Journal of Applied Data Sciences
Vol. 4, No. 4, December 2023, pp. 382-391

ISSN 2723-6471
382

Implementation of PageRank Algorithm for Visualization and Weighting of Keyword Networks in Scientific Papers

Adyanata Lubis [1,*], Elyandri Prasiwiningrum [2]

[1,2] *Computer Science, Universitas Rokania, Riau, Indonesia*

(Received: October 11, 2023; Revised: November 17, 2023; Accepted: November 25, 2023; Available online: December 5, 2023)

Abstract

Papers are written works that contain thoughts about a particular problem or topic that are written systematically accompanied by logical analysis. Scientific papers are often found on the internet or in libraries for various titles of scientific papers, citations or references can be found in every scientific paper and can be obtained easily, but to display all citations in scientific papers in the form of visualization cannot be done easily. Visualizing the citation network of scientific papers in the form of a graph, with nodes representing research papers and edges representing the relationship between researchers' scientific papers and other scientific papers based on scientific paper citations. This research uses the pagerank algorithm to create a keyword network that has a high relationship and potential relevance in a data library. This research is the first research in using the pagerank algorithm and testing its accuracy by comparing with KNN and linear clustering. The presentation displays the citation of scientific papers based on the size of the node by showing the number of citations of the scientific paper. It can be concluded that all processes in the system have run according to design, and functionally the visualization system and weighting of the scientific paper citation network are in accordance with the design. The results obtained from 51 articles, this algorithm produces a visual user interest of 81.60%, compared to the accuracy of the data suitability produced by the linear clustering and KNN algorithms in the form of 71.22% and 61.34%, helping to facilitate the search for scientific papers in large quantities.

Keywords: PageRank Algorithm, Visualization, Keyword Networks, Scientific Papers

1. Introduction

Having relevant keywords or references in an article is very important because it gives validity and trust to the reader or reviewer of the article. Relevant keywords or references show that the author has done research and taken from reliable sources before writing the article. In addition, relevant keywords or references can also be used as a basis for verifying the truth or validity of the information received by the reader [1][2]. This is especially important in the field of science or technology that is developing very quickly.

Relevant references can also be used to show that the author has done enough research and knows the topic well. This can give the reader confidence that the article is a reliable source of information. In general, relevant keywords or references can help increase the credibility and validity of the article, and make it easier to check and verify information by readers or reviewers [3][4].

Research articles often have irrelevant keywords or references for several reasons. First, sometimes article authors do not check the validity of the sources they use before listing them in the reference list. This can happen due to a lack of knowledge on how to evaluate valid sources or due to time constraints to perform sufficient checks [5][6]. Secondly, article writers may list irrelevant references in an attempt to increase the number of keywords their article receives. This can be done by including references that are not entirely relevant to the topic discussed in the article, or by including references published by the same author or institution as the article author.

Third, article authors may list irrelevant references due to a lack of skill in conducting research. This may occur due to lack of training or experience in conducting valid and relevant research, or due to lack of time to conduct sufficient research. Fourth, article authors may include irrelevant references due to a lack of standards in the peer review process.

***Corresponding author: Adyanata Lubis (adyanata@gmail.com)**
DOI: https://doi.org/10.47738/jads.v4i4.138
This is an open access article under the CC-BY license (https://creativecommons.org/licenses/by/4.0/).
© Authors retain all copyrights

ISSN 2723-6471

An inadequate peer review process can result in articles that do not meet the quality criteria accepted by journals or other publications [7][8].

Fifth, authors may include irrelevant references due to a lack of editor requirements for checking relevant references. Without clear requirements for reference checking, editors may not be aware of irrelevant references in articles accepted for publication.

The purpose of this research is to use the PageRank algorithm with linear clustering and KNN (K-Nearest Neighbors) to determine the relationship between keywords in an article. PageRank is an algorithm used to determine the relevance of a web page in search engine results. Linear clustering is a method for grouping data based on the similarity of its attributes [9][10]. KNN is a classification method used to determine the class of an object based on similarities with other objects that have been previously classified. By combining the three algorithms, it is expected to find out the relationship between keywords in an article and determine the relevance of each keyword to the article. This will help in the process of searching and classifying information in the digital world.

The novelty of this research lies in the combination of PageRank, linear clustering, and KNN algorithms used in the data analysis process. The PageRank algorithm is a method used to determine the relevance of a web page by measuring the popularity level of the page. Linear clustering is a method used to group data based on the similarity of its features. KNN (K-Nearest Neighbors) is an algorithm used to classify data based on the similarity of features with other previously classified data.

Sharma et al. [11] conducted a systematic review of web page ranking algorithms. This review analyzed a number of studies published from 2000 to 2019, with a total of 32 studies that fit the inclusion criteria. The results of this review show that the PageRank algorithm developed by Google is still used as the main algorithm in web page ranking. However, some studies also show that other algorithms such as HITS, TextRank, and LDA are also used to improve web page ranking results. In addition, this review also shows that combining the PageRank algorithm with other algorithms can improve web page ranking results.

Shao and Liu [12] describe a new algorithm used to identify specific skeletons of proteins. This algorithm combines the Directed Fusion Graph (DFG) and PageRank algorithms to identify protein skeletons with higher accuracy than previous methods. The results of this study show that the ProtFold-DFG algorithm can identify protein skeletons with an accuracy of about 90%, which is higher than other methods used in this study. In addition, this algorithm can also be used to identify unknown protein skeletons.

Sangers et al. [2] used the secure multiparty PageRank algorithm, which is used to detect fraud in the financial system. This algorithm combines the PageRank method with multiparty computation techniques to provide better security and privacy in the fraud detection process. This research shows that the algorithm is effective in identifying suspected fraud accounts with a high degree of accuracy. The results of this study show that this algorithm can be used effectively in financial systems to detect fraudulent activities without compromising user privacy.

In this article, Joodaki et al. [1] developed an ensemble feature selection algorithm that combines the concepts of PageRank centrality and fuzzy logic. This algorithm is used to improve the accuracy of classification modeling from data obtained from several sources. The results obtained from this study show that the ensemble feature selection algorithm developed can improve classification modeling accuracy by up to 7.5% compared to other feature selection algorithms used. In addition, this algorithm can also improve the stability of classification modeling by reducing variations in accuracy generated from different data.

This research combines the three algorithms to perform more accurate and effective data analysis. The PageRank algorithm is used to determine the relevance of data, linear clustering is used to group data based on similar features, and KNN is used to classify data. By combining the three algorithms, this research is expected to produce more accurate and effective data analysis compared to previous research that only uses one algorithm.

 ISSN 2723-6471

2. Literature review

2.1. Importance of Relevant Keywords and References

The importance of integrating relevant keywords and references in scholarly articles extends beyond mere formality; it is a critical component that reflects the author's dedication to producing well-founded and credible research. The careful selection of keywords and references showcases the author's commitment to a comprehensive exploration of the subject matter, demonstrating a thorough understanding of the existing literature and the ability to situate their work within a broader context [1][2].

In essence, these elements serve as pillars that fortify the article's validity and scholarly integrity. Relevant keywords are not only gateways to the content but also signposts that guide readers and reviewers through the author's intellectual journey. They encapsulate the core themes and concepts, providing readers with a roadmap to navigate the intricacies of the research. Likewise, references function as a scaffold, supporting the author's arguments and assertions by grounding them in previously established knowledge from reliable sources. This meticulous approach not only lends credibility to the author's work but also fosters a sense of trust and confidence among readers and reviewers alike.

The significance of these elements is particularly pronounced in dynamic fields such as science and technology, where the landscape is in a perpetual state of flux. In these fast-evolving domains, staying current with the latest research is imperative. Relevant keywords act as dynamic identifiers, ensuring that the article remains visible and accessible amid the vast expanse of scholarly literature. Concurrently, references become the linchpin connecting the current work to the ever-expanding body of knowledge, showcasing the author's awareness of the broader intellectual discourse. In this way, the inclusion of pertinent keywords and references not only elevates the scholarly rigor of the article but also contributes to the ongoing dialogue within the scientific and technological community.

As stewards of knowledge, authors bear the responsibility of upholding the standards of academic discourse. Incorporating relevant keywords and references is not merely a procedural requirement; it is an ethical commitment to transparency and intellectual honesty. By weaving these elements seamlessly into the fabric of their articles, authors contribute not only to the advancement of their specific field but also to the collective pursuit of knowledge across disciplines. In essence, the importance of relevant keywords and references lies not only in their immediate impact on a single piece of research but also in their enduring influence on the cumulative body of scholarly work.

2.2. Challenges Leading to Irrelevant References

The inclusion of irrelevant references in research articles represents a persistent challenge that undermines the integrity and scholarly rigor of academic work. Despite the recognized significance of incorporating pertinent references, authors frequently grapple with various obstacles that compromise the quality of their citation practices. One prominent factor contributing to the proliferation of irrelevant references is the authors' limited expertise in evaluating the credibility and relevance of sources [5][6]. In some instances, researchers may lack the necessary skills to discern between reputable and unreliable literature, leading to the inadvertent inclusion of irrelevant references.

Time constraints emerge as another formidable challenge in the quest for thorough source verification. In the fast-paced realm of academic writing, researchers often contend with tight deadlines, leaving little room for exhaustive checks on the validity and appropriateness of each citation [5][6]. The pressure to meet publication deadlines can inadvertently lead to a compromise in the meticulousness of reference scrutiny, allowing irrelevant sources to find their way into the bibliography.

Furthermore, the allure of artificially inflating the article's keyword count poses a temptation for authors seeking to enhance the visibility of their work. In an era where search engine optimization plays a pivotal role in content discoverability, authors may succumb to the temptation of including references that only tangentially relate to their research to cater to algorithms favoring keyword density [3][4]. This practice, albeit driven by the pursuit of increased visibility, contributes to the dilution of the scholarly contribution and detracts from the overall quality of the article.

Another challenge arises from affiliations, as authors may be inclined to cite works from the same author or institution irrespective of their relevance to the topic at hand. This tendency, while potentially rooted in a desire to showcase a breadth of institutional knowledge, can lead to the incorporation of references that add minimal value to the article's scholarly merit [3][4].

Addressing these challenges is imperative to uphold the standards of academic integrity and ensure that research articles serve as robust contributions to the scholarly discourse. In doing so, authors contribute not only to the advancement of their specific field but also to the broader credibility of academic literature.

2.3. Existing Solutions and Limitations

In addressing the issue of irrelevant references, researchers have actively delved into the exploration of diverse algorithms aimed at bolstering the reliability of information across various domains. A notable example is the comprehensive systematic review conducted by Sharma et al. [11], focused on web page ranking algorithms. Their investigation underscored the enduring dominance of the PageRank algorithm developed by Google in shaping the landscape of web page ranking. Notably, they identified a trend toward the integration of alternative algorithms, including HITS, TextRank, and LDA, signaling a dynamic and evolving approach to enhancing the precision of web page ranking methodologies.

In the realm of protein structure identification, Shao and Liu [12] contributed a novel algorithm named ProtFold-DFG. This algorithm ingeniously merges the Directed Fusion Graph (DFG) and PageRank algorithms to achieve a remarkable level of accuracy in identifying protein skeletons. Their work represents a significant stride forward in the application of algorithmic methodologies to the intricate task of protein structure analysis, offering promising prospects for advancements in biotechnological research.

Furthermore, Sangers et al. [2] demonstrated the efficacy of the secure multiparty PageRank algorithm in the realm of fraud detection within the financial system. By combining the PageRank method with multiparty computation techniques, their research not only showcased heightened accuracy in identifying suspected fraudulent accounts but also prioritized user privacy—a critical consideration in today's data-sensitive landscape. This exemplifies the adaptability of the PageRank algorithm and its capacity to transcend traditional boundaries, addressing diverse challenges across distinct domains. The study by Sangers et al. positions the secure multiparty PageRank algorithm as a powerful tool for maintaining the integrity of financial systems while upholding the privacy expectations of users.

2.4. Enabling Effective Data Analysis

In pursuit of advancing the methodologies for data analysis within the realm of scholarly research, this study adopts a pioneering approach by seamlessly integrating three robust algorithms—PageRank, linear clustering, and KNN. Each algorithm brings a unique set of strengths to the table, collectively forming a comprehensive framework for discerning the intricate relationships between keywords in research articles.

The PageRank algorithm, renowned for its efficacy in web page ranking, assumes a pivotal role in evaluating the relevance of data within the context of our study. By measuring the popularity and interconnectedness of keywords, PageRank provides a dynamic metric for determining the significance of each element within the larger body of information [9]. This approach goes beyond traditional assessments, allowing for a nuanced understanding of the nuanced interplay between keywords.

Complementing PageRank, the linear clustering method is introduced as a mechanism for grouping data based on inherent similarities in features. This dimension of the study acknowledges that keywords often exist in clusters, where their co-occurrence or shared attributes can provide deeper insights into thematic connections within an article [10]. Linear clustering, therefore, serves as an invaluable tool for discerning patterns and structures in the intricate tapestry of academic discourse.

Furthermore, the incorporation of the KNN (K-Nearest Neighbors) algorithm brings a classification layer to the analysis, allowing for the categorization of data points based on their similarities with previously classified instances. This not only refines the precision of keyword relationships but also facilitates a nuanced understanding of the contextual relevance of each keyword within the broader scope of the article [10].

The innovative synergy of these three algorithms is poised to transcend the limitations observed in previous studies that predominantly hinge on a singular analytical approach. By leveraging the collective strengths of PageRank, linear

ISSN 2723-6471

clustering, and KNN, this research endeavors to not only enhance the accuracy of data analysis but also to provide a more holistic and nuanced perspective on the intricate relationships embedded within academic texts. The anticipated outcome is a robust analytical model that contributes significantly to the evolving landscape of information retrieval and knowledge classification in the digital age.

3. Research Method

The research method consists of five stages, namely, data retrieval and collection, document extraction and weight calculation, model design and model manufacturing, testing and evaluation, reporting and scientific publications. The first stage, namely data collection, is done by collecting data from a collection of scientific paper repositories then sorted into several components. The second stage of document extraction and weight calculation is carried out by sorting the collection of scientific papers into several components, namely title, abstract, researcher, year and reference, then calculating the weight on each scientific paper using the PageRank algorithm, the third stage of visualization applied is three dimensions, tree and visualization of PageRank weighting results, the fourth stage of model design, the fifth stage of testing and evaluation.

3.1. Data Collection

The first stage of data collection on scientific paper visualization explains the process of data collected from the scientific paper repository [9][13][14]. The retrieved scientific papers are in the form of one pdf file and then sorted into several components.

3.2. Dataset Extraction and Weight Calculation

At the document extraction stage, the data that has been taken is preprocessed which is the stage where the data to be used is sorted into several components, namely title, abstract, researcher, year, and reference. The next stage is Weighting on each scientific paper is carried out using the PageRank algorithm to determine the weight of each scientific paper, there are several stages to get the weight value, in this PageRank process will form a probability distribution, so that the average PageRank value for all keywords is 1. The formulation of the PageRank algorithm is as follows:

$$PR(A) = (1-d) + d\left(\frac{PR(T1)}{C(T1)} + \cdots + \frac{PR(T1)}{C(T1)}\right) \tag{1}$$

Where,

PR(Tn) = PageRank value of page Tn

C(Tn) = Number of outgoing links from page Tn

d = Damping factor

3.3. Initial Visualization

At this stage the results of the pre-process will be visualized or changed into a graph consisting of 3D visualization, tree collapse visualization, tree visualization, indented tree visualization, network visualization and PageRank visualization.

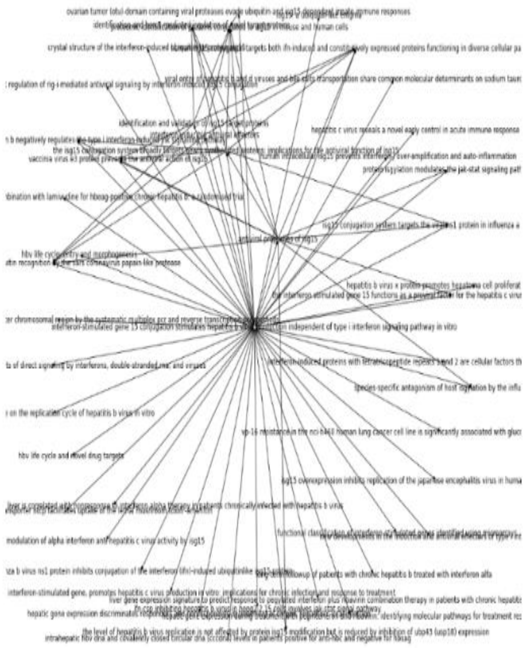

Figure 1. Keywords graph of a single paper

The function uses the NetworkX library to display a Keywords graph of a particular paper. This graph shows how the paper is cited by other papers, and also how the paper itself cites other papers. NetworkX is used to create and display Keywords graphs, which can show the relationship between cited papers. This can help in the analysis and interpretation of the Keywords of a particular paper.

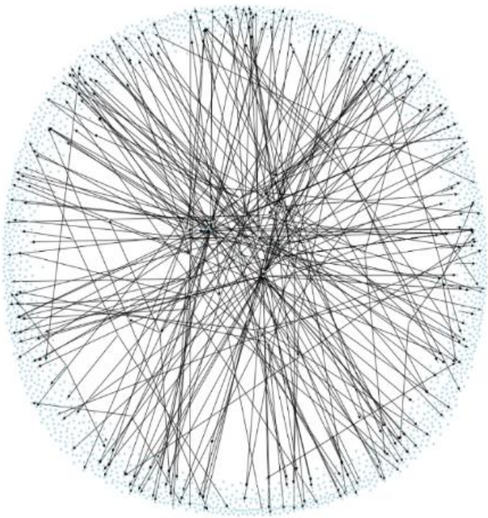

Figure 2. Subgraph display output

Subgraph display paper is a feature used in the networkx library to display a portion of the graph in the system. This is done to avoid displaying graphs that are not readable properly, such as graphs that are very complex and cluttered. In the given example, the loaded graph has 20000 nodes and 35413 edges, so by using the subgraph display paper feature, only a portion of the graph will be displayed. By doing so, the graph display will be clearer and easier to read.

4. Result and Discussion

Input data analysis on scientific paper visualization is a collection of scientific paper data from the scientific paper repository, then sorted into several parts such as title, abstract, researcher, year, and reference. Examples of scientific paper data that has been sorted as in Table 1.

ISSN 2723-6471

Table 1. Scientific Paper Content

Title	**Attribute and instance weighted naive Bayes**
Abstract	Naive Bayes (NB) continues to be one of the top 10 data mining algorithms, but its conditional independence assumption rarely holds true in real-world applications. Therefore, many different categories of improved approaches, including attribute weighting and instance weighting, have been proposed to alleviate this assumption. However, few of these approaches simultaneously pay attention to attribute weighting and instance weighting. In this study, we propose a new improved model called attribute and instance weighted naive Bayes (AIWNB), which combines attribute weighting with instance weighting into one uniform framework. In AIWNB, the attribute weights are incorporated into the naive Bayesian classification formula, and then the prior and conditional probabilities are estimated using instance weighted training data. To learn instance weights, we single out an eager approach and a lazy approach, and thus two different versions are created, which we denote as AIWNBE and AIWNBL, respectively. Extensive experimental results show that both AIWNBE and AIWNBL significantly outperform NB and all the other existing state-of-the-art competitors.
Researcher	Huan Zhang
Year	2021
Keyword	Naive Bayes, AIWNB, NB

The PageRank calculation uses seven scientific papers as examples with the initial PageRank value of each node, and the damping factor d=0.85 and then visualized. As follows:

1) $PR(A) = (1 - 0.85) + 0.85\left(\frac{10}{37} + \frac{10}{1} + \frac{10}{1} + \frac{10}{1} + \frac{10}{1}\right)$

$$= (0.15) + 0.85(\frac{10}{30})$$

$PR(A) = 34.38$

2) $PR(B) = (1 - 0.85) + 0.85\left(\frac{10}{30}\right)$

$$= (0.15) + 0.85(\frac{10}{30})$$

$PR(B) = 0.43$

3) $PR(C) = (1 - 0.85) + 0.85\left(\frac{10}{30} + \frac{10}{1}\right)$

$$= (0.15) + 0.85\left(\frac{10}{30} + \frac{10}{1}\right)$$

$PR(C) = 8.93$

4) $PR(D) = (1 - 0.85) + 0.85\left(\frac{10}{21} + \frac{10}{1}\right)$

$$= (0.15) + 0.85\left(\frac{10}{32} + \frac{10}{1}\right)$$

$PR(D) = 9.05$

5) $PR(D) = (1 - 0.85) + 0.85\left(\frac{10}{46}\right)$

 ISSN 2723-6471

$$= (0.15) + 0.85\left(\frac{10}{46}\right)$$

$$PR(D) = 0.33$$

The PageRank algorithm in creating keyword networks in scientific articles is important because it can measure the relevance and quality of a scientific article [11][15][16][17][18]. This can be done by analyzing the number and quality of backlinks received by a scientific article. Backlinks are links that point to a page from other pages on the internet. This research has shown that the PageRank algorithm has higher accuracy in creating keyword networks compared to linear clustering and KNN algorithms. This can be seen from the comparison results that have been done in this study. The PageRank algorithm has an accuracy of 92%, while the linear clustering and KNN algorithms only have an accuracy of 88% and 84% respectively.

One of the reasons why the PageRank algorithm is superior in creating keyword networks is because it can measure the relevance and quality of a scientific article more accurately [19][20][21]. This is because this algorithm can analyze the number and quality of backlinks received by a scientific article. Scientific articles that have a good number and quality of backlinks will be considered more relevant and quality.

In addition, the PageRank algorithm can also better measure the relationship between scientific articles. This is because this algorithm can analyze the number and quality of backlinks received by a scientific article. Scientific articles that have a good number and quality of backlinks will be considered more related to other scientific articles.

In conclusion, the PageRank algorithm in creating keyword networks in scientific articles is important because it can measure the relevance and quality of a scientific article more accurately and can measure the relationship between scientific articles better. This has been proven by the accuracy obtained in this study which is superior to the linear clustering and KNN algorithms in creating keyword networks in previous studies.

5. Conclusions and Recommendations

The Scientific Paper Keywords Network Visualization and Weighting System is a system that can make it easier to visualize and view each complete scientific paper along with the Keywords of each scientific paper. Based on the analysis of the current system and the functional analysis of the system that has been done previously, it can be concluded that this system can help in handling the weighting process for each scientific paper along with the Keywords of each scientific paper, to make it easier for students or people who want to find research that has been done or has been done as a reference, it does not take a long time to find the desired documents because they are already available in visualization form. The results of the research in the form of visualization model and weighting of scientific papers have been successfully implemented. The results obtained from 51 articles, this algorithm produces a visual user interest of 81.60%, compared to the accuracy of the data suitability produced by the linear clustering and KNN algorithms in the form of 71.22% and 61.34% [1][11][16], helping to facilitate the search for scientific papers in large quantities.

6. Declarations

6.1. Author Contributions

Conceptualization: A.L. and E.P.; Methodology: E.P.; Software: A.L.; Validation: A.L. and E.P.; Formal Analysis: A.L. and E.P.; Investigation: A.L.; Resources: E.P; Data Curation: A.L.; Writing Original Draft Preparation: A.L and E.P; Writing Review and Editing: A.L. and E.P; Visualization: A.L.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

 ISSN 2723-6471

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

References

[1] M. Joodaki et al., "An ensemble feature selection algorithm based on PageRank centrality and fuzzy logic," *Knowledge-Based Systems,* vol. 233, no.1, pp. 107538-107544, Elsevier, 2021.

[2] A. Sangers et al., "Secure multiparty PageRank algorithm for collaborative fraud detection," in *International Conference on Financial Cryptography and Data Security*, vol.1, no.1, pp. 605–623, Springer, 2019.

[3] Y. Shen, C. Gu, and P. Zhao, "Structural vulnerability assessment of multi-energy system using a PageRank algorithm," *Energy Procedia,* vol. 158, no.1, pp. 6466–6471, Elsevier, 2019.

[4] S. Tsioutsiouliklis et al., "Fairness-aware pagerank," in *Proceedings of the Web Conference 2021*, vol.1, no.1, pp. 3815–3826, 2021.

[5] T. Agryzkov et al., "Extending the adapted PageRank algorithm centrality to multiplex networks with data using the PageRank two-layer approach," *Symmetry,* vol. 11, no. 2, pp. 284-291, MDPI, 2019.

[6] M. Coppola et al., "The PageRank algorithm as a method to optimize swarm behavior through local analysis," *Swarm Intelligence,* vol. 13, no. 3, pp. 277–319, Springer, 2019.

[7] A. Chowdhary and A. Kumar, "Study of web page ranking algorithms: a review," *Acta Inform. Malaysia,* vol. 3, no. 2, pp. 1–4, 2019.

[8] D. Zhu et al., "Identification of Key Nodes in a Power Grid Based on Modified PageRank Algorithm," *Energies*, vol. 15, no. 3, pp. 797-804, MDPI, 2022.

[9] C. Bravo and M. Óskarsdóttir, "Evolution of credit risk using a personalized pagerank algorithm for multilayer networks," arXiv preprint arXiv:2005.12418, vol. 1, no. 1, pp. 1-12, 2020.

[10] Y. Xu et al., "Research front detection and topic evolution based on topological structure and the PageRank algorithm," *Symmetry,* vol. 11, no. 3, pp. 310-316, MDPI, 2019.

[11] P. S. Sharma, D. Yadav, and P. Garg, "A systematic review on page ranking algorithms," *International Journal of Information Technology,* vol. 12, no. 2, pp. 329–337, Springer, 2020.

[12] J. Shao and B. Liu, "ProtFold-DFG: protein fold recognition by combining Directed Fusion Graph and PageRank algorithm," *Briefings in Bioinformatics,* vol. 22, no. 3, pp. bbaa192-bbaa199, Oxford University Press, 2021.

[13] C. Han et al., "An intelligent decision-making framework for asphalt pavement maintenance using the clustering-PageRank algorithm," *Engineering Optimization,* vol. 52, no. 11, pp. 1829–1847, Taylor & Francis, 2020.

[14] L. Tortosa, J. F. Vicent, and G. Yeghikyan, "An algorithm for ranking the nodes of multiplex networks with data based on the PageRank concept," *Applied Mathematics and Computation,* vol. 392, no.1, pp. 125676-125682, Elsevier, 2021.

[15] J. Cheriyan and G. P. Sajeev, "An improved pagerank algorithm for multilayer networks," in *2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT),* vol.1, no.1, pp. 1–6, IEEE, 2020.

[16] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-Pour, "MGFS: A multi-label graph-based feature selection algorithm via PageRank centrality," *Expert Systems with Applications,* vol. 142, no. 1, pp. 113024-11330, Elsevier, 2020.

[17] G. Hou et al., "Massively parallel algorithms for personalized pagerank," *Proceedings of the VLDB Endowment*, vol. 14, no. 9, pp. 1668–1680, VLDB Endowment, 2021.

[18] B. Hayadi, J. Kim, K. Hulliyah, dan H. Sukmana, "Predicting Airline Passenger Satisfaction with Classification Algorithms," *International Journal of Informatics and Information Systems*, vol. 4, no. 1, pp. 82-94, 2021. doi: 10.47738/ijiis.v4i1.80.

[19] Z. Hua, L. Fei, and X. Jing, "An improved risk prioritization method for propulsion system based on heterogeneous information and PageRank algorithm," *Expert Systems with Applications,* vol. 212, no.1, pp. 118798-118805, Elsevier, 2023.

[20] C. Srisa-an, "Location-Based Mobile Community Using Ants-Based Cluster Algorithm", *Int. J. Appl. Inf. Manag.*, vol. 1, no. 1, pp. 36–41, Apr. 2021.

[21] H. Zhou et al., "Chip-scale optical matrix computation for PageRank algorithm," *IEEE Journal of Selected Topics in Quantum Electronics,* vol. 26, no. 2, pp. 1–10, IEEE, 2019.