Identifying Influential Nodes in Complex Networks: Combining Global and Local Views

[1]D. Varshitha, [2]A. Iswarya,[3] P. JanardhanaRao, [4]J. Anil Kumar, [5]D. Anusha
[1,2,3,4] B. Tech Students, Department of CSE(AI&ML), SRK Institute Technology, Vijayawada, A.P, India
[5] Assistant Professor, Department of CSE(AI&ML), SRK Institute Technology, Vijayawada, A.P, India

Abstract: Influential node identification in complex networks is an important problem with implications in disease spread, viral marketing, rumor spreading, and opinion tracking. Classical centrality measures tend to miss the subtle significance of nodes by considering only global or local structural features. This article presents two measures of centrality: Global Relative Average Centrality (GRAC) and Local Relative Average Centrality (LRAC). GRAC measures the relative change in a node's centrality at the global level of the network after its removal, and LRAC measures the effect of node removal on local neighborhood structure. With the SIR model, we illustrate how GRAC and LRAC perform better than basic centrality measures in finding influential nodes and provide a more holistic measure of node importance in complex networks.

Keywords: Influential Nodes, Centrality Measures, Global Relative Average Centrality (GRAC), Local Relative Average Centrality (LRAC), Susceptible-Infected-Recovered (SIR) Model.

INTRODUCTION

Complex networks, such as social, biological, and technological networks, are dynamic systems where the flow of information, resources, and influence plays a pivotal role in shaping their behavior and evolution. Identifying key nodes that drive these processes is essential for a wide range of applications, including controlling disease outbreaks, optimizing communication networks, and maximizing the spread of innovations.

Conventional measures of centrality, including degree centrality, betweenness centrality, and closeness centrality, have been extensively used to measure node importance in networks. These conventional measures tend to give an incomplete view by paying attention to only global or local structural characteristics. This shortcoming has motivated more advanced centrality measures that bring together both the global and the local viewpoints.

In this work, we introduce two centrality measures: Global Relative Average Centrality (GRAC) and Local Relative Average Centrality (LRAC). GRAC quantifies the relative difference in a node's centrality at the global network level when it is removed, whereas LRAC assesses the effect of node removal on the local neighborhood structure. Combining both global and local views, these measures offer a more comprehensive insight into node significance in complex networks.

LITERATURE SURVEY

On node identification in networks, the importance of these nodes is paramount for disease control and marketing applications. The usual centrality measures, like degree, closeness, and betweenness centrality, concentrate either on global or local importance but not both. Newer centrality measures, like Eigenvector Centrality and PageRank, take into account a node's links but still have trouble with time-varying networks. Freeman (1977, 1979) and Bonacich (1987) laid down the foundation for centrality measures, while Page et al. (1999) came up with the concept of PageRank. Later on, the authors introduced in 2019 by Lv et al. and Zhao et al. in 2020 proposed centrality measures such as average shortest path centrality and global importance of nodes (GIN), thus providing a good basis for influential spreaders. Local centrality measures, such as semi-local centrality and local neighbor contribution (LNC), gained more prominence for their effectiveness. Hajarathaiah et al.

(2023) introduced GRAC and LRAC, where global and local views are combined in centrality measurement, which involves changes in centrality when a node is removed. Future research will delve into exploring dynamic networks, context-aware mechanisms, and machine learning in an effort to enhance the identification of key nodes in the changing networks.

EXISTING SYSTEM

Degree centrality, closeness centrality, and betweenness centrality are the traditional centrality measures that have been the backbone of network analysis. They deal with individual facets of node significance:

- Degree Centrality: Quantifies the number of edges connected to a node.
- Closeness Centrality: Quantifies the proximity of a node to every other node in the network.
- Betweenness Centrality: Calculates the frequency with which a node is found on the shortest path between other nodes.

Although these measures are informative, they tend to miss the complete picture of node importance, particularly in complex networks where global and local influences both play vital roles.

DRAWBACKS TO EXISTING SYSTEM

1. Incomplete Picture: Global and local centrality metrics concentrate on a single aspect - either global or local - failing to capture node significance's true complexity.
2. Static Nature: These metrics are usually computed from static snapshots of the network, not reflecting the dynamic nature of network interactions.
3.Computational Complexity: In large-scale networks, global centrality metrics such as betweenness centrality can be computationally costly.
4.Context Ignorance: Conventional metrics fail to take into account how node importance varies based on the application.

PROPOSED SYSTEM METHODOLOGY

The suggested system uses GRAC and LRAC to integrate both local and global views. While LRAC considers the impact of node removal on the local structure of the neighborhood, GRAC measures the relative change in centrality of a node at the global level of the network when the node is removed. The following are the steps in the methodology:

1. Centrality Calculation: Find the GRAC and LRAC for every graph node.
2. Weighted Combination: Combine the assigned weights with the centrality scores.
3. Finding Influential Nodes: Using the weighted centrality scores, choose the nodes that have the greatest influence.
4. Simulation: Use the chosen influential nodes to model the contagion spread using the SIR model.

PROPOSED SYSTEM ARCHITECTURE

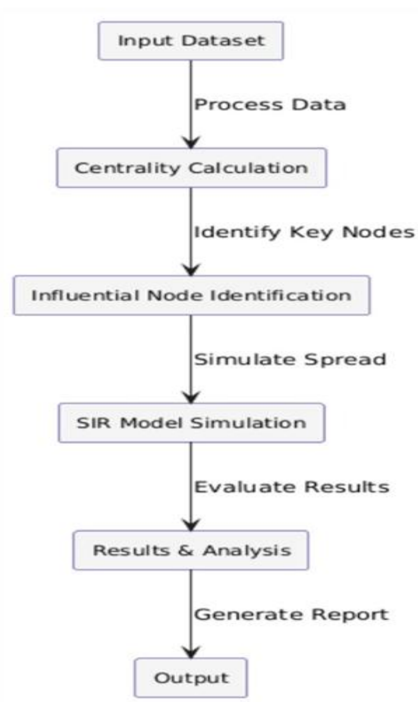

Fig:[1]

PROPOSED ALGORITHMS EXPLANATION

1.Global Relative Average Centrality (GRAC):
GRAC is used to measure the global influence of a node by taking the relative change in the network's average centrality when removing the node. The GRAC formula is represented as:

$$GRAC = \frac{|AV[G'v]-AV[G]|}{AV[G]}$$

- GRAC is Global Relative Average Centrality of vertex v in graph G. $G^{|}$ is the graph removing vertex v.
- AV[$G^{|}$] and AV[G] represents the average centrality measures of the graphs $G^{|}$ and G respectively

2.Local Relative Average Centrality (LRAC):
LRAC is concerned with the local effect of a node by considering the effect of its removal on the structure of the local neighborhood. LRAC is described by the equation:

$$\text{LRAC} = \frac{|AV[Hl/v]-AV\ [Hl]|}{AV\ [Hl]}$$

- LRAC is Local Relative Average Centrality. Hl is a graph from G which contains up to l levels of neighborhood of vertex v.
- Hl/v is a graph from G which contains up to l levels of neighborhood of vertex v except vertex v.

3.Simulation Using the SIR Model:
In order to analyze the efficiency of GRAC and LRAC, we use the Susceptible-Infected-Recovered (SIR) model, which mimics the transmission of contagion in a network. Nodes in this model are classified as Susceptible (S), Infected (I), or Recovered (R). Infected nodes try to spread the infection to their neighbors with probabilities predefined. We choose seed nodes on the basis of their GRAC and LRAC scores and mimic the contagion spread to check their impact on information spread.

BENEFITS OF THE PROPOSED SYSTEM

The proposed system, which centralizes around GRAC and LRAC, relates to some very important aspects of improvement with reference to other forms of centrality.
1.Comprehensive Node Importance Assessment:
GRAC and LRAC have a more complete and multidimensional understanding of node importance, hence, instead of solely covering a holistic or broad perspective taking into account both the global network structure and local neighborhood, it also captures the local neighborhood dynamics to provide an overall view of the node effect.
2.Improved Identification of Influential Nodes:
The system identifies nodes that can more effectively spread contagion or information much more effectively than any of the other centrality measures and hence offers extremely useful applications in disease control, viral marketing, and rumor management in other areas.
3. Robustness Across Different Transmission Rates:
GRAC and LRAC are postulating robustness across different beta values in the SIR model. Overall, these measures are gained across distinct transmission rates as they are always above those traditional centrality measures, making them reliable in targeting sources across scenarios.
4.Active Network Analysis:
The system is extensible to a dynamic analysis where, over time, the structure changes. This is significant because it helps in interpreting the changing dynamics of real networks and the roles of top nodes.

RESULTS

Our findings show that GRAC, LRAC and SIR Model, always perform better than the conventional centrality measures, including degree centrality, closeness centrality, and betweenness centrality, in detecting influential nodes. For different transmission rates (beta values), GRAC, LRAC and SIR Model, have the highest number of infected nodes, reflecting their better capacity to identify nodes that are most important in disease transmission.

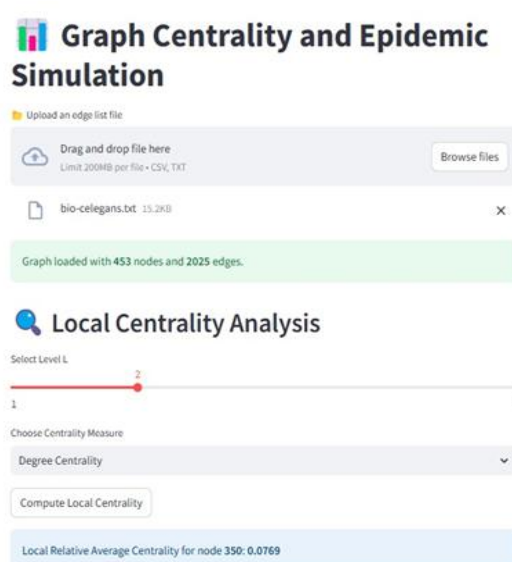

Fig:[2]

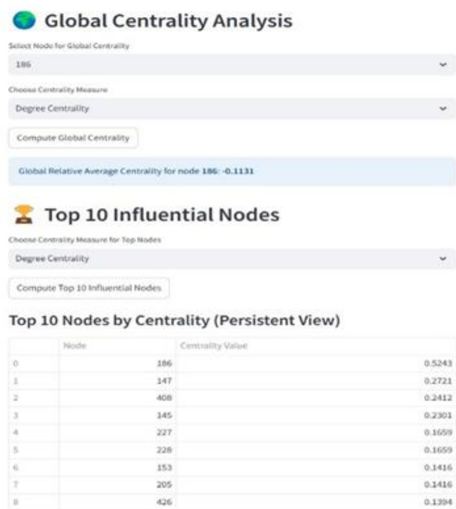

Fig:[3]

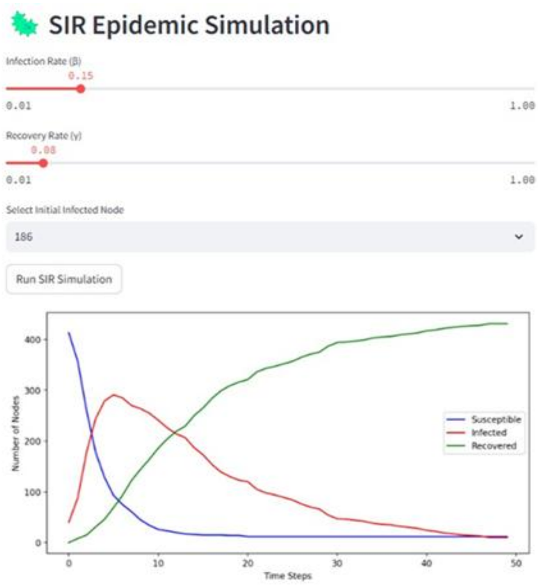

Fig:[4]

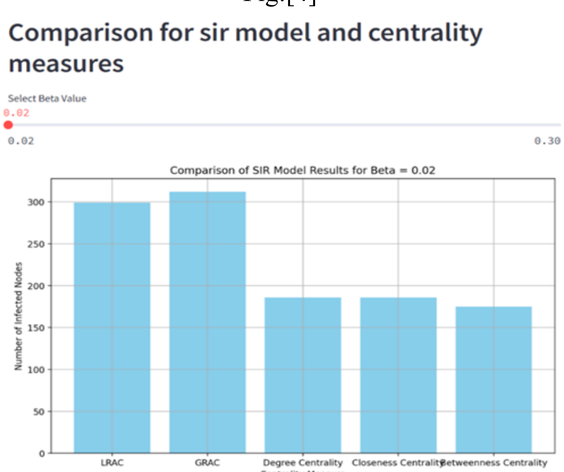

Fig:[5]

CONCLUSION

In this work, we used two centrality metrics, GRAC and LRAC, that incorporate global and local views to detect influential nodes in networked systems. Our findings show that these metrics outperform existing centrality metrics consistently in detecting nodes that are critical to disease spreading and information spreading.

Future research will investigate dynamic network analysis, in which network structure evolves over time, and test our conclusions using real-world networks from a wide range of domains. We also intend to study other epidemic models and tackle the problem of scalability of computing centrality measures for massive networks.

FUTURE SCOPE

Further research might focus on refining GRAC and LRAC to their applications in dynamic networks, such as networks where evolving structures influence the propagation of influences. Using modified epidemic models would make the models far more robust to different transmission dynamics. Distributed computing will address scalability which will thus enable applications at a larger scale such a social media and financial networks. Next step would be automating and adaptive learning of models of influence detection integrated with machine learning, particularly GNNs. In addition, they would further ensure the realization of the applications of GRAC ,LRAC and SIR Model ,which would then render them practical under real-world conditions like cyber threat detection, misinformation controls, and even recommendation systems, in a world becoming more digital than ever.

REFERENCES

[1] Freeman, L. C (1977). A set of measures of centrality based on betweenness. Sociometry, 35(1), 41-56,1977.

[2] Freeman, L. C. (1979). Centrality in social networks conceptual clarification. Social Networks, 1(3), 215-239.

[3] Newman, M. E. (2005). A measure of betweenness centrality based on random walks. Social Networks, 27(1), 39-54.

[4] Bonacich, P. (1987). Power and centrality: A family of measures. American Journal of Sociology, 92(5), 1170-1182.

[5] Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank citation ranking: Bringing order to the web. Stanford InfoLab.

[6] Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London, Series A, 115(772), 700-721

[7] Centola, D. (2010). The spread of behavior in an online social network experiment. *Science, 329*(5996), 1194-1197.

[8] Hajarathaiah, K., Enduri, M. K., Dhuli, S., Anamalamudi, S., & Cenkeramaddi, L. R. (2023). Generalization of Relative Change in a Centrality Measure to Identify Vital Nodes in Complex Networks. *IEEE Access, 11*, 808-824.

[9] Liu, H. L., Ma, C., Xiang, B. B., Tang, M., & Zhang, H. F. (2018). Identifying multiple influential spreaders based on generalized closeness centrality. *Physica A: Statistical Mechanics and Its Applications, 492*, 2237–2248.

[10] Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, 555–564.

[11] Lv, Z., Zhao, N., Xiong, F., & Chen, N. (2019). A novel measure of identifying influential nodes in complex networks. *Physica A: Statistical Mechanics and Its Applications, 523*, 488–497.

[12] Saramäki, J., Kivelä, M., Onnela, J. P., Kaski, K., & Kertész, J. (2007). Generalizations of the clustering coefficient to weighted complex networks. *Physical Review E, 75*(2), 027105.

[13] Hajarathaiah, K., Enduri, M. K., Anamalamudi, S., Reddy, T. S., & Tokala, S. (2022). Computing Influential Nodes Using the Nearest Neighborhood Trust Value and PageRank in Complex Networks. *Entropy, 24*(5), 704.

[14] Zhao, J., Wang, Y., & Deng, Y. (2020). Identifying influential nodes in complex networks from a global perspective. *Chaos, Solitons & Fractals, 133*, 109637.

[15] Chen, D., Lu, L., Shang, M. S., Zhang, Y. C., & Zhou, T. (2012). Identifying influential nodes in complex networks. *Physica A: Statistical Mechanics and Its Applications, 391*(4), 1777–1787.

[16] Berahmand, K., Bouyer, A., & Samadi, N. (2018). A new centrality measure based on the negative and positive effects of clustering coefficient for identifying influential spreaders in complex networks. *Chaos, Solitons & Fractals, 110*, 41–54.

[17] Dai, J. et al. (2019). Identifying Influential Nodes in Complex Networks Based on Local Neighbor Contribution. *IEEE Access, 7*, 131719–131731.

[18] Zeng, A., & Zhang, C. J. (2013). Ranking spreaders by decomposing complex networks. *Physics Letters A, 377*(14), 1031–1035.

[19] Yang, Y., Wang, X., Chen, Y., Hu, M., & Ruan, C. (2020). A Novel Centrality of Influential Nodes Identification in Complex Networks. *IEEE Access, 8*, 58742–58751.

[20] Kumar, S., & Panda, B. S. (2020). Identifying influential nodes in social networks: Neighborhood Coreness based voting approach. *Physica A: Statistical Mechanics and Its Applications, 553*, 124215.