ICAI 2023
12th International Conference on Applied Informatics
Eger, Hungary, 2023

The performance of modern centrality measures on different information models

Péter Marjai[a], **Máté Nagy-Sándor**[b], **Attila Kiss**[c]

[a]Department of Information Systems, ELTE Eötvös Loránd University, Budapest 1117, Hungary
g7tzap@inf.elte.hu

[b]Department of Information Systems, ELTE Eötvös Loránd University, Budapest 1117, Hungary
olj0im@inf.elte.hu

[c]Department of Information Systems, ELTE Eötvös Loránd University, Budapest 1117, Hungary
kiss@inf.elte.hu

Abstract

For the last few years networks became integral parts of our everyday life. They are used in communication, transportation, marketing, and the list goes on. They are also becoming bigger, and more complex and dynamic networks also start to appear more. In light of this, the problem of finding the most influential node in the network remains of high interest however, it is getting more and more difficult to find these nodes. It is hard to grasp the true meaning of what is really being the most influential node means. There are several approaches to define what the most vital nodes are like having the most edges connected to them or having the shortest paths running through them. They can be also identified by calculating the influence of their neighbors, or evaluating how they contribute to the whole of the network. Over recent years various new centrality measures were proposed to order the importance of the nodes of a network.

Local Fuzzy Information Centrality (LFIC) [6] uses a box for every node that contains the node's closest neighbors. The information that can be found in a node's box is used to evaluate the significance of the node. To calculate the uncertainty of the amount of information in the boxes, and to calculate the contribution of a node's neighbors, an improved Shannon entropy is used. A lot of centrality

measures take the whole network into account, but in real life's huge networks, these are not applicable because of their sheer size. Local Clustering H-index Centrality (LCH) [5] only takes the local information into account. While calculating the node's importance, it considers the quality, influence, and topology of first-order and second-order neighbor nodes. Global Structure Model (GSM) [4] not only uses a node's self-influence to rank the nodes but also the node's influence on the whole network. To achieve this, the method utilizes k-shell clusterization.

The information model that is used in a network can also affect the behavior of the nodes. The SIR model [3] starts with a non-empty array of infected nodes. In each turn, the infected nodes try to infect their neighbors with a fixed probability. They also have a fixed probability to recover. Recovered nodes can not be infected again. Independent Cascade model [1] is a stochastic information diffusion model that uses cascading to flow the information through the network. Each node can have two states, active or inactive. In each step, the active nodes have a fixed probability to activate their passive neighbors. An active node can only try to activate its neighbors once. Another information model is the linear threshold model [2]. It also works in iterations. The nodes became active after the ratio of their active and passive exceeded the pre-defined threshold.

In this paper, we evaluate the performance of the modern centrality measures on different information models and compare them with conventional centrality measures. In our experiments, we investigate the similarity between the top-n ranking nodes of the measures, the influential capacity of these nodes as well as the frequency of the nodes with the same centrality value.

References

[1] S. Bikhchandani, D. Hirshleifer, I. Welch: *A theory of fads, fashion, custom, and cultural change as informational cascades*, Journal of political Economy 100.5 (1992), pp. 992–1026.

[2] M. Granovetter, R. Soong: *Threshold models of diffusion and collective behavior*, Journal of Mathematical sociology 9.3 (1983), pp. 165–179.

[3] W. O. Kermack, A. G. McKendrick: *A contribution to the mathematical theory of epidemics*, Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character 115.772 (1927), pp. 700–721.

[4] A. Ullah, B. Wang, J. Sheng, J. Long, N. Khan, Z. Sun: *Identification of nodes influence based on global structure model in complex networks*, Scientific Reports 11.1 (2021), pp. 1–11.

[5] G.-Q. Xu, L. Meng, D.-Q. Tu, P.-L. Yang: *LCH: A local clustering H-index centrality measure for identifying and ranking influential nodes in complex networks*, Chinese Physics B 30.8 (2021), p. 088901.

[6] H. Zhang, S. Zhong, Y. Deng, K. H. Cheong: *LFIC: Identifying influential nodes in complex networks by local fuzzy information centrality*, IEEE Transactions on Fuzzy Systems (2021).