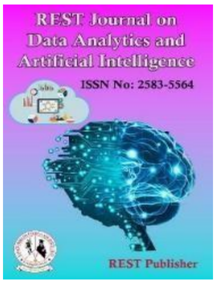
REST Journal on Data Analytics and Artificial Intelligence
Vol: 3(1), March 2024
REST Publisher; ISSN: 2583-5564
Website: http: //restpublisher.com/journals/jdaai/
DOI: https://doi.org/10.46632/jdaai/3/1/7



Graph Theory and Algorithms for Social Network Analysis

***Sreenath Devineni, Bhargavi Gorantla**
Lead Data Engineer, USA.
**Corresponding author: Srinathdevineni@gmail.com*

Abstract: *Graph theory and algorithms play a crucial role in the analysis of social networks, which are complex systems composed of individuals or entities connected by various types of relationships. This paper provides an overview of graph theory concepts and algorithms commonly used in social network analysis. It discusses how these tools can be applied to understand the structure and dynamics of social networks, identify key actors or communities, and analyze information flow and influence propagation. The paper also highlights some of the challenges and future directions in the field of social network analysis using graph theory. The study of graph theory and algorithms for social network analysis is of significant importance due to its wide-ranging applications in various fields. Understanding the structure and dynamics of social networks can provide valuable insights into human behavior, information diffusion, organizational dynamics, and the spread of diseases or ideas. By applying graph theory concepts such as centrality measures, community detection algorithms, and link prediction techniques, researchers can uncover hidden patterns, identify influential nodes or groups, and predict future interactions in social networks. This knowledge can be used to design more effective strategies for marketing, public health interventions, social media management, and other areas where understanding and influencing human behavior are critical. Versatile with distinctive options. MOORA is a new optimization approach that has been proposed. This objective method signifies a matrix of potential replies, but it also proposes better policies based on the rates employed. Well established. For comparison, the reference point method is another multi-objective optimization technique. After several competitions, this proved to be the best pick among the methods. From the result network models is in first position whereas Graph density is in 5th position*

Keywords: *Social network analysis, Graph sampling algorithms, Gephi, Pajek, and NodeXL*

1. INTRODUCTION

Social network analysis tools are software applications that facilitate the study of social networks. These tools provide functionalities for data collection, visualization, and analysis of network structures and dynamics. They often include algorithms for calculating network metrics such as centrality, clustering coefficient, and community detection. Some popular social network analysis tools include Gephi, Pajek, and NodeXL, which offer user-friendly interfaces for researchers and practitioners to explore and analyze social networks. These tools play a crucial role in understanding complex social systems and can inform decision-making in various fields, including sociology, anthropology, and marketing. [1] Graph sampling algorithms for social network analysis are methods used to extract a subset of nodes or edges from a large graph. These algorithms aim to preserve the structural properties of the original graph while reducing its size for more efficient analysis. Common techniques include random node sampling, snowball sampling, and random walk-based sampling [2] Graph pattern matching in social network analysis is the process of finding specific subgraph structures within a larger graph that match a predefined pattern. This is essential for identifying motifs, recurring configurations, or specific relationships of interest in social networks. Algorithms such as subgraph isomorphism and graph simulation are commonly used to efficiently search for these patterns. Pattern matching helps in understanding the underlying structure and dynamics of social networks, aiding in various analyses and predictions.[3] Connectedness refers to the property of a graph where there is a path between every pair of vertices. Connectivity, on the other hand, is a measure of how connected a graph is and is often quantified by the minimum number of vertices or edges that need to be removed to disconnect the graph. Understanding connectedness and connectivity is crucial in social network analysis as they help determine the robustness, information flow, and overall structure of social networks, influencing various network measures and analyses. Efficient algorithms for incremental all pairs shortest paths, closeness, and between' s in social network analysis focus on updating these metrics efficiently when the graph changes. These algorithms aim to reduce the computational cost of re-calculating these metrics from scratch by leveraging the existing information and structure of the graph.[5] Social network analysis encompasses various aspects, including network structure, node attributes, and dynamics. Network structure examines how nodes are connected, identifying key features like density, centrality, and clustering. Node attributes consider characteristics of individual nodes, such as age, gender, or interests, which influence network behaviors. Dynamics explore how networks evolve over time, studying phenomena like information diffusion or community formation. Understanding these aspects helps in predicting behavior, identifying influential nodes, and designing interventions in social networks. Social network analysis thus provides a comprehensive framework for studying complex interactions within social systems. [6] A random graph generation algorithm for the analysis of social networks is a method to create synthetic networks with properties similar to real-world social networks. These algorithms typically generate graphs based on specified parameters such as the number of nodes, average degree, and clustering coefficient, allowing researchers to compare and validate their findings against known network properties.[7] Social influence modeling using information theory in mobile social networks involves quantifying the flow of influence between individuals based on information diffusion. By applying concepts like entropy and mutual information, researchers can analyze how information spreads through the network, identifying key influencers and understanding the dynamics of information flow in mobile social networks.[8] Efficient algorithms for social network coverage and reach aim to identify the minimum set of nodes that can maximize the coverage or reach of information in a network. These algorithms consider factors like node influence and network structure to optimize the selection of nodes for broadcasting messages or maximizing the spread of information [9] Efficient algorithms for social network coverage and reach focus on selecting a subset of nodes in a network to maximize the spread or coverage of information. These algorithms consider factors such as node influence and network structure to identify the most influential nodes or optimal paths for disseminating information efficiently in a social network [10] An analysis of overlapping community detection algorithms in social networks involves evaluating techniques that identify nodes belonging to multiple communities. These algorithms, such as the Clique Percolation Method or the Node-Centric Overlapping Community Detection algorithm, aim to uncover the complex and overlapping nature of communities in social networks, providing insights into the network's structure and function.[12] "Nature-inspired link prediction and community detection algorithms for social networks: a survey" is a comprehensive review of computational methods inspired by natural phenomena, such as evolutionary algorithms or swarm intelligence, used to predict links or detect communities in social networks. These algorithms mimic natural processes to enhance the understanding of complex network structures and dynamics.[14]

2. MATERIALS AND METHODS

2.1. Alternative parameter: Graph Density, Centrality Measures, Community Detection Algorithms, Network Models, Graph Embedding Techniques,

2.2. Evaluation parameter: Scalability, Accuracy, Robustness, Interpretability.

2.3. Graph Density: Graph density is a measure of how "dense" or connected a graph is. It is defined as the ratio of the number of edges present in the graph to the total number of possible edges in the graph.

2.4. Centrality Measures: Centrality measures in graph theory are metrics used to quantify the importance or influence of individual vertices (nodes) within a graph. These measures help identify key nodes that play significant roles in the network.

2.5. Community Detection Algorithms: Community detection algorithms are techniques used to identify groups of nodes in a network that are more densely connected to each other than to the rest of the network. These groups are often referred to as "communities" or "clusters," and community detection algorithms aim to uncover these structures in complex networks.

2.6. Network Models: Network models refer to mathematical or computational representations of networks, which are comprised of nodes (vertices) and edges (links) connecting these nodes. These models are used to study the structure, behavior, and properties of networks in various fields, including computer science, sociology, biology, and transportation, among others.

2.7. Graph Embedding Techniques: Graph embedding techniques are methods used to represent nodes and edges in a graph as low-dimensional vectors in a continuous vector space. These techniques aim to capture the structural and relational information of the graph in a way that preserves important properties and relationships between nodes. Graph embedding techniques are particularly useful for tasks such as node classification, link prediction, and graph visualization, as they allow machine learning algorithms to operate on graphs in a vector space, where traditional machine learning techniques can be applied.

2.8. Scalability: Scalability refers to the ability of a system, network, or process to handle a growing amount of work or its potential to accommodate growth. In other words, scalability is the capability of a system to efficiently scale up or scale out in order to meet increased demand or workload.

2.9. Accuracy: Accuracy is a measure of the correctness of a model's predictions compared to the actual outcomes in a dataset. It is commonly used to evaluate classification models but can be applied to any model that produces discrete outcomes.

2.10. Robustness: Robustness can be evaluated using various techniques, including sensitivity analysis, where the model's performance is tested under different conditions or with perturbed inputs, and adversarial testing, where the model is tested against intentionally crafted inputs designed to fool the model. Improving the robustness of a model often involves techniques such as data augmentation, regularization, ensemble methods, and designing models that are inherently more resistant to noise and adversarial attacks.

2.11. Interpretability: Interpretability refers to the ability to understand and explain the decisions or predictions made by a machine learning model. A model is considered interpretable if its inner workings can be easily understood by humans, typically through transparent rules, visualizations, or feature importance scores.

2.12. Method: Rational multi-objective analysis (MOORA). This optimization was accomplished. The second MOORA characteristic is dimensionless numbers. This will serve as the foundation. Finally, it contrasts the discrepancies in well-being throughout Lithuania's ten counties in light of all of the objectives. The three opulent districts are in stark contrast to those of the less fortunate. The movement of workers from different regions to Vilnius is a crucial issue that represents income. Automatic redistribution is deemed unacceptable. Instead, commercialization and industrialization should emerge in specific areas [15]. Concrete multi-objective optimization. The system can be improved simultaneously within constraints or with more conflicting attributes (notes). There are several areas where the optimal selections must be taken, including product design issues and multi-goal optimization. 2. When there are commercial transactions, there may be conflicts of interest. Increasing sales and decreasing product costs, improving performance while lowering automotive fuel consumption, reducing weight while exacerbating difficulties, and [16]. First, Moora refers to a brand-new MCTM technique that was developed with knowledge of the shortcomings of more classic procedures. We therefore believed it should be completely practical. The second reason is the processing time required by MOORA to resolve the issue, as documented in the MCDM literature. Finally, MOORA requires little to no setup, as the literature suggests that it takes time and has a consistent personality [17]. The institution has a decision-making tool called MOORA, which may be utilized to address a wide range of situations. Scholarship candidates can be chosen quickly using a machine selection technique, thereby improving educational achievement and aiding needy pupils [18]. MOORA is simply amazing. A green multi-criteria selection method for a comprehensive study of possibilities that takes into account high heterogeneity and a variety of useful components. The MOORA technique is presented to effectively resolve complex decision-making challenges. This method usually yields grades that are rigidly contradictory. Considers and strives to select the best solution while taking into account both positive and negative standards. Some MOORA decisions are awarded based on their technique [19]. MOORA is a method for multi-objective optimization. There are several characteristics and approaches that some people utilize to go through and improve at the same time. MOORA is all about trying new things. An effective approach strategy. Constraints [20]. The MOORA approach can recall all attributes and their weights, allowing for a more thorough examination of alternatives. The MOORA technique may be straightforward to learn and implement. The proposed strategy is generic and adaptable to any size or quality. Combining the features results in more precise targeting and a simpler decision-making process. Furthermore, this method is applicable to all types of choice problems [21]. MOORA, or multiple criteria or multiple features, refers to multi-goal optimization based mostly on ratio analysis. Optimization is an upgrade process that takes into account two or more disputed attributes at the same time (notes). This timed provides a wide range of tools for decision-making in the disputed and complex environment of the supply chain. Examples include selecting the warehouse location, supplier, product, and procedure design. MOORA can be used when the finest options are required [22]. According to the failure prioritization produced by the usage of the extension in MOORA, it is clear that every single identified failure is ranked as an excellent priority. In other words, the suggested solution aims to alleviate a number of key shortcomings of RPN score as well as the selection process in regular MOORA. Provides dependability by connecting the use of range concepts. Finally, present logical conclusions to the decision-maker. Using this procedure When the results are compared to the two different traditional approaches, it is clear that disasters are fully prioritized and discovered [23].MOORA'S ANALYSIS Again, earlier scholars' research is more recent, and as a result, MOORA and MOOSRA approaches are assumed to use the most current statistics accessible. For the first method of selection Basically. As a result of the discussion above, the MOORA and MOOSRA methods are utilized to solve the choice problem. Complementary, resulting in diversity and non-conventional This method is quite reliable in a production environment. When articulated, this ratio is beneficial at the expense of the denominator. It is a preferred method of assessing economic welfare because the ratio's value becomes the same. As a result, MOORA and MOOSRA methodology are ideologically consistent with other mounting performance evaluation approaches [24].Both the ratio device and the benchmark MOORA approach include components. We select the type and importance of goals and options because our simulation of port planning is all that matters to us. The concerned parties include municipal, state, and federal governments, as well as cooperative organizations. Only implicitly is consumer sovereignty linked to the industrial process. However, authorities have also been recognized as valid clients' representatives [25]. Teamwork by MOORA Subjective, inconsistent, and contradictory information is provided by CNC machine tool to solve value issues and create an atmosphere conducive to decision making. Because this time period mixes fuzziness and helps decision-makers integrate a range of fuzzily articulated language variables. This page discusses the various MULTI-MOORA Ranking orders supplied by the regions. The results are summarized by comparison [26].

3. RESULT AND DISCUSSION

TABLE 1. Graph theory and algorithms for social network analysis

	Scalability	Accuracy	Robustness	Interpretability
Graph Density	0.75	0.9	0.8	50.00
Centrality Measures	0.7	0.85	0.75	45.00
Community Detection Algorithms	0.65	0.8	0.7	40.00
Network Models	0.6	0.75	0.65	30.00
Graph Embedding Techniques	0.55	0.7	0.6	35.00

Table 1 shows compare above values. Scalability: Graph Density has the highest value (0.75), indicating better scalability compared to the other alternatives. Accuracy: Graph Density also has the highest value (0.9), indicating better accuracy. Robustness: Graph Density and Centrality Measures tie for the highest value (0.8), indicating equal robustness. Interpretability: Graph Density has the highest value (50.00), indicating better interpretability. Based on this comparison, Graph Density appears to perform the best across the different evaluation parameters, followed by Centrality Measures, Community Detection Algorithms, Network Models, and Graph Embedding Techniques.

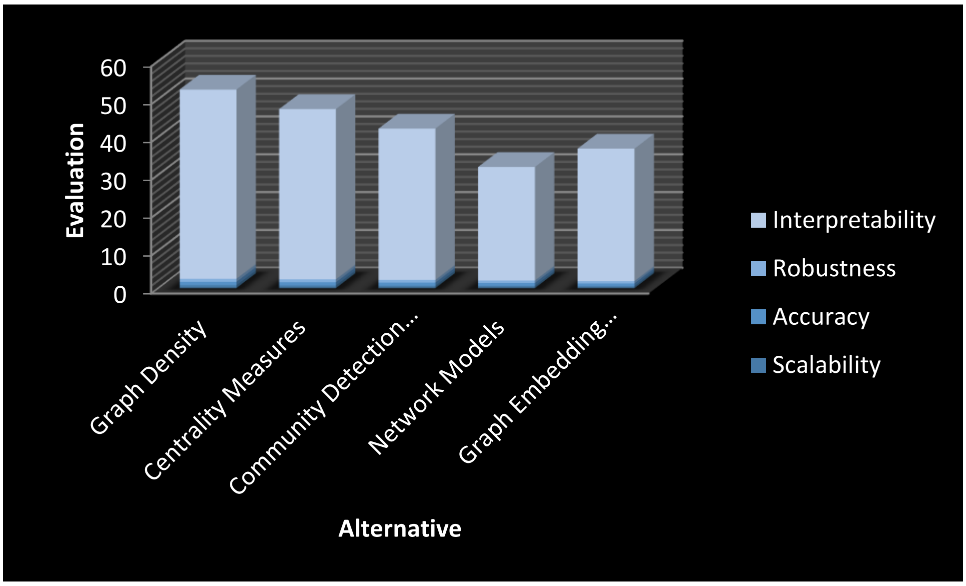


FIGURE1. Graph theory and algorithms for social network analysis

Figure1 illustrates the graphical representation of Graph theory and algorithms for social network analysis

TABLE 2. Divide & Sum

0.5625	0.8100	0.6400	2500.0000
0.4900	0.7225	0.5625	2025.0000
0.4225	0.6400	0.4900	1600.0000
0.3600	0.5625	0.4225	900.0000
0.3025	0.4900	0.3600	1225.0000
2.1375	**3.2250**	**2.4750**	**8250.0000**

Table 2 shows divide and sum To calculate the average value for each alternative across the evaluation parameters, Divide the sum of the values for each alternative by the number of parameters. Graph Density: Sum: 0.5625 + 0.81 + 0.64 + 2500 = 3101.0125, Average: 3101.0125 / 4 = 775.2531. Centrality Measures: Sum: 0.49 + 0.7225 + 0.5625 + 2025 = 2518.715, Average: 2518.715 / 4 = 629.6788. Community Detection Algorithms: Sum: 0.4225 + 0.64 + 0.49 + 1600 = 2041.5525, Average: 2041.5525 / 4 = 510.3881. Network Models: Sum: 0.36 + 0.5625 + 0.4225 + 900 = 1863.885, Average: 1863.885 / 4 = 465.9713. Graph Embedding Techniques: Sum: 0.3025 + 0.49 + 0.36 + 1225 = 2017.8525, Average: 2017.8525 / 4 = 504.4631.

TABLE 3. Normalized Data

	Normalized Data			
	Scalability	**Accuracy**	**Robustness**	**Interpretability**
Graph Density	0.5130	0.5012	0.5085	0.5505
Centrality Measures	0.4788	0.4733	0.4767	0.4954
Community Detection Algorithms	0.4446	0.4455	0.4449	0.4404
Network Models	0.4104	0.4176	0.4132	0.3303
Graph Embedding Techniques	0.3762	0.3898	0.3814	0.3853

$$X_{n1} = \frac{X1}{\sqrt{(X1)^2 + (X2)^2 + (X3)^2 \ldots}} \ldots \ldots \ldots \ldots \ldots (1)$$

Table 3 shows the various Normalized Data, Alternative: Graph Density, Centrality Measures, Community Detection Algorithms, Network Models, and Graph Embedding Techniques. Evaluation preference: Scalability, Accuracy, Robustness, Interpretability. Normalized value is obtained by using the formula (1).

TABLE 4. Weight

Weight			
0.25	0.25	0.25	0.25
0.25	0.25	0.25	0.25
0.25	0.25	0.25	0.25
0.25	0.25	0.25	0.25
0.25	0.25	0.25	0.25

Table 4 shows weight matrix

TABLE 5. Weighted normalized decision matrix

Weighted normalized decision matrix			
0.1282	0.1253	0.1271	0.1376
0.1197	0.1183	0.1192	0.1239
0.1111	0.1114	0.1112	0.1101
0.1026	0.1044	0.1033	0.0826
0.0940	0.0974	0.0953	0.0963

$$X_{wnormal1} = X_{n1} \times w_1 \dots\dots\dots\dots \quad (2)$$

Table 5 shows the weighted normalized decision matrix. Alternative: Graph Density, Centrality Measures, Community Detection Algorithms, Network Models, and Graph Embedding Techniques. Evaluation preference: Scalability, Accuracy, Robustness, Interpretability. The weighted default result is calculated using the matrix formula (2)

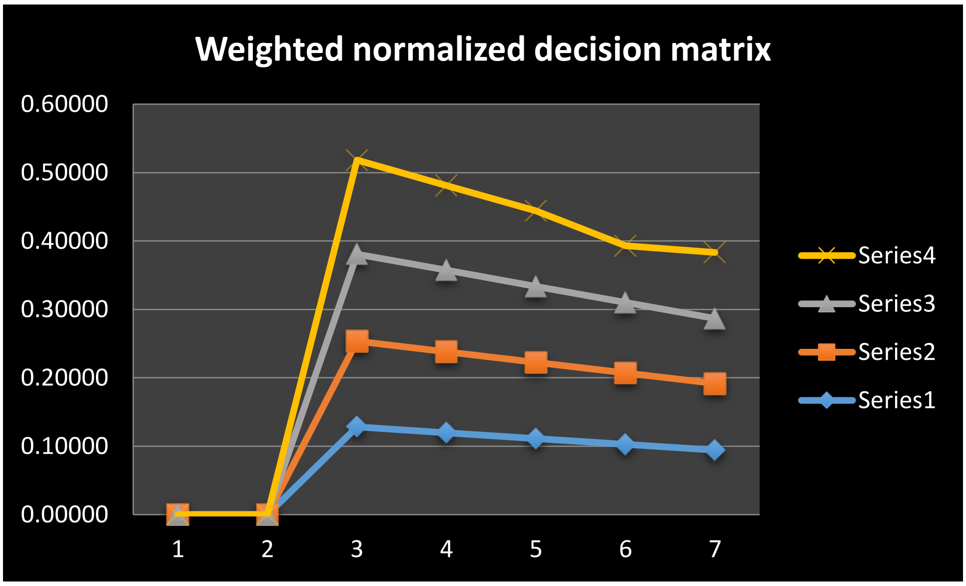


FIGURE 2. Weighted normalized decision matrix.

Figure 2 illustrate the graphical representation of weighted normalized decision matrix.

TABLE 6. Assessment value& Rank

	Assessment value	**Rank**
Graph Density	-0.0112	5
Centrality Measures	-0.0050	4
Community Detection Algorithms	0.0012	2
Network Models	0.0211	1
Graph Embedding Techniques	-0.0002	3

$$\text{Assesment value} = \sum X_{wn1} + X_{wn2} - X_{wn3} \quad (3).$$

Table 6 shows the Assessment value& Rank used. The Assessment value for Graph Density -0.0112, Centrality Measures -0.0050, Community Detection Algorithms 0.0012, Network Models.0.0211, Graph Embedding Techniques -0.0002, the final rank of this paper, Graph Embedding Techniques is in 3rd rank, Community Detection Algorithms is in 2nd rank, Centrality Measures is in 4th rank, Graph Density is in 5th rank, and the Network Models is in 1st rank. The final result is done by using the Moora method.

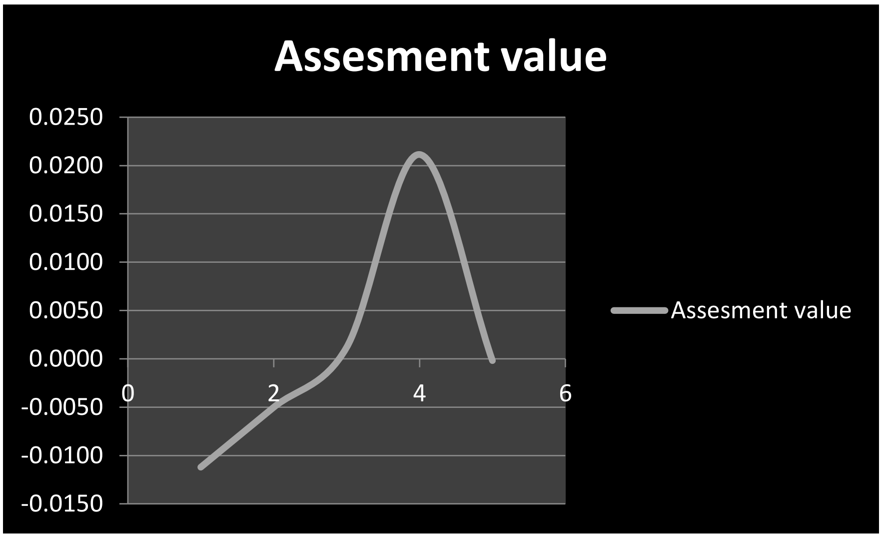


FIGURE 3. Assessment value

The Assessment value for Graph Density -0.0112, Centrality Measures -0.0050, Community Detection Algorithms 0.0012, Network Models.0.0211, Graph Embedding Techniques -0.0002.

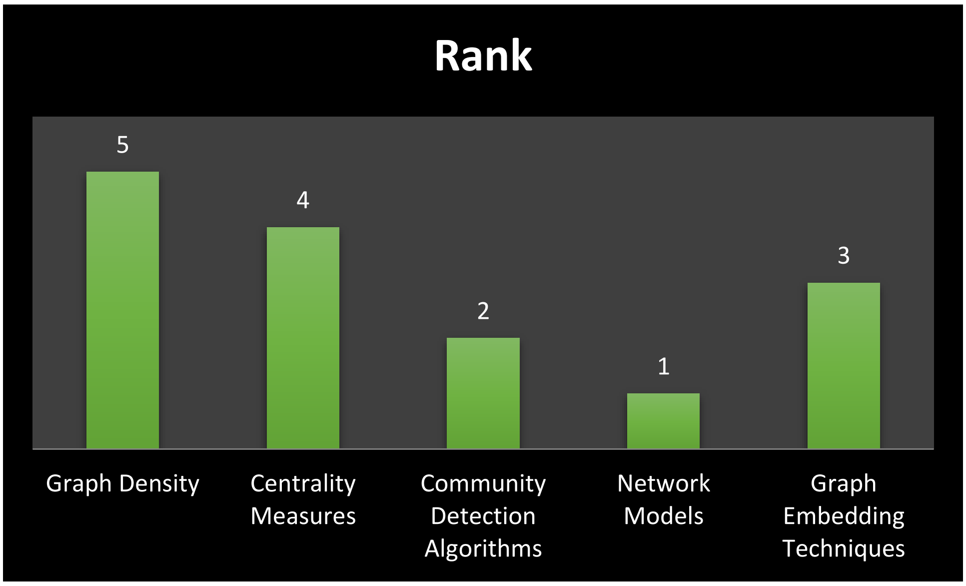


FIGURE 4. Rank

Figure 4 shows rank. Graph Embedding Techniques is in 3rd rank, Community Detection Algorithms is in 2nd rank, Centrality Measures is in 4th rank, Graph Density is in 5th rank, and the Network Models is in 1st rank. The final result is done by using the Moora method.

4. CONCLUSION

In analysing graph theory and algorithms for social network analysis, it is evident that different algorithms and approaches have varying impacts on the analysis process. Algorithms like Page Rank and community detection algorithms are crucial for identifying key nodes and communities within networks, aiding in understanding network structures and dynamics. However, the choice of algorithm should be aligned with the specific goals of the analysis, as some algorithms may prioritize computational efficiency over accuracy or vice versa. Overall, a nuanced understanding of graph theory and algorithm selection is essential for effective social network analysis, ensuring that the chosen methods align with the desired outcomes.

REFERENCES

[1]. Akhtar, Nadeem. "Social network analysis tools." In 2014 fourth international conference on communication systems and network technologies, pp. 388-392. IEEE, 2014.
[2]. Wang, Tianyi, Yang Chen, Zengbin Zhang, Tianyin Xu, Long Jin, Pan Hui, Beixing Deng, and Xing Li. "Understanding graph sampling algorithms for social network analysis." In 2011 31st international conference on distributed computing systems workshops, pp. 123-128. IEEE, 2011.
[3]. Fan, Wenfei. "Graph pattern matching revised for social network analysis." In Proceedings of the 15th International Conference on Database Theory, pp. 8-21. 2012.
[4]. Barnes, John A. "Graph theory and social networks: A technical comment on connectedness and connectivity." Sociology 3, no. 2 (1969): 215-232.
[5]. Khopkar, Sushant S., Rakesh Nagi, Alexander G. Nikolaev, and Vaibhav Bhembre. "Efficient algorithms for incremental all pairs shortest paths, closeness and betweenness in social network analysis." Social Network Analysis and Mining 4 (2014): 1-20.
[6]. Jamali, Mohsen, and Hassan Abolhassani. "Different aspects of social network analysis." In 2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2006 Main Conference Proceedings)(WI'06), pp. 66-72. IEEE, 2006.
[7]. Morris, James F., Jerome W. O'Neal, and Richard F. Deckro. "A random graph generation algorithm for the analysis of social networks." The Journal of Defense Modeling and Simulation 11, no. 3 (2014): 265-276.

[8]. Peng, Sancheng, Aimin Yang, Lihong Cao, Shui Yu, and Dongqing Xie. "Social influence modeling using information theory in mobile social networks." Information Sciences 379 (2017): 146-159.
[9]. Puthal, Deepak, Surya Nepal, Cecile Paris, Rajiv Ranjan, and Jinjun Chen. "Efficient algorithms for social network coverage and reach." In 2015 IEEE International Congress on Big Data, pp. 467-474. IEEE, 2015.
[10]. Puthal, Deepak, Surya Nepal, Cecile Paris, Rajiv Ranjan, and Jinjun Chen. "Efficient algorithms for social network coverage and reach." In 2015 IEEE International Congress on Big Data, pp. 467-474. IEEE, 2015.
[11]. Devi, J. Chitra, and E. Poovammal. "An analysis of overlapping community detection algorithms in social networks." Procedia Computer Science 89 (2016): 349-358.
[12]. Pulipati, Srilatha, Ramasubbareddy Somula, and Balakesava Reddy Parvathala. "Nature inspired link prediction and community detection algorithms for social networks: a survey." International Journal of System Assurance Engineering and Management (2021): 1-18.
[13]. Chakraborty, Shankar. "Applications of the MOORA method for decision making in manufacturing environment." The International Journal of Advanced Manufacturing Technology 54, no. 9 (2011): 1155-1166.
[14]. Akkaya, Gökay, BetülTuranoğlu, and Sinan Öztaş. "An integrated fuzzy AHP and fuzzy MOORA approach to the problem of industrial engineering sector choosing." Expert Systems with Applications 42, no. 24 (2015): 9565-9573.
[15]. Siregar, Victor MarudutMulia, Mega RomaulyTampubolon, EkaPratiwiSeptaniaParapat, Eve Ida Malau, and Debora Silvia Hutagalung. "Decision support system for selection technique using MOORA method." In IOP Conference Series: Materials Science and Engineering, vol. 1088, no. 1, p. 012022. IOP Publishing, 2021.
[16]. Dabbagh, Rahim, and Samuel Yousefi. "A hybrid decision-making approach based on FCM and MOORA for occupational health and safety risk analysis." Journal of safety research 71 (2019): 111-123.
[17]. Sutarno, S., M. Mesran, S. Supriyanto, Y. Yuliana, and A. Dewi. "Implementation of Multi-Objective Optimazation on the Base of Ratio Analysis (MOORA) in Improving Support for Decision on Sales Location Determination." In Journal of Physics: Conference Series, vol. 1424, no. 1, p. 012019. IOP Publishing, 2019.
[18]. Attri, Rajesh, and Sandeep Grover. "Decision making over the production system life cycle: MOORA method." International Journal of System Assurance Engineering and Management 5, no. 3 (2014): 320-328.
[19]. Dey, Balaram, BipradasBairagi, Bijan Sarkar, and SubirSanyal. "A MOORA based fuzzy multi-criteria decision making approach for supply chain strategy selection." International Journal of Industrial Engineering Computations 3, no. 4 (2012): 649-662.
[20]. Ghoushchi, SaeidJafarzadeh, Samuel Yousefi, and Mohammad Khazaeili. "An extended FMEA approach based on the Z-MOORA and fuzzy BWM for prioritization of failures." Applied Soft Computing 81 (2019): 105505.
[21]. Sarkar, Asis, S. C. Panja, Dibyendu Das, and Bijon Sarkar. "Developing an efficient decision support system for non-traditional machine selection: an application of MOORA and MOOSRA." Production & Manufacturing Research 3, no. 1 (2015): 324-342
[22]. Brauers, Willem Karel M. "Multi-objective seaport planning by MOORA decision making." Annals of Operations Research 206, no. 1 (2013): 39-58.
[23]. Sahu, Anoop Kumar, Nitin Kumar Sahu, and Atul Kumar Sahu. "Application of modified MULTI-MOORA for CNC machine tool evaluation in IVGTFNS environment: an empirical study." International Journal of Computer Aided Engineering and Technology 8, no. 3 (2016): 234-259.