Identification of the most influential nodes involving all topological dimensions of a network

Abbas Salavaty[1,2], Mirana Ramialison[1,2], and Peter D Currie[1,3]
[1]Australian Regenerative Medicine Institute-, [2]Systems Biology Institute Australia-, [3]EMBL Australia-Monash University

ABSTRACT

Biological systems are composed of highly complex networks and several algorithms have been designed to identify the most influential regulatory points within them. However, current methods do not address all the topological dimensions of a network or correct for inherent positional biases. We developed the Integrated Value of Influence (IVI), which integrates the most important network centrality measures in an unbiased way and captures all of the topological dimensions of a network to identify the most influential nodes. IVI is accessible on the Influential Software Package web portal at: https://influential.erc.monash.edu/IVI/

CONTACT

Adrian (Abbas) Salavaty
ARMI, Monash University

Personal Website

@mania_abbas

IVI Shiny App

INTRODUCTION

Network science has been used for the study of complex networks in almost all domains (1, 2). In all network analyses the topology of a network is analyzed using different centrality algorithms to identify the most influential nodes and get a deeper insight from the network (3). In this context, two major concepts are defined, hubness and spreading potential. While hub nodes have high connections with other nodes, the spreader nodes have the greatest impact on the flow of information throughout the network (4, 5). In a study published in Patterns, Cell Press, we for the first time demonstrated that nodes with a simultaneously large number of connections and high spreading potential are the most influential and vital nodes in a network (5). IVI is the synergistic product of integration of six network centrality measures.

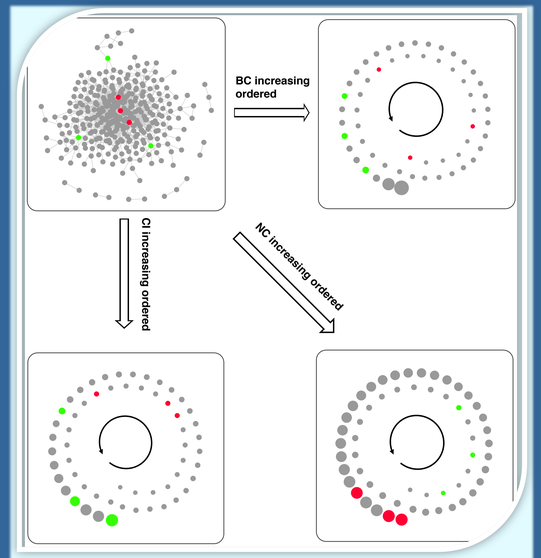

Fig. 1. The positional bias of betweenness centrality and collective influence. This figure has been adapted from (5). BC and CI are positionally biased in the network and this bias is in contrast to the behavior of NC.

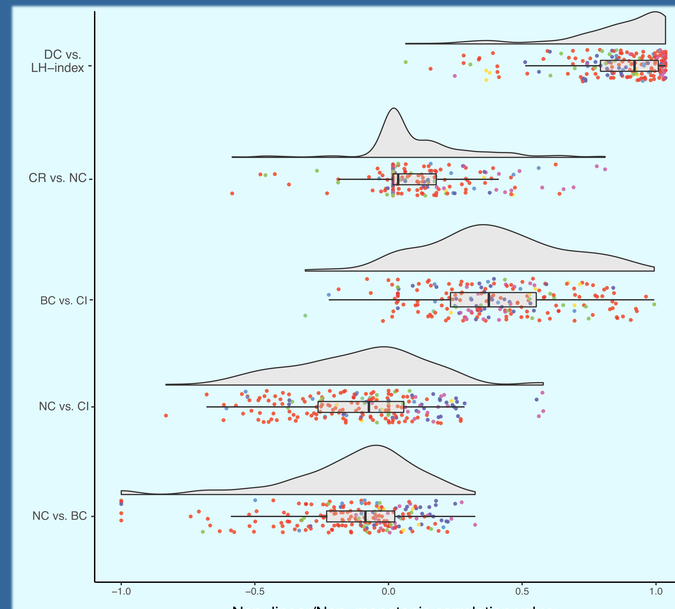

Fig. 2. Correlation between different centrality Measures across 200 networks. This figure has been adapted from (5).

- The association and dependence of six centrality measures including degree centrality (DC), local H-index (LH-index), neighborhood connectivity (NC), ClusterRank (CR), betweenness centrality (BC), and collective influence (CI) were thoroughly assessed.
- BC and CI are positively and negatively correlated with each other and with NC across 200 networks, respectively. Also, CR was the only centrality measure that was positively correlated with NC in most of the networks. Moreover, DC and LH-index had strong and consistent positive correlation with each Other.

IVI algorithm

- The centrality measures were integrated using the *Addition* and *Multiplication* functions to synergize their effects and produce the IVI.
- The *Spreading* and *Hubness* scores emerged as the product of four and two different centrality measures, respectively, as follows.

$$Spreading_{score_i} = (NC'_i + CR'_i)(BC'_i + CI'_i)$$

$$Hubness_{score_i} = DC'_i + LH'_{index_i}$$

- The integration of *Spreading* and *Hubness* scores resulted in the IVI, which synergizes the effects of most important centrality indices

$$IVI_i = (Hubness_{score_i})(Spreading_{score_i})$$

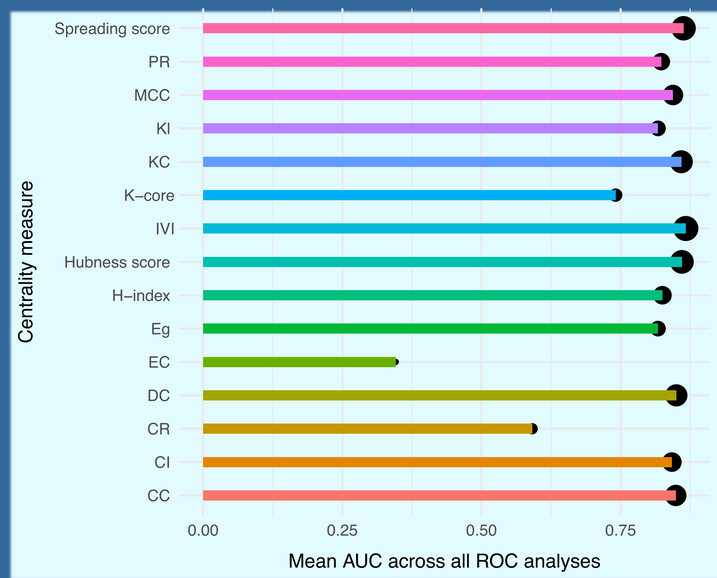

Fig. 3. Performance of IVI in comparison with 14 other current methods in identification of the most influential nodes. This figure has been adapted from (5).

- IVI outperformed 12 other contemporary methods in detecting influential network nodes. Also, *Spreading* and *Hubness* scores, the components of the IVI formula, came up as the second and third ranked metrics, respectively.
- AUC: Area Under the Curve
- corresponding to receiver operating characteristic (ROC) analyses.

CONCLUSIONS

- IVI algorithm integrates six of the most important centrality measures in such a way that their strengths are synergized and positional biases are removed
- IVI algorithm outperforms other algorithms in identifying the most influential nodes
- IVI could be a beneficial tool for all future network analyses including systems biology studies

REFERENCES

1. Balaban AT. Applications of graph theory in chemistry. J Chem Inf Comput Sci. 1985;25(3):334-343.
2. Frainay C, Jourdan F. Computational methods to identify metabolic sub-networks based on metabolomic profiles. Brief Bioinform. 2017;18(1):43-56.
3. Tieri P, Farina L, Petti M, Astolfi L, Paci P, Castiglione F. Network Inference and Reconstruction in Bioinformatics. Elsevier; 2019. p. 805-813.
4. Kitsak M, Gallos LK, Havlin S, Liljeros F, Muchnik L, Stanley HE, et al. Identification of influential spreaders in complex networks. Nature Phys. 2010;6(11):888-93.
5. Salavaty A, Ramialison M, Currie PD. Integrated Value of Influence: An Integrative Method for the Identification of the Most Influential Nodes within Networks. Patterns. 2020;1(5):100052.