STARS, NEIGHBORHOOD INCLUSION, AND NETWORK CENTRALITY

David Schoch, Ulrik Brandes

SIAM Workshop on Network Science 2015
May 16-17 · Snowbird

Summary

The numerous network centrality indices proposed in the literature have little in common but the operational quantification of an intuition that nodes in better positions are more central, where "position" is relative to a particular conceptualization. We propose to discriminate centrality indices from other indices by focusing on a minimal requirement for "better" instead. Formally, we posit that a centrality ranking must preserve the neighborhood-inclusion preorder, and discuss advantages of such an approach.

Additional details

For ease of exposition, we consider only networks represented as simple undirected graphs. Vertex centrality is commonly defined via mappings $c : V \to \mathbb{R}_{\geq 0}$ assigning a non-negative number to every vertex such that

$$c(u) > c(v) \iff u \text{ considered more central than } v\ .$$

Thus, any index c induces a ranking of the vertices, but not every such ranking represents a plausible concept of structural importance. Several attempts have been made to delineate and break down the space of centrality indices.

One such line of research was initiated by Sabidussi [8] and has focused on axiomatic characterization. Ideally, a combination of (intuitively plausible) axioms describes the behavior of centrality indices to an extent that facilitates interpretative statements about centrality rankings and aids in the selection of indices [3]. On the other hand, axiomatic approaches are typically restrictive, often leaving only few possibilities to satisfy a combination of axioms. This impedes general theorems about centrality indices and sometimes only shifts the focus from the definition of indices to the definition of axioms.

A second line of research is more conceptually oriented [4, 2] providing terminology and intuition to reason about the features embodied in centrality indices. It thus relates formal definitions with substantive motivations, but does not allow for sharp distinctions and provable statements.

It appears that the only requirement that is both formally established and substantively accepted is the *star property*. In the words of Freeman [4],

> *"A person located in the center of a star is universally assumed to be structurally more central than any other person in any other position in any other network of similar size."*

As we will argue below, however, it is not strong enough to ensure plausibility and interpretability by itself.

We therefore introduce a new approach to characterize centrality concepts. It is positioned between the weak star property and restrictive axiomatizations, but deviates from both in several respects. Most importantly, it weakens the scale of measurement from quantitative to ordinal, and thus allows for the progressive tightening of feasible rankings rather than intersecting regions in the space of indices. The core element of our approach is a well-studied preorder on the vertices of a graph.

Definition 1. *Let $G = (V, E)$ be a simple undirected graph and $u, v \in V$. Denote by $N(u)$ the neighborhood of u and by $N[u] = N(u) \cup \{u\}$ the closed neighborhood of u. Then, u* dominates *v, $u \succcurlyeq v$, if $N[u] \supseteq N(v)$.*

Domination defines a preorder $\succcurlyeq$ (i.e., a partial ranking) on the set of vertices, often referred to as the neighborhood-inclusion or vicinal preorder. If the shared meaning of all centrality concepts is, that a better position is characterized by having better relations, a vertex dominating another should (a) be comparable to the other and (b) never be less central than it. In the extended version of this paper we show that these requirements are indeed satisfied for all common centrality indices. This motivates the following formulation.

Proposition 2. *For a simple undirected graph $G = (V, E)$, an index $c : V \to \mathbb{R}_{\geq 0}$ is a centrality, if and only if it respects the neighborhood-inclusion preorder $\succcurlyeq$, i.e.,*

$$u \succcurlyeq v \implies c(u) \geq c(v)$$

for all $u, v \in V$.

In the remainder, we provide three arguments to make the case for our criterion.

Star property. An immediate consequence of our requirement is that the rankings of all centrality indices coincide on networks for which the neighborhood-inclusion preorder is complete. This holds, in particular, for stars so that the star property is maintained as a necessary requirement. It is strengthened, however, to the much larger class of threshold graphs which are completely ranked as well [6].

To illustrate the strengthening of the star property, consider a crafted index that we refer to as hyperbolic centrality,

$$c_{hyp}(u) = \mathbb{D}(G[N[u]]) \cdot \left[\sum_{v \in N[u]} \sum_{k=0}^{\infty} \frac{A(G[N[u]])^{2k}_{vv}}{(2k)!} \right] ,$$

where $G[N[u]]$ is the subgraph induced by $N[u]$, $A(G[N[u]])$ is its adjacency matrix, and $\mathbb{D}(G[N[u]])$ its density (also known as the clustering coefficient of u). The index thus corresponds to a density-weighted sum over all length-scaled closed walks of even length within the closed neighborhood of a vertex. It thus combines ideas from other centrality indices such as total communicability [9], and it satisfies the star property. We even find that, experimentally, it compares well with centrality indices previously applied to a biological prediction task. Given that the walks considered here may start anywhere in the neighborhood it is difficult to argue, though, that it matches any intuitive notion of centrality and, indeed, it fails to preserve the proposed dominance criterion.

Axiomatization. The goal of centrality axiomatization is comparative in general: which properties characterize a particular index and distinguish it from others? Sabidussi's [8] seminal work appears to be the first along these lines, and many others have followed (cf. the review sections in [5, 3]).

In line with Vigna and Boldi [3, p. 11], we avoid prescribing a set of axioms, but only postulate a necessary requirement. Moreover, the requirement in Proposition 2 disregards the actual values being assigned and assumes only an ordinal scale of measurement. It also does away with the analytically inconvenient differentials after graph transformation such as edge addition or switching.

Empirical research. Centrality is commonly used as explanatory, independent, as well as intermediate variable in empirical research. Research hypotheses typically state that the level of some variable (say, trust placed in an organization) is either positively or negatively associated with some centrality index (say, in a business network the organization is part of). The selection of a centrality index is usually the weakest part of a research design, as little reliable knowledge exists that places one index over another. Moreover, if the association cannot be confirmed empirically, one is at loss.

With a strengthened necessary condition, it is possible to make more powerful general statements about *all* centrality indices. If, for example, it turns out the ordering of levels in the other variable is already inconsistent with the partial ranking obtained from the dominance relation, we can conclude that there is *no* centrality that supports the hypothesis, no matter which ones we try.

Moreover, the shared partial ranking of centrality indices can serve to explain frequently observed correlations of centrality indices. The closer a graph to a threshold graph, i.e., a graph with a unique centrality ranking, the fewer degrees of freedom exist for indices. If multiple centrality indices operationalizing different ideas of what makes a vertex central yield similar rankings, they represent competing explanations. Distance from a threshold graph [1, 7] or sparseness of the neighborhood-inclusion preorder thus provide a level confidence in an explanation offered by a particular index.

References

[1] Srinivasa R. Arikati and Uri N. Peled. Degree sequences and majorization. *Linear algebra and its applications*, 199:179–211, 1994.

[2] Stephen P. Borgatti and Martin G. Everett. Graph-Theoretic Framework for Classifying Centrality Measures. *Social Networks*, 28(4):466–484, 2006.

[3] Paolo Boldi and Sebastiano Vigna. Axioms for Centrality. *Internet Mathematics*, 10(3-4):222–262, 2014.

[4] Linton C. Freeman. Centrality in social networks I: Conceptual clarification. *Social Networks*, 1(3):215–239, 1979.

[5] Dirk Koschützki, Katharina A. Lehmann, Dagmar Tenfeld-Podehl, and Oliver Zlotkowski. Advanced centrality concepts. In: *Network Analysis*, LNCS Tutorial 3418. Elsevier, 1995.

[6] Nadimpalli V.R. Mahadev and Uri N. Peled. *Threshold graphs and related topics*, Annals of Discrete Mathematics 56. Elsevier, 1995.

[7] Assaf Natanzon, Ron Shamir, and Roded Sharan. Complexity classification of some edge modification problems. *Discrete Applied Mathematics*, 113(1):109–128, 2001.

[8] Gert Sabidussi. The centrality index of a graph. *Psychometrika*, 31(4):581–603, 1966.

[9] Michele Benzi and Christine Klymko. Total communicability as a centrality measure. *Journal of Complex Networks*, 1(2):124–149, 2013.