Social Networks 54 (2018) 50–60

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

Centrality without indices: Partial rankings and rank probabilities in networks☆

David Schoch

ETH Zürich Department of Humanities, Social and Political Sciences Chair of Social Networks Weinbergstraße 109 8006 Zürich, Switzerland

ARTICLE INFO

Article history:

Keywords:
Network centrality
Neighborhood-inclusion
Centrality indices
Rank probabilities
Partial orders

ABSTRACT

We present an alternative approach to assess centrality in networks which does not rely on traditional indices. The work is based on neighborhood-inclusion, a partial ranking inducing relation of nodes, which was shown to be preserved by many existing centrality indices. As such, it can serve as the shared basis for centrality in networks. We argue that evaluating this partial ranking by itself allows for a generic assessment of centrality, avoiding several pitfalls that can arise when indices are applied. Additionally, we illustrate how to derive further partial rankings and introduce some probabilistic methods to, among others, compute expected centrality ranks of nodes.

1. Introduction

Network centrality is commonly defined in terms of indices, assigning numerical scores to the nodes of a network. These scores induce a ranking which is meant to reflect the structural importance of entities comprising a network. There is, however, little agreement on what constitutes "structural importance" and many indices exist, assessing it on various levels (Lü et al., 2016). Much effort was put into conceptual clarifications (Freeman, 1979), axiomatizations (Sabidussi, 1966; Nieminen, 1974; Ruhnau, 2000; Landherr et al., 2010; Kitti, 2012; Boldi and Vigna, 2014) and classifications (Borgatti, 2005; Borgatti and Everett, 2006), yet formal guidelines to restrict the set of possible indices or alternative methods are scarce.

Recently, Schoch and Brandes (2016) introduced a simple characterization of centrality: if an actor has the same (and possibly more) ties than another actor, it can never be less central. The authors showed that there exists a partial ranking of nodes based on neighborhood-inclusion which is preserved by many known centrality indices. This finding can be used to characterize possible centrality rankings (hence indices) as the set of rankings extending the partial ranking given by neighborhood-inclusion.

Preorders as given above are a simple and well studied mathematical structure (Davey and Priestley, 2002; Grätzer, 2002), emerging in many different fields where objects need to be put in order. A prime example provides the field of *multi-criteria decision making* (MCDM) (Belton and Stewart, 2002; Triantaphyllou, 2013). Given a set of alternatives valued by a set of criteria, the objective is to find the "best" alternative. This is commonly done using indices aggregating the criteria of each alternative into a *preference ranking*. The preference rankings are supposed to preserve an intrinsic dominance order, that is if an object is better for each criterion than another, it should always be preferred to the dominated one. In a sense, the idea of these indices is conceptually related to centrality indices.

While indices are well established to build preference rankings, they are perpetually under scrutiny for various fundamental issues. Stewart (1992) summarizes that there exists "a plethora of approaches", where "some of these are *ad hoc*, and largely unjustified on theoretical and/or empirical grounds". Patil and Taillie (2004) argue that indices are typically "adopted on grounds of mathematical convenience or simplicity" without proper justifications. Triantaphyllou and Mann (1989) uncovered the infamous *decision making paradox* which states that different index based decision methods can yield different results when fed with the same data. Put in other terms, the choice of index affects the choice of the most preferred alternative. The paradox is also related to the problem of *rank reversals*, i.e. the most preferred alternative changes when an actually inferior alternative is added. This issue was first described by Belton and Gear (1983).

The outlined issues in MCDM are fundamentally linked to network centrality. We too are faced with a plethora of indices that are, according to Freeman (1979), "often unnecessarily complicated", "absolutely unintelligible from any theoretical perspective whatever", or "tend to add unnecessary and confusing complica-

☆ Part of this research was done while affiliated with the University of Konstanz and with the financial support from Deutsche Forschungsgemeinschaft (DFG) under grant Br 2158/6-1, which I gratefully acknowledge. I also want to thank the anonymous reviewers whose comments/suggestions helped improve and clarify this manuscript.

E-mail address: david.schoch@gess.ethz.ch

https://doi.org/10.1016/j.socnet.2017.12.003
0378-8733/

tions that make them difficult to interpret." Moreover, choosing an appropriate index in an empirical setting becomes a daunting task when faced with the overabundance of indices. This may foster trial and error approaches, probing different indices until a satisfactory result is obtained.

Many alternative methods exist to analyse and determine (partial) rankings which do not necessarily rely on indices. In this work, we focus on methods that build upon an intrinsic ordering of some pairs of objects. These methods range from simple analyses of their structure (Patil and Taillie, 2004; Pavan and Todeschini, 2004), how to handle incomparability (Brüggemann and Carlsen, 2014; Bartel and Mucha, 2014), constructing non-numerical rankings (Fishburn and Gehrlein, 1975; Janicki, 2008, 2009) to computing all possible rankings that preserve the intrinsic ordering (Bubley and Dyer, 1999; Habib et al., 2001). The latter are mostly of interest to derive (relative) rank probabilities and expected ranks (Brüggemann et al., 2003; Brüggemann et al., 2004; Brüggemann et al., 2005; De Loof et al., 2006, 2008; De Loof, 2009).

In the upcoming parts, we illustrate how these existing methods can be used in the context of network centrality, offering an alternative or complementary approach to centrality indices. We start in Section 2 by introducing the intrinsic ordering of nodes based on neighborhood-inclusion. We illustrate how to decompose indices into three building blocks which facilitate further theoretical considerations. Lastly, we show how the neighborhood-inclusion preorder can be extended to denser partial rankings. The remainder of the work builds on the theoretical results from this section. In Section 3, we offer first alternatives to indices by introducing Hasse diagrams and assessing rank ambiguities with rank intervals. Subsequently, we introduce probabilistic methods for centrality in Section 4. These include relative rank probabilities (How likely is it, that a node is more central than another?) and expected ranks of nodes (How central is a node expected to be considering all possible rankings?). In Section 5, we illustrate how the presented methods can be employed in empirical research. Finally, we end with some concluding remarks and a discussion in Section 6.

2. Centrality indices and partial rankings

Centrality is commonly defined in terms of mappings $c : V \rightarrow \mathbb{R}_{\geq 0}$ assigning real valued scores to the set of nodes V of a graph $G=(V, E)$ and are interpreted as

$$c(u) > c(v) \Leftrightarrow u \text{ is more central than } v.$$

The most widely used indices are degree, betweenness (Freeman, 1977) which was earlier introduced as rush by Anthonisse (1971), closeness (Bavelas, 1948; Sabidussi, 1966), and eigenvector centrality which was introduced for the first time by Wei (1952), later generalized by Berge (1958) and reintroduced by Bonacich (1972). Recently, Schoch and Brandes (2016) showed that there exists a shared basis among a large group of indices, including these prototypical ones, by a comparison of neighborhoods in graphs. The idea is based on a decomposition of indices into the following three generic steps:

(1) Deriving indirect relations via path algebras.
(2) Defining vertex positions via coordinates evaluating indirect relations.
(3) Defining centrality scores by aggregating values of positions.

In this section, we formally introduce this framework and illustrate how each of these steps leads to the definition of a partial ranking on networks. Additionally, we illustrate how this framework can be used to derive additional partial rankings, which are preserved by specific classes of indices.

2.1. Indirect relations

Indirect relations on graphs can be characterized with so called *path algebras* based on the algebraic structure of semirings.[1]

Definition 1. An algebraic structure $(S, \oplus, \odot, \bar{0}, \bar{1})$ on a set of values S is called a *semiring* if and only if

(i) $\oplus, \odot : S \times S \rightarrow S$ are closed and associative
(ii) $\bar{0}, \bar{1} \in S$ are neutral elements of $\oplus$ and $\odot$, respectively
(iii) $\bar{0}$ is absorbing for $\odot$
(iv) $\oplus$ is commutative and $\odot$ distributes over $\oplus$

Path algebras characterize indirect relations on graphs $G=(V, E)$ for vertices $s, t \in V$ by associating a value from a semiring with every (s, t)-path and then aggregating over all such paths. We associate each edge with an element $\bar{e} \in S$, the *edge value*, so that we can characterize a graph $G=(V, E)$ by a matrix $A \in S^{V\times V}$ with entries

$$a_{st} = \begin{cases} \bar{e} & \{s, t\} \in E \\ \bar{0} & \text{otherwise} \end{cases}$$

for all $s, t \in V$. An (s, t)-path P with vertex sequence $s = v_0, v_1, \ldots, v_{k-1}, v_k = t$ is then evaluated by

$$a(P) = \odot_{i=1}^{k} a_{v_{i-1}, v_i}, \tag{1}$$

where $a(P) = \bar{1}$ if $k=0$. The indirect relation between s and t is then obtained from

$$a^*_{st} = \oplus_P a(P), \tag{2}$$

where all possible (s, t)-paths are aggregated. If no such path exists, $a^*_{st} = \bar{0}$ holds. The term a^*_{st} is also known as the *closure* of a_{st}. A convenient joint formulation can be derived in terms of matrices. Let $A \in S^{V\times V}$ be a matrix as defined above, and let $\mathbf{1} \in S^{V\times V}$ be the matrix with entries $\bar{1}$ on the diagonal and $\bar{0}$ elsewhere. This formulation gives rise to a semiring of matrices $(S^{V\times V}, \oplus, \odot, \mathbf{0}, \mathbf{1})$, where $\oplus$ and $\odot$ replace the usual matrix addition and multiplication. By defining $A^0=\mathbf{1}$ and $A^k = A \odot A^{k-1}$ for $k \geq 1$, we obtain the closure $A^* = \oplus_{k=0}^{\infty} A^k$ with entries a^*_{st} as above.

For later purposes, we need two additional properties of semirings. The first defines an order relation on the set S.

Definition 2. The canonical preorder of a semiring $(S, \oplus, \odot, \bar{0}, \bar{1})$ is given by

$$a \leq b \text{ if } a \oplus c = b \text{ for some } c \in S.$$

Definition 3. Let $(S, \oplus, \odot, \bar{0}, \bar{1})$ be a semiring and $\bar{e} \in S$ be an edge value. The semiring is called *decreasing* if

$$\bar{e} \odot a \leq a$$

holds for all $a \in S$.

As an example, consider the geodesic semiring $(\mathbb{N}_0 \cup \infty, \min, +, \infty, 0)$ with edge value $\bar{e} = 1$. By *concatenation* with Eq. (1) and *aggregation* via Eq. (2), we obtain entries $a^*_{st} = \text{dist}(s, t)$ and thus the distance matrix as the closure A^* of the corresponding matrix semiring. The geodesic semiring is decreasing, since adding an edge to a path increases the distance. It should be noted though

[1] A comprehensive introduction of semirings and path algebras is given by Gondran and Minoux (2008).