as a parameter of complexity in the form of number of iterations of the algorithm to judge the how it converges towards the targeted top-retrieval of pages in SE results in considerable amount of retrieval time. The experimental outcome is further assessed for a comparative study under different conditions.

3.2. Experimental assesment and analysis

The strategy for result analysis considers implementing this proposed page rank idea to enhance the search engine performance for both retrieval efficiency with respect to ranking and reducing the computational time to minimize the time of retrieval. The Figure 1 shows the page rank measure obtained through the proposed page rank algorithm of SEO towards retrieving the top-8 pages from the SE. The degree of information of node computes the average InDegree and OutDegree measure are 20 and 14 respectively for the retrieved pages. The proposed study also further normalizes the page rank score of the individual pages to enhance the SE results considering the (6).

$$\overline{r_i} = \frac{r(i)}{\mu_{\forall r}} \quad (6)$$

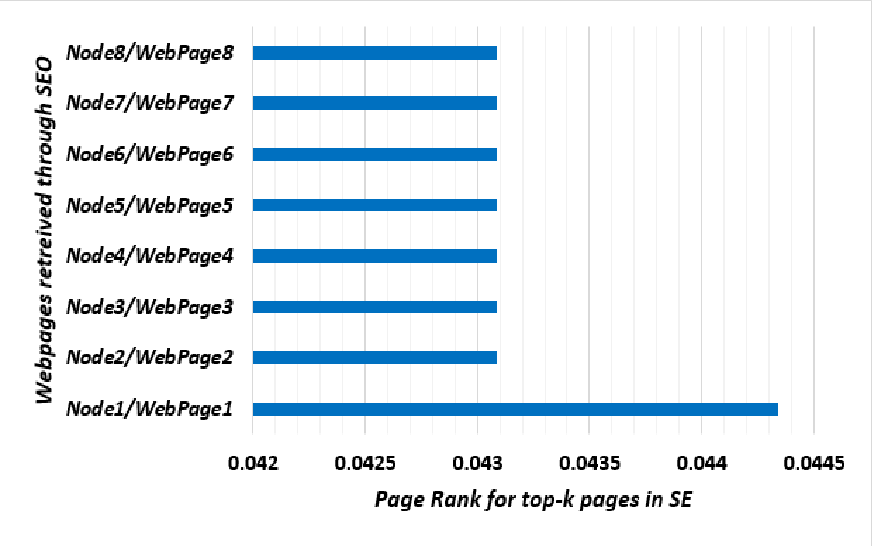

Figure 1. PageRank measure for top-k pages retreived in SE

The normalization of page rank score significantly improves the performance of SE. The study further analyses the graph structure of connectivity for degree of information among the nodes. The study also extended the analysis of the proposed optimized page rank algorithm of SEO for implementing it on another website having set of webpages linked. In this regard it refers to a data which is obtained through automatic web crawler for a particular website. The experiment here constructs an adjacency matrix of V_{ij} where the connections between 100 nodes are explored and their information are extracted. The nodes in the clique are highly connected with each other and it also depicts the fact that if a random web user uses SE, then that user has got a probability factor of approximately 4.5% chance to arrive on the retrieved pages. The study also evaluates the measure of computational time complexity for iterations which also reflects the algorithm's influence on minimizing the retrieval time.

The Figure 2 shows the comparative analysis among the conventional page rank algorithms and the proposed PageRank algorithm for the measure of computational time complexity measure in average processing time (sec). The Figure 2 shows that the proposed optimized PageRank strategy converges towards the target retrieved pages with minimal number of iterations which is comparatively more in the case of conventional PageRank algorithm and HITS. One of the primary reasons of optimizing the iteration process is that the proposed PageRank strategy simplifies the execution process with explicit functions where the core computational efforts are reduced to a greater extent. It also applies a normalization function for page ranking and retrieve the most significant pages in SE. This is how the proposed strategy also reduces the retrieval time for SE results. The study also performs a comparative analysis with significant related studies as highlighted in the literature section of the study is shown in Table 1. Table 1 highlights the comparison with the state-of-the-art methods presented in section 1. From the outcome shown in Table 1, the following inference of novelty is drawn:

- The closer analysis of the performance of the proposed PageRank strategy in SEO shows that unlike conventional approaches it accomplishes a well-balanced performance between Retrieval Accuracy and Retrieval Time. The performance of Retrieval Time is influenced by optimizing the iterations of the proposed PageRank algorithm design.
- The study also shows that the simplified design modeling of PageRank strategy is cost-effective and has reduced dependencies of iterative computation towards reaching to the convergence factor of the high-ranked target page on SE.

– Unlike conventional machine learning based designs for SEO, the proposed PageRank strategy also optimizes the retrieval time with simplified and progressive analytical operations which makes it more significant and lesser complex.
– The light-weight analytical operations not only ensure better retrieval score of PageRank but also reduces the complexity execution modeling as comparison to the existing PageRank strategy.

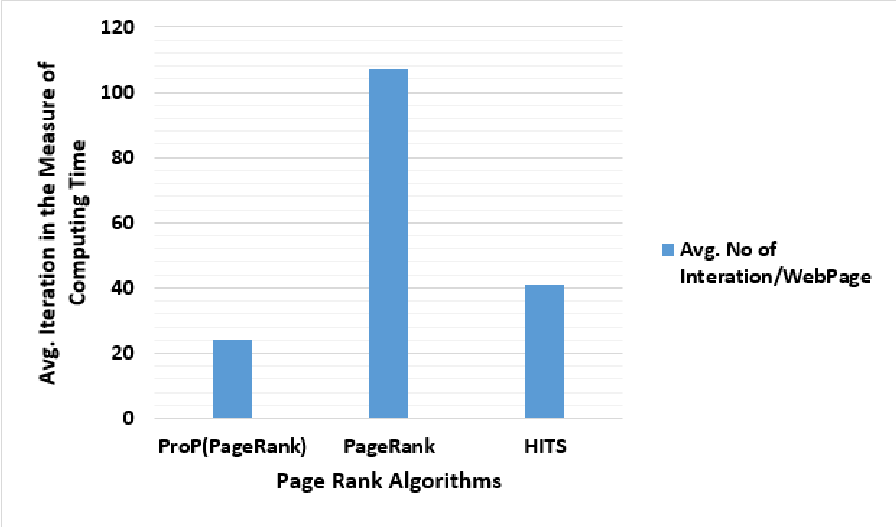

Figure 2. Comparative study of computing time measure for iterations

Table 1. Comparison with state-of-art

Method	Computational complexity	Retrieval accuracy	Retrieval time
ProP (PageRank)	Low	High	Low
Alghamdi and Alhaidari [19]	High	High	High
Suri *et al*. [20]	High	Low	High
Jayaraman *et al*. [18]	Moderate	Low	Moderate
Banaei and Honarvar [14]	High	Moderate	High
Alhaidari *et al*. [32]	High	Moderate	High

4. CONCLUSION

The study in this research work introduces an optimized PageRank algorithm for effective search engine operations. It designs the SEO based on a normalized PageRank analytical strategy where the flow of execution is optimized to balance the performance of retrieval efficiency for top-k pages and the retrieval time. This research approach offers an extensive study of the conventional related work on page ranking and further derives the problem from its core findings. The proposed research methodology uses the advantageous factors of Markov process modeling and the Perron–Frobenius theorem to offer better search results over SE. Unlike the existing system, the study also offers a well-balanced performance between retrieval efficiency and retrieval time by normalizing the page rank scores, significantly reducing the convergence time for retaining the top-k pages through SEO. In the end, the study also compares the performance of the proposed PageRank strategy with the state-of-the-art page rank designs. It justifies the proposed PageRank strategy's effectiveness in enhancing SEO performance. Future research will explore the formulated PageRank algorithm with more varying execution parameters and analyze its performance.

REFERENCES

[1] L. Page and S. Brin, "The anatomy of a large-scale hypertextual Web search engine," *Computer Networks*, vol. 30, no. 1–7, pp. 107–117, 1998, doi: 10.1016/s0169-7552(98)00110-x.
[2] R. S. Bhandari and A. Bansal, "Impact of search engine optimization as a marketing tool," *Jindal Journal of Business Research*, vol. 7, no. 1, pp. 23–36, 2018, doi: 10.1177/2278682117754016.
[3] S. Suri, A. Gupta, and K. Sharma, "Comparative analysis of ranking algorithms used on web," *Annals of Emerging Technologies in Computing*, vol. 4, no. 2, pp. 14–25, 2020, doi: 10.33166/AETiC.2020.02.002.

[4] N. Papagiannis, *Effective SEO and content marketing: the ultimate guide for maximizing free web traffic*. Hoboken, USA: John Wiley & Sons, 2020.
[5] S. Gupta, N. Agrawal, and S. Gupta, "A review on search engine optimization: basics," *International Journal of Hybrid Information Technology*, vol. 9, no. 5, pp. 381–390, 2016, doi: 10.14257/ijhit.2016.9.5.32.
[6] M. R. Baye, B. D. L. Santos, and M. R. Wildenbeest, "Search engine optimization: what drives organic traffic to retail sites?," *Journal of Economics and Management Strategy*, vol. 25, no. 1, pp. 6–31, 2016, doi: 10.1111/jems.12141.
[7] A. Veglis and D. Giomelakis, "Search engine optimization," *Future Internet*, vol. 12, no. 1, 2020, doi: 10.3390/fi12010006.
[8] T. Kelsey, *Introduction to search engine optimization-a guide for absolute beginners*. Berkeley, California: Apress, doi: 10.1007/978-1-4842-2851-7.
[9] P. Rani and E. S. Singh, "An offline SEO (Search engine optimization) based algorithm to calculate web page rank according to different parameters," *International Journal of Computers & Technology*, vol. 9, no. 1, pp. 926–931, 2013, doi: 10.24297/ijct.v9i1.4161.
[10] A. Jain, "The role and importance of search engine and search engine optimization," *International Journal of Emerging Trends & Technology in Computer Science (ITETTCS)*, vol. 2, no. 3, pp. 1–102, 2013.
[11] M. N. A. Khan and A. Mahmood, "A distinctive approach to obtain higher page rank through search engine optimization," *Sadhana - Academy Proceedings in Engineering Sciences*, vol. 43, no. 3, 2018, doi: 10.1007/s12046-018-0812-3.
[12] E. D. Ochoa, *An analysis of the application of selected search engine optimization (SEO) techniques and their effectiveness on Google's search ranking algorithm*, Northridge: California State University, 2012.
[13] N. Craswell, F. Crimmins, D. Hawking, and A. Moffat, "Performance and cost tradeoffs in web search," in *Proceedings of the 15th Australasian database conference*, 2004, pp. 161–169.
[14] H. Banaei and A. R. Honarvar, "Web page rank estimation in search engine based on SEO parameters using machine learning techniques," *International Journal of Computer Science and Network Security*, vol. 17, no. 5, pp. 95–100, 2017.
[15] S. Zhang and N. Cabage, "Search engine optimization: Comparison of link building and social sharing," *Journal of Computer Information Systems*, vol. 57, no. 2, pp. 148–159, 2017, doi: 10.1080/08874417.2016.1183447.
[16] J. Zilincan, "Search engine optimization," in *CBU International Conference Proceedings*, vol. 3, 2015, pp. 506–510, doi: 10.12955/cbup.v3.645.
[17] H. Dubey and P. B. N. Roy, "An improved page rank algorithm based on optimized normalization technique," *International journal of Computer Science and Information technology (IJCSIT)*, vol. 2, no. 5, pp. 2183–2188, 2011.
[18] S. Jayaraman, M. Ramachandran, R. Patan, M. Daneshmand, and A. H. Gandomi, "Fuzzy deep neural learning based on goodman and Kruskal's gamma for search engine optimization," *IEEE Transactions on Big Data*, vol. 8, no. 1, pp. 268–277, 2022, doi: 10.1109/TBDATA.2020.2963982.
[19] H. Alghamdi and F. Alhaidari, "Extended user preference based weighted page ranking algorithm," *2021 IEEE 4th National Computing Colleges Conference, NCCC 2021*, 2021, doi: 10.1109/NCCC49330.2021.9428844.
[20] X. Zhang, H. Yu, C. Zhang, and X. Liu, "An improved weighted HITS algorithm based on similarity and popularity," *2nd International Multi-Symposiums on Computer and Computational Sciences, IMSCCS'07*, pp. 477–480, 2007, doi: 10.1109/IMSCCS.2007.4392647.
[21] A. Chowdhary and A. Kumar, "Study of web page ranking algorithms: a review," *Acta Informatica Malaysia*, vol. 3, no. 2, pp. 1–4, 2019, doi: 10.26480/aim.02.2019.01.04.
[22] A. Kelotra, "Upgradation of PageRank algorithm based upon time spent on web page and its link structure," *International Journal of Computer Applications*, vol. 109, no. 11, pp. 7–9, 2015, doi: 10.5120/19230-0952.
[23] Z. Hao, P. Qiumei, Z. Hong, and S. Zhihao, "An improved pagerank algorithm based on web content," *14th International Symposium on Distributed Computing and Applications for Business, Engineering and Science, DCABES 2015*, pp. 284–287, 2016, doi: 10.1109/DCABES.2015.78.
[24] M. Usha and N. Nagadeepa, "A hybrid web page ranking algorithm to achieve effective organic search result," *International Journal of Scientific Research in Science and Technology*, vol. 3, no. 6, pp. 392–402, 2017.
[25] S. Tuteja, "Enhancement in weighted PageRank algorithm using VOL," *IOSR Journal of Computer Engineering*, vol. 14, no. 5, pp. 135–141, 2013, doi: 10.9790/0661-145135141.
[26] R. Singh and D. K. Sharma, "Enhanced-ratioRank: enhancing impact of inlinks and outlinks," *2013 IEEE Conference on Information and Communication Technologies, ICT 2013*, pp. 287–291, 2013, doi: 10.1109/CICT.2013.6558107.
[27] R. Joshi and V. K. Gupta, "Improving PageRank calculation by using content weight," *International Journal of Science and Research (IJSR)*, vol. 3, no. 6, pp. 2319–7064, 2014.
[28] B. Jaganathan and K. Desikan, "Weighted page rank algorithm based on in-out weight of webpages," *Indian Journal of Science and Technology*, vol. 8, no. 34, 2015, doi: 10.17485/ijst/2015/v8i34/86120.
[29] V. K. Nagappan and P. Elango, "Agent based weighted page ranking algorithm for web content information retrieval," *Proceedings of the International Conference on Computing and Communications Technologies, ICCCT 2015*, pp. 31–36, 2015, doi: 10.1109/ICCCT2.2015.7292715.
[30] D. Gupta and D. Singh, "User preference based page ranking algorithm," *IEEE International Conference on Computing, Communication and Automation, ICCCA 2016*, pp. 166–171, 2017, doi: 10.1109/CCAA.2016.7813711.
[31] I. Mahajan, S. Gupta, H. Kaur, and D. Kumar, "Extended weighted page rank based on VOL by finding user activities time and page reading time, storing them directly on search engine database server," *International Journal of Engineering Works*, vol. 4, no. 2, pp. 41–48, 2017.
[32] F. Alhaidari, S. Alwarthan and A. Alamoudi, "User preference based weighted page ranking algorithm," *2020 3rd International Conference on Computer Applications & Information Security (ICCAIS)*, Riyadh, Saudi Arabia, 2020, pp. 1-6, doi: 10.1109/ICCAIS48893.2020.9096823.
[33] G. R. Amin and A. Emrouznejad, "Optimizing search engines results using linear programming," *Expert Systems with Applications*, vol. 38, no. 9, pp. 11534–11537, 2011, doi: 10.1016/j.eswa.2011.03.030.
[34] A. S. Bozkir and E. A. Sezer, "Layout-based computation of web page similarity ranks," *International Journal of Human Computer Studies*, vol. 110, pp. 95–114, 2018, doi: 10.1016/j.ijhcs.2017.10.008.
[35] M. W. Ahmad, M. N. Doja, and T. Ahmad, "Enumerative feature subset based ranking system for learning to rank in presence of implicit user feedback," *Journal of King Saud University - Computer and Information Sciences*, vol. 32, no. 8, pp. 965–976, 2020, doi: 10.1016/j.jksuci.2017.12.006.
[36] J. B. Killoran, "How to use search engine optimization techniques to increase website visibility," *IEEE Transactions on Professional Communication*, vol. 56, no. 1, pp. 50–66, 2013, doi: 10.1109/TPC.2012.2237255.
[37] X. Gui, J. Liu, Q. Lv, C. Dong, and Z. Lei, "Probabilistic top-k query: Model and application on web traffic analysis," *China Communications*, vol. 13, no. 6, pp. 123–137, 2016, doi: 10.1109/CC.2016.7513208.

[38] A. M. Ozdemiray and I. S. Altingovde, "Explicit search result diversification using score and rank aggregation methods," *Journal of the Association for Information Science and Technology*, vol. 66, no. 6, pp. 1212–1228, 2015, doi: 10.1002/asi.23259.
[39] H. Ishii and R. Tempo, "Markov chains and ranking problems in web search," *Encyclopedia of Systems and Control*, pp. 691–697, 2015, doi: 10.1007/978-1-4471-5058-9_135.
[40] S. U. Pillai, T. Suel, and Seunghun Cha, "The Perron-Frobenius theorem: some of its applications," *IEEE Signal Processing Magazine*, vol. 22, no. 2, pp. 62–75, 2005, doi: 10.1109/msp.2005.1406483.

BIOGRAPHIES OF AUTHORS

Vinutha Mysore Srinivas received B.E. degree in Electrical & Electronics Engineering from VTU, M.E. in Information Technology from Bangalore university. She is currently pursuing the Ph.D. in the Department of Computer science & Engineering, PET Research foundation, PESCE, Mandya Affiliated to University of Mysore. She is currently Assistant Professor in Department of Computer Science & Engineering at Dr Ambedkar Institute of Technology, Bengaluru, Karnataka, India. She has published several papers in various reputed international journals and conferences. Her current research interest includes data science, natural language processing, data analytics, and graph theory. She can be contacted at email: vinupraj2014@gmail.com.

Dr. Padma Muthalambikasheta Halli Cheluvae Gowda received the B.E. degree in computer science and engineering from PESCE, Mandya, M.Sc. Tech from University of Mysore and Ph.D. from Visvesvaraya Technological University (VTU), Belgaum. She is currently professor in Department of Computer science & Engineering and Dean Research, PESCE, Mandya. She has served on several panels of academic bodies for universities and autonomous colleges as a BOS and BOE member. She has published several papers in various reputed international journals and conferences. Her research interests include in pattern recognition, image processing (document image analysis and recognition), natural language processing, data mining, data analytics, and sentiment analysis. She is serving as an editor and a reviewer for various prestigious international journals. She has delivered many keynote addresses and invited talks throughout India on a variety of topics related to computer science and engineering. She can be contacted at email: padmamc@pesce.ac.in.