Network positions

Ulrik Brandes

Abstract
A formal definition of position in networks is proposed. Based on the conceptualization of networks as data on intersecting dyads, it leads to integrated means of analysis for various kinds of relational and attribute data and to a simultaneous unification and generalization of a wide range of network-analytic methods. In fact, by defining network positions explicitly, various methods can be broken down into sequences of generic steps that are more amenable to substantive theorizing, integration of qualitative insight, and empirical validation.

Keywords
social networks, network science, social position, centrality

Introduction
Network science is an emerging discipline experiencing rapid growth (Brandes et al., 2013). This growth is driven by the development of new models and methods on the one hand, and their application to ever more data from an expanding range of disciplines on the other. Alas, despite noteworthy exceptions (Borgatti et al., 2014; Borgatti and Foster, 2003; Emirbayer and Goodwin, 1994; Erikson, 2013), reflections on theory and methodology are less common.

Social network analysis in particular (Borgatti et al., 2013; Degeene and Forsé, 1999; Hennig et al., 2012; Scott, 2013; Wasserman and Faust, 1994) is often attributed to be all about methods. While this view may be dismissed as ignorant, the traditional name of the area does not help, and it is certainly fair to say that methods are often selected and applied with insufficient justification. Sadly, this is not a recent development (e.g. Granovetter, 1979; Salancik, 1995).

Potential explanations point to the perception of network analysis as a peculiarity, a special method (Granovetter, 1990), rather than a normal science (Hummon and Carley, 1993). The graphical appeal of network images (Mayer, 2012) and the availability of special-purpose software (Huisman and Van Duijn, 2011) seem to reinforce the impression of an approach very different from mainstream research among beginners and those viewing the field from a safe distance.

It is indeed problematic, however, that few methods are derived from first principles and that there appears to be a general dearth of overarching concepts. Consider as arbitrarily selected examples the detailed reviews of methods for community detection in Schaeffer (2007) and Fortunato (2010). They are incredibly useful as references, but their organization speaks volumes of the rather organic growth of the area.

My cavalier assessment of the state of network science and its application to social structure is not meant to be developed, at least for now, into any kind of deep analysis. It explains, however, the premise under which the present contribution is being made.

As an attempt toward reinforcing the foundations of network science, a formal concept of network position is proposed. It generalizes previous formal and informal uses of the term and leads to restructuring of methods. Network positions bear the potential to form an overarching concept that facilitates the organization and context-specific development of network-analytic procedures and identifies the loci of theory.

The concept of position in networks is all but new, although uses of the term are varied. Despite differences in the details and consequences, they appear to share a common idea. A position characterizes what the network looks like
from the point of view of an actor, and multiple actors sharing traits and perceiving comparable environments are considered to be subject to similar opportunities and constraints, resulting in similar actions and the assumption of similar roles. This idea is made explicit, for instance, in von Hayek (1952) and Nadel (1957), and it is the common ground of formalizations (Borgatti and Everett, 1992; Breiger and Pattison, 1986; Burt, 1976; Doreian et al., 2005; Faust, 1988; Lorrain and White, 1971).

My main aim is to reconcile formal concepts of structural position with the companion notion of position in social space (Blau, 1977a; McPherson, 1983), and to demonstrate that this generalization does not come at the cost of precision or practicality, but may in fact provide a guiding frame of reference that facilitates the increase of both sophistication and rigor.

We begin by outlining in Section “Network Data and Indirect Relations” the understanding of network science as a mathematical science delineated by the format of the data that is dealt with. This allows us to define a generic concept of position in Section “Network Positions and Positional Dominance.” The implications of refocusing network studies on the definition and evaluation of network positions are exemplified for the case of centrality analysis in Section “Centrality” and discussed more generally in Section “Discussion.”

Network data and indirect relations

The concept of position put forward in this contribution is meant to be completely general. It applies to asymmetric, valued, temporal, two-mode, ego-centric, and other kinds of networks alike. We therefore start by defining, or redefining, basic terminology and notation so that it extends more straightforwardly. This approach reiterates that we delineate network science by the structure of the data being considered rather than, say, an application domain or a grand network theory. Moreover, it suggest to address the major obstacle of method particularization by factoring out temporal aspects because they generally involve more complex, non-scalar data such as time intervals. That said, the following definition unifies the most common types of networks considered in the literature.

Definition 1 (Network). A network is a mapping \( x : S \to W \) assigning values in a range \( W \) to dyads from a finite domain \( S \subseteq N \times A \) comprised of ordered pairs of nodes \( N \) and affiliations \( A \).

If \( N = A \), then \( S \) is called an interaction domain and \( x \) a one-mode network. If \( N \cap A = \emptyset \), then \( S \) is called an affiliation domain and \( x \) a two-mode network.

Note that we consider a dyad to be an ordered pair and thus capture possible asymmetry in relations such as subordination. Networks may be represented equivalently as vectors \( x \in W^S \) with a set of two-dimensional indices \( S \), and we will use these representations interchangeably.

In the simplest case, the range is dichotomous with values \( W = \{0,1\} \) representing the presence or absence of a relationship. Values may, however, be larger sets of numerals, intervals, or any other kind of data objects.

For these values to facilitate distinctions and comparison, we assume the existence of a preorder \( \preceq \) on \( W \). A preorder is a partial ranking, that is, a binary relation that is reflexive and transitive but need not be complete or anti-symmetric. This is not a strong assumption as we can, for example, default to the diagonal preorder \( \text{diag}(W) = \{(i,i) : i \in W\} \) for nominal variables, and whatever the association between values and preferability otherwise, it just needs to be consistent (i.e. transitive).

For convenience we do assume, though, that the preorder on the values is either bounded from below by a designated element \( 0 \in W \) with \( 0 \preceq a \) for all \( a \in W \) or from above by...
a designated element \( \alpha \in \mathcal{W} \) with \( a \leq \alpha \) for all \( a \in \mathcal{W} \). The special value will be interpreted as the absence or impossibility of the quality or quantity expressed in a dyadic observation. The two situations are distinguished by the interpretation that larger values represent more beneficial relationships (as in, for example, frequency of contact) or smaller values represent less costly relationships (as in, for example, physical distance). Unless explicitly stated otherwise, we will assume the benefit perspective.

The main distinction of different classes of networks is in the composition of nodes and affiliations. For instance, egocentric networks are two-mode networks in which \( N \) is the set of all alters and \( \mathcal{A} \) is the set of all alters but \( \mathcal{S} \) is such that each alter is associated with exactly one ego. Personal networks, on the other hand, are preferably represented as one-mode networks of alters in which no dyad relates alters of different egos. In this case, ego–alter relationships are conveniently included as nodal attributes since the corresponding ego is implied.

Finally, we would like to point out that, by symmetry of the incidence structure of a dyadic domain, nodes \( i \in N \) need not be first-class elements but can be interpreted as the aggregate of incident dyads.

**Running examples**

Two tutorial network data sets will be used for illustration in later sections. We introduce them here to reinforce our delineation of network data.

The first data set has been selected in part because it is a boundary case that illustrates the discrimination of general data tables from those that usefully admit a network perspective. The second is a highly representative type of social network data set that has been analyzed numerous times.

**Example 1 (La Balance des Peintres).** In 1708, the French art critic Roger de Piles published a textbook on painting that concludes with a table in which 56 painters are assigned values ranging from 0 to 20. It is reproduced in Figure 1.

The values encode de Piles’ subjective ratings of each painter’s ability with regard to the four dimensions deemed essential throughout the book, namely composition, drawing, color, and expression.

It is important to note that the rating scale is introduced only once and therefore applies jointly to all four dimensions, with 20 representing perfection (and thus considered unachievable) and 19 representing the highest level conceivable (though it is not attained). Moreover, we may safely assume that 0 represents the absence of any skill, again independent of the dimension.

De Piles refers to the table as “La Balance des Peintres” but it is not certain whether an overall ranking was meant to be read into it.

We may view this data table as a two-mode network \( x : N \times A \rightarrow \{0,...,20\} \) in which painters make up the nodes in \( N \) and skill dimensions make up the affiliations in \( A \). This perspective is viable because de Piles’ introduction to the table and the single definition of a common scale imply that each entry, no matter in which column, should be interpreted as an expression of a joint variable \( x \) corresponding to a “degree of perfection.” We therefore insinuate that equal ratings express equal levels across dimensions, at least by de Piles’ subjective standards, but that the scales are ordinal as evidenced by the leap from 19 and 20.

The rating network will serve as an example for two-mode networks which are the focus, for instance, in the study of interlocking directorates or belief systems. It is also an example of a cases-by-variables matrix that can be viewed as a two-mode network because the entries are comparable across columns (see, for example, Breiger et al., 2014).

The other example is a one-mode network with an additional node attribute.

**Example 2 (Medical Innovations).** Coleman et al. (1966) collected data on private practice physicians in four cities in Illinois, USA, to study whether network ties had an influence on the diffusion of a medical innovation. Although the study provides more extensive data on the involvement of physicians in the medical community, we restrict ourselves to the two key variables, contact networks and the adoption of the then newly introduced antibiotic tetracycline.

The relations for which data were collected are advice seeking, work-related discussion, and friendship.

For 125 of the 246 physicians, a nodal attribute represents the first confirmed subscription of tetracycline in the period from November 1953 to February 1955, chunked in monthly intervals with one exception in the final three months. Values from 1 to 17 indicate that prescription sampling turned up a first prescription in the corresponding month, whereas a value of 18 is assigned to those 16 physicians for whom no prescription was found. They are assumed to have adopted the drug either later or not at all.

The data set was edited and published in the course of a follow-up study of Burt (1987), and it is depicted in Figure 2.

Formally, the amalgamated nominations for advice, discussion, and friendship give rise to an asymmetric one-mode contact network \( x : S \rightarrow \{0,1\} \), where \( S = (N \times N) \setminus \text{diag}(N) \) and \( s \rightarrow t \) indicate that the physician represented by node \( \sigma \) nominated the one represented by \( t \) in at least one of the three questionnaire items.

The single node attribute we will use is the month \( z : N \rightarrow \{1,...,18,\perp\} \) of the first confirmed prescription, where \( z(i) = \perp \) indicates that, according to pharmacy records sampled in the study, physician \( i \) prescribed tetracycline no
Methodological Innovations

The value \( \bot \) is assigned to all physicians without prescription records (white nodes in Figure 2) and incomparable to other values.

In Section “Centrality,” the two example network data sets will be used to exemplify two novel kinds of analysis facilitated by the positional approach.

Figure 1. Painter ratings from La Balance des Peintres (de Piles, 1708). Names adapted to current use in the English-language Wikipedia and dates added, some of which are approximate. Two ratings are blank in the original table, Annibale Carracci is actually rated together with Agostino and Lodovico Carracci, and Paolo Veronese’s double appearance (different names but same ratings) has been consolidated. Later editions, including the English translation published posthumously in 1743, deviate in some of the ratings. Does the summation of ratings (i.e. degree centrality) yield an intended quality ranking?

Later than \( J \). The value \( \bot \) is assigned to all physicians without prescription records (white nodes in Figure 2) and incomparable to other values.
Representations

We have defined networks as data on a sample of dyads that are related by incidence and thus emphasized that a network is not one but many observations.

There is hence a fine line between network data and some of the representations that are employed to make the application of methods from various areas of mathematics more convenient but can be suggestive in how to interpret and analyze a network.

For instance, although networks are often defined as graphs, we think that the above definition is more expedient because it emphasizes that a network is not a single object and that each dyad carries a data point, even if only a zero value. Notwithstanding, a network can of course be represented in a graph.

Definition 2 (Graph of a Network). For a network $x \in \mathcal{W}^S$ with $S \subseteq \mathcal{N} \times \mathcal{A}$, its (weighted directed) graph $G(x) = (V, E; \omega)$ with edge weights $\omega : E \rightarrow \mathcal{W} \setminus \{0, \infty\}$ is defined by

$$V = \mathcal{N} \cup \mathcal{A}$$

$$E = \{ (s, t) \in S : x_{st} \notin \{0, \infty\} \}$$

$$\omega(s,t) = x_{st} \text{ for all } (s,t) \in E$$

If $x$ is symmetric, an undirected graph $G(x)$ is defined accordingly. If $x$ is dichotomous, the uniform edge weights are often omitted.

Note that a graph representation (of a single network) may require loops but no multiple edges.

Unobserved and zero-value data, however, are indistinguishable. There is of course the conventional trick to introduce a special numeral representing missing data, but it requires careful documentation and special treatment. Such special treatment is common for diagonal entries in matrices representing one-mode networks of relationships that nodes cannot have with themselves.

We will not utilize matrix representations here and refrain also from defining any of the various graphical representations (see Figure 3 and, for example, Brandes et al., 2013).

Figure 2. Contact nominations among physicians in the Medical Innovation Study (Coleman et al., 1966). Contacts represent the presence of at least one nomination in questions on advice seeking, discussion, or friendship. Node color indicates first known date of tetracycline prescription. Earlier adopters are darker, no confirmed prescription for white nodes. Do central physicians tend to adopt earlier?
Instead, for its graphic notation and expressiveness, we will draw on a representation as a valued relation for the most part.

**Definition 3 (Relation of a Network).** For a network \( x \in \mathcal{W}^S \), its (valued) relation \( R(x) \subseteq S \times (\mathcal{W} \setminus \{0, \infty\}) \) is defined by

\[
((s,t),a) \in R(x) \iff (s,t) \in S \land x_{st} \notin \{0, \infty\}
\]

We denote \(((s,t),a) \in R(x)\) by \((s,t;a)\) or \(s \rightarrow^a t\). A subscript \((s,t;a)_x\) or \(s \rightarrow^a t_x\) is added if the network is not clear from context. In the latter case, we may omit the implied value \(a = x_{st}\). If \(x\) is dichotomous or it only matters that a relationship exists, we often write \(s \rightarrow t\) for \(x_{st} \neq 0, \infty\).

Note that, technically, relation representations do discriminate between unobserved and zero-valued dyads because they are defined over \(S\) rather than \(\mathcal{N} \times A\).

Since a network \( x \in \mathcal{W}^S \) and its relation \( R(x) \) are equivalent, we often refer to \( x \) as a relation, too. Each element \( x_{st}\) is a relationship, and if \(x_{st} \neq 0, \infty\), we call it a link. In social science contexts, nodes generally represent (social) actors, whereas links represent (social) ties.

Most representations are easily transformed from one into the other. Since the distinction of unobserved and zero-valued dyads may be lost, our use of the term representation, however, is slightly more liberal than the common implication of an isomorphism.

**Derived relations**

The relations of interest are not necessarily directly observable. Borgatti et al. (2009) list a number of relations on which data is commonly collected, and many methods start by transforming them, explicitly or implicitly, into other relations.

For instance, when interested in access to information held by individuals, direct interactions via personal or institutionalized contacts may be observable but an actor’s access to additional information via intermediaries is generally difficult or too costly to observe, for example, because respondents might not even be aware of it. In such a scenario, it is common to model indirect access to information via a combination of direct interactions and a model of information propagation.

We attempt to reformulate methods for analysis in such a way that necessary transformations are performed prior to, rather than as part of, an analytic procedure. If analytic procedures are indifferent to whether a relation has been observed directly or derived from, say, a proxy or a flow-process model, the reasoning about methods becomes independent from the reasoning about the social processes that establish the indirect relations, and the overt use of indirect relations caters to their empirical validation. This should also be understood as a generalization of the ideas of individual distances from Burt (1976) and ego algebras (Breiger and Pattison, 1986).

A possible preprocessing is therefore the transformation of an observed or otherwise given network \(x\) into a derived network \(\tau(x)\), during which the dyadic domain might be extended and the value range might be changed. The derived network captures the indirect relationships established by, say, an assumed flow process.

By making this step explicit, the consequences of the transformation can be studied separately from the other aspects of an analysis. A flow process, or any other rationale for establishing derived relations, is thus more open to specific theorizing (Borgatti and Halgin, 2011) and empirical validation on the dyad level.

Derived relations are therefore a topic of themselves. Only to illustrate how the concept of network position stresses the importance of indirect relations, we here distinguish two important types of transformations, neighborhood-based and walk-based. These are not exhaustive but do cover the bulk of transformations in use today.

In neighborhood-based transformations, values associated with dyads are rescaled, normalized, classified, symmetrized, or otherwise modified by considering local configurations and overall value distributions. For instance, in a network variable \(x\) obtained from questionnaires in which actors self-report their close relationships, respondents may have different baselines in the assessment of the strength of their relationships, which may result in asymmetric or even unconfirmed ties although, in theory, the relation is symmetric. If this asymmetry is not of interest in itself, network \(x\) might therefore be transformed into \(\tau(x)\) by shifting values and symmetrizing relationships.

Note that the network in Example 2 is the result of transforming three observed relations (nominations for advice, discussions, and friendship) into one (contact nominations) dyad for dyad.

Walk-based transformations, on the other hand, are obtained from the combination of values along concatenations of relationships. A (directed) walk is a sequence of non-zero relationships such that the target of one is the source of the next. A walk that contains each node no more than once is called a (directed) path. A (directed) graph is said to be (strongly) connected, if there is a (directed) path from every vertex to every other vertex.

The concatenation of values along walks and the aggregation of values of multiple walks with the same source and target give rise to a multitude of indirect relations, many of which can be obtained from appropriately defined semirings (Batagelj, 1994; Gondran and Minoux, 2008).

The length of a path, for instance, is defined as the sum of the values of relationships in the path. The (geodesic or shortest path) distance, \(\text{dist}(s,t)\), from a vertex \(s\) to a vertex \(t\) is the minimum length over all paths from \(s\) to \(t\). This is, in fact, the relation we will be interested in for the Medical Innovation data.
Definition 4 (Distance Transform). Let \( x \in \mathbb{R}^S \) be a one-mode network on the domain \( S \subseteq (N \times N) \diagdown (N) \) with a range of positive real values representing, for instance, the length, transfer duration, or cost of a link. The (shortest path or geodesic) distance transform is the network \( \text{dist}(\cdot) : N \times N \to \mathbb{R}_{\geq 0} \) with
\[
\text{dist}(s, t) = \begin{cases} 
0 & \text{if } s = t \\
\text{dist}(s, t) & \text{if } t \text{ is reachable from } s, \text{i.e., } s \xrightarrow{x} t \\
\infty & \text{otherwise}
\end{cases}
\]
where \( \text{dist}(s, t) \) is the minimum sum of values along any directed \((s, t)\) path. Note that both domain and range may be extended in the transformation.

Figure 3 gives the network \( \text{dist}(\cdot) \) of shortest path distances derived from the direct interactions \( x \) in one of the cities from Example 2. Our units of analysis below will be the rows or columns of such matrices, whereas previous studies have aggregated them into numerical indicators of centrality and similarity.

Other commonly used indirect relations include reachability, walk numbers, effective resistance, reliability, and maximum network flow.

As a special type of derived relationships, we mention imputations, where unobserved dyads are often assigned values that are obtained from dyad-based or walk-based transformations.

Network positions and positional dominance

This section introduces our main contribution, a formal definition of position in networks. It is derived as a generalization of position in social space by the addition of relational dimensions. It aligns naturally with the dyad-centric view of network data and facilitates more uniform, systematic, and transparent analysis.

By generalizing the concept of dominance accordingly, we are paving the way to describe classes and hierarchies of social network positions. These ultimately lead to a reconceptualization and refinement of network centrality analysis as outlined in the next section.

Positions

Formally, a position can be defined as a vector of coordinates that describe a location by its distance from an origin along each of several dimensions. In reference to this idea, the term “social position” has been applied to evoke the image of a location in society or in a status hierarchy, although the reference is often rather metaphorical.

If social structure is meaningfully studied through the lens of data on actors and their relations, as is the underlying proposition of any empirical social network research, there is an implied congruence between an actor’s social position and the role enacted. This idea naturally extends to similar positions and similar roles.

The following formal notion of social space and positions therein was proposed by Peter Blau. We will generalize his position concept by extending it with relational dimensions that are qualitatively different.

Example 3 (Blau Space). Blau (1977a, 1977b) conceptualizes a social structure as the distribution of social positions in a multidimensional space (Blau, 1977b: 4). The structure of positions in this space and relations among these positions are explicitly conceived as relative to the dimensions selected.
The dimensions of any such space are made up of sociodemographic, socioeconomic, or other variables associated with individual actors and are referred to as structural parameters. In Blau’s terminology, nominal variables represent heterogeneity, whereas quantitative variables represent inequality in a dimension.

A position thus characterizes a homogeneous set of equivalent individuals having the exact same parameter values, and different positions signal differentiation into groups and social strata. Positions in this space are referred to as social positions because they are assumed to reflect and affect role relationships among groups of individuals.

In contrast to microsociological approaches, where the interest would be in describing the social relationships of individuals, the macrosociological perspective focuses on social associations between positions, that is, the degree to which ties are formed between aggregates of individuals occupying certain positions.

As before, the social space perspective posits that the characteristics of social actors coincide with aspects of their relationships. As this association is generally unknown, changing, or does not exist a priori, the social space should be extended to accommodate both attributes and relations for more general investigations. We therefore conceive of positions in social networks as positions in a space made up of both attribute dimensions and relationships.

We start by defining positions in a single network.

**Definition 5 (Network Position).** Given a network \( x \in \mathcal{W}^S \) on a domain \( \mathcal{S} \subseteq \mathcal{N} \times \mathcal{A} \), the position of \( i \in \mathcal{N} \) in \( x \) is defined as

\[
\mathcal{1} = \text{pos}(i | x) = \left\{ \begin{array}{l} x_i \in x \\
\quad \text{s.t.} : (i, t) \in \mathcal{S} \end{array} \right. 
\]

The symbol \( \mathcal{1} \) is chosen because it resembles a graphical representation of a node in a network. It is used as shorthand notation when the space in which the position is located is clear from context.

For each network \( x \in \mathcal{W}^S \) with \( \mathcal{S} \subseteq \mathcal{N} \times \mathcal{A} \), the position of \( s \in \mathcal{N} \) has coordinates \( x_s \in \mathcal{W} \) along dimensions \( t \in \mathcal{A} \). Recall that \( \mathcal{A} \) might equal \( \mathcal{N} \), and that the network may represent a derived relation.

Example 1 can be viewed as positioning painters in a four-dimensional social space in which each dimension is a separate variable. But it can also be viewed as a two-mode network \( x \) so that a painter’s position is defined by coordinates that correspond to its valued relationships (degrees of perfection) with distinct affiliations (skills). Clearly, these views are equivalent. The network perspective is contingent, however, on the assumption that the rating values have the same interpretation in all four dimensions, that is, \( x \) maps dyads to values that are comparable not only within but also across skill dimensions.

In Example 2, the position \( \text{pos}(i | x) \) of a physician \( i \) in contact network \( x \) is defined by the presence and absence of ties to the other physicians in the same city. As stated before, we will not be interested in the observed contact network \( x \) but in the distance relation \( \text{dist}(x) \) derived from it. A position \( \text{pos}(i | \text{dist}(x)) \) in this network consists of \( i \)’s shortest path distances to all other nodes, and we may or may not choose to include physicians from other cities at infinite distance to make positions comparable across cities.

By defining positions only for nodes in relation to their affiliations, undue comparison of different modes (nodes vs affiliations) or directions (outgoing vs incoming) is prevented. In a two-mode network, affiliations define the space in which nodes are positioned, and by duality, these roles can be exchanged via transposition, but they cannot be mixed. In a one-mode network, only outgoing relationships are considered, and by reversal, the focus can be shifted to incoming relationships, again via transposition, but the meaning of an incoming relationship cannot be equated with that of an outgoing one.

It should also be noted that we follow Blau (1977b) and Burt (1976) in defining a position as a particular combination of coordinates. Hence, a position corresponds to multiple nodes only if they all agree on these values. The seemingly more common understanding of positions in networks as emerging from classes of nodes with equivalent relations (Borgatti and Everett, 1992; Breiger et al., 1975; Doreian et al., 2005; Faust, 1988) will later be obtained via equivalence relations on our finer-grained position.

Extension to multiple networks and the inclusion of nodal attributes are straightforward because positions are sets in which the elements are not only values but also tuples encoding the relation from which they originate, the source, the target, and the value. An attribute \( z : \mathcal{N} \rightarrow W(z) \) can be represented as a diagonal relation with relationships \( i \sim i \), \( i \in \mathcal{N} \), so that positions become

\[
\mathcal{1} = \text{pos}(i | x^{(1)} \ldots x^{(k)}, z^{(1)} \ldots z^{(l)}) \\
= \bigcup_{y \in x^{(1)} \ldots x^{(k)}, z^{(1)} \ldots z^{(l)}} \left\{ i \sim t : (i, t) \in \mathcal{S} \right\}
\]

While we will define most concepts for single networks, it should therefore be apparent that they generalize to multiple networks and attributes. Our notion of position is therefore closely related to, but considerably more general than, those presented in Burt (1976) or Faust (1988).

**Dominance**

Since we define the position of a node in terms of all its relevant attributes and relationships in basic or derived networks, subsequent analyses in which the unit of analysis is the node can focus on comparison and evaluation of positions.
The following definition extends a preorder on the value range (which allows to compare two relationships in terms of their value) to a preorder on the set of positions and thus represents an aggregate, though only relative, cost or benefit even before introducing any quantification.

**Definition 6 (Positional Dominance).** Let \( x \in W^S \) be a network on a dyadic domain \( S \subseteq N \times A \) with values in a range \( W \) that is preordered by \( \preceq \).

For \( i, j \in N \), we say that \( \{i\} \) is dominated by \( \{j\} \) (under the assumption of total homogeneity), denoted \( \{i\} \preceq \{j\} \), if there exists a permutation \( \pi: A \rightarrow A \) such that for every \( (i, t) \in S \) we have

\[
(j, \pi(t)) \in S \quad \text{with} \quad x_i \preceq x_{j \pi(t)}
\]

In one-mode networks, the dyads \((i, j), (j, i)\) and the diagonal might require special treatment when comparing the positions of \( i \) and \( j \). If \( \pi \) is restricted to the identity permutation, we say that \( \{i\} \) is dominated by \( \{j\} \) under the assumption of total heterogeneity and write \( \{i\} \preceq \{j\} \).

Positional dominance under heterogeneity is the usual notion of dominance among vectors and corresponds to the most basic requirement in conjoint measurement which is referred to as monotonicity or independence. It extends the preorder on values \( W \) without making any additional assumptions. Hence, the use of the same symbol.

Clearly, dominance under heterogeneity implies dominance under homogeneity

\[
\{i\} \preceq \{j\} \Rightarrow \{i\} \preceq \{j\}
\]

but not vice versa. Note that dominance under total homogeneity is conveniently checked by sorting relationship values in weakly decreasing order and comparing them front to back.

In the network of ratings in Example 1, Penni’s positions are dominated by Raphael’s, but not Rubens’, under the heterogeneity assumption, that is, when skill dimensions need to be considered separately. This is because de Piles rated Penni at 15 in drawing and thus higher than Rubens.

If we allow any skill to substitute for another, that is, under the assumption of total homogeneity, Penni’s position is dominated by both Raphael’s and Rubens’. The positions of Raphael and Rubens, however, remain incomparable even under the homogeneity assumption.

For dichotomous variables such as the contact network in Example 2, we get

\[
pos(i | x) \preceq pos(j | x) \Leftrightarrow pos(i | \text{dist}(x)) \preceq pos(j | \text{dist}(x))
\]

if the dyads involving \( i \) and \( j \) are treated appropriately. In other words, the distance transformation does not add any information under the heterogeneity assumption. This is because the first edge on any shortest path from \( i \) can be matched by an edge from \( j \) to the same vertex and because unit distances correspond directly to ties. Under the common assumption of homogeneity, however, the dominance orders of positions in \( x \) and \( \text{dist}(x) \) are unrelated in general.

Restricting the admissible permutations \( \pi \) of affiliations is a powerful way to integrate degrees of homogeneity by making some affiliations substitutable by certain others. For instance, if affiliations (which may be nodes themselves) are grouped, we may allow to compare the relationship of \( i \) and some affiliation with a relationship of \( j \) and another affiliation as long as the two affiliations belong to the same group. For the physicians in Example 2, a relevant grouping may be obtained from their specialization in the medical field.

The approach is thus open to the use of more complex comparison operators for positions which may include node attributes and multiple relations. In fact, this is the key to obtain standard node equivalence relations from positional comparison. We here limit ourselves to the case of comparing single relationships in terms of their values under total homogeneity or heterogeneity only. Though inappropriate in general, this represents the most common cases and serves solely to simplify the exposition. General heterogeneities and multivariate data are beyond the scope of this introduction to the positional approach.

**Centrality**

Under the term of centrality analysis, we subsume any method that is designed to identify the relative structural importance of nodes in a network. Depending on the substantive meaning of structural importance, it might also be described by terms such as status, prominence, social capital, power, or influence.

According to Wasserman and Faust (1994), centrality is “[o]ne of the primary uses of graph theory in social network analysis” (p. 169). It is remarkable, therefore, that the assessment of Freeman (1979) that “there is very little agreement on the proper procedure for its measurement” still applies.

In the positional approach advocated here, the primary use of graph theory is in the determination of indirect relations, in particular, those related to flow processes, whereas centrality is treated as a problem of ranking the positions that are defined in terms of these relations.

**Centrality indices**

The chapter on centrality in Borgatti et al. (2013) opens with the statement that “[c]entrality is a property of a node’s position in a network.” What follows, however, is focused on the usual approach of constructing an index of centrality, that is, the restriction of “property” to a quantitative variable and the somewhat circular definition of position in terms of the assigned number.
Given a network, a centrality index is a procedure that yields a mapping \( c : \mathcal{N} \rightarrow \mathbb{R} \) of nodes to real numbers. The underlying assumption of centrality analyses based on indices is that positions are manifestations of a property that is measurable, that is, a quantity associated with the position. Hence, the equivocal reference to centrality indices as centrality measures.

In measurement-theoretic terms (e.g. Hand, 2010), this is an operational measurement at best because the precise definition of the procedure to calculate a centrality score starts only after the dyadic data has already been obtained. Moreover, the outcome is usually treated as a ratio-scale quantity that can be used in regressions and other analytic tasks. Justifying a measurement procedure for a quantitative variable is a daunting task, though.

Avoiding the difficulties and context-specific limitations of measurement theory, several authors have taken axiomatic approaches to characterize centrality indices (e.g. Kishi, 1981; Ruhnau, 2000; Sabidussi, 1966), and such characterizations may help selecting or ruling out indices for specific purposes. See Boldi and Vigna (2014) for a recent account with many references. A more substantively oriented approach is to discriminate centrality indices by social processes that are modeled in their definition (Borgatti, 2005; Borgatti and Everett, 2006; Friedkin, 1991).

The only universally accepted requirement shared by all approaches, however, appears to be that the center of a star network must be maximally central (Freeman, 1979).

Before attempting to place the positional approach in this discussion, let us consider four prototypical centrality indices. Three of them are discussed extensively in Freeman (1979), and the fourth is the main representative of a group of indices based on feedback. For ease of expositions, we follow the common approach of defining the indices as graph invariants on a connected simple undirected graph \( G = (V, E) \). A centrality index thus yields a mapping \( c : V \rightarrow \mathbb{R} \).

**Degree centrality.** The basic idea is to evaluate the importance or involvement of an actor simply by counting incident ties, \( c_{\text{deg}}(i) = \deg(i) \). In directed graphs, indegree and outdegree may be distinguished, and in a weighted graph, the common generalization is to add up the weights on edges incident to \( i \). Variants trade-off the number of ties with their values (Opsahl et al., 2010) and, therefore, effectively combine two networks into one.

**Closeness centrality.** In closeness centrality, a node is considered to be in a beneficial position if it is close to others in terms of geodesic distance. This is interpreted either as efficient reachability or as relative independence from intermediaries. The purest definition of closeness is \( c_{\text{cl}}(i) = \sum_{t \neq i} \text{dist}(i, t) \) (Sabidussi, 1966), so that higher closeness centrality is attributed to nodes with smaller values. Again, generalizations distinguish between incoming and outgoing distances, and valued networks may be considered. Note, however, that comparison of distance sums is meaningful only if all distances are finite.

**Betweenness centrality.** If nodes relate to each other via shortest paths only, the dependency of a pair \( s,t \in V \) on a broker \( i \in V \) can be modeled as \( \delta(s,t | i) = \frac{\sigma(s,t | i)}{\sigma(s,t)} \), where \( \sigma(s,t) \) is the number of shortest paths from \( s \) to \( t \), and \( \sigma(s,t | i) \), the number of those containing \( i \neq s,t \). Betweenness centrality is then defined as the total dependency of pairs of other nodes on \( i \), \( c_{\text{B}}(i) = \sum_{s,t \in V} \delta(s,t | i) \) (Freeman, 1977) and generalizes to directed and valued networks without modification. The conceptual control versus dependence duality of betweenness and closeness actually holds also in a formal sense (Brandes et al., 2016).

**Eigenvector centrality.** Degree centrality can be generalized by taking into account the centrality of neighbors in a feedback loop. In other words, an actor is considered central to the extent to which it is directly connected to central others. If \( \lambda \) is the largest positive eigenvalue of the adjacency matrix \( A(G) \) of a connected graph \( G \), the principal eigenvector \( c_{\text{e}} \) of \( A(G) \) is the unique (up to scaling) positive vector satisfying \( A(G)c_{\text{e}} = \lambda c_{\text{e}} \). Eigenvector centrality is thus defined as \( c_{\text{e}}(i) = \lambda^{-1} \sum_{t \in \mathcal{N}(i)} c_{\text{e}}(t) \) (Bonacich, 1972), where \( \mathcal{N}(i) \) denotes the neighbors of \( i \). It generalizes to weighted graphs but requires connectivity (strong connectivity for directed graphs).

These four indices appear to be based on very different criteria: many strong ties, short distances, frequent brokerage, and many ties to central actors. Like most centrality indices, however, the above can be expressed as representations of positions. A position is represented numerically if the dominance relation among positions is respected by the natural order of the assigned numbers

\[
1 \leq 1 \Rightarrow c(i) \leq c(j)
\]

The above indices are actually easy to obtain as sums

\[
c(i) = \sum_{x \in \mathcal{N}(i)} \tau(x) x
\]

over suitably defined indirect relations \( \tau(x) \) derived from the input relation \( X \). For degree centrality, \( \tau \) is simply the identity transformation, and for closeness, it is the distance transform \( \tau(x) = \text{dist}(x) \). We can define a dyadic dependency \( \delta(s,t) = \sum_{x \in \mathcal{N}(s)} \delta(s,t | x) \), so that betweenness becomes degree centrality in this dependency transform. Even eigenvector centrality is degree centrality in a transformed network, where each entry \( o(x)_{st} \) represents the limit proportion of \((s,t)\) walks of length \( k \) for \( k \to \infty \) (Benzi and Klymko, 2015). This transformation is obtained directly from the power iteration method to compute the principal eigenvector (e.g. Golub and Van Loan, 1996).
The four prototypical indices, and in fact most centrality indices, are therefore additive representations of positions in a derived network.

Factoring out that the indirect relation is an important aspect of the restructuring of centrality analysis via network positions. It suggests to validate the often implicit assumptions on what constitutes an indirect relation on the dyadic level and to replace the indirect relation with another if the fit is poor. It also suggests that representations of positions other than the sum of relationship values may yield a better quantification. Eccentricity centrality (Hage and Harary, 1995), for instance, is given by the maximum value in positions based on distance.

The other important aspect, however, is that we may even avoid the quantitative representation of a position altogether. Using positional dominance instead, comparisons can be made independent of a specific representation because dominance is preserved by definition. Quantitative representations are limiting because they impose a one-dimensional ranking structure and require a justifiable aggregation operator such as the sum or maximum to turn a vector into a scalar.

This is illustrated below by applying positional dominance to the two example data sets. In the first example, the enforced unidimensionality of centrality rankings via indices disguises important structural patterns that are caught by the step-wise approach via positional dominance. In the second example, we can conclude that network distance has little to no effect on innovation adoption without the trial-and-error approach via particular indices.

The balance of painters

Graddy (2013) relates the ratings of artists from Example 1 to art market returns for their paintings and finds that, overall, the subjective ratings of de Piles “have withstood the test (...) of time.” Using regression analysis on returns over time, she finds that only two of the four dimensions have a relatively constant value as predictors, whereas the influence of drawing is declining and the influence of color is rising.

As in most other works on de Piles’ table, there is recurrent reference to the notion of a ranking by totals. When giving each skill dimension equal weight, this corresponds to degree centrality in the valued two-mode network and therefore directly relates to the way in which, say, university or product rankings are constructed. The painter ranking based on degree is represented in the ordering from left to right in Figure 4.

Quite generally, the availability of numerical scores appears to trigger an impulse to combine them into a ranking, even though de Piles’ introduction is ambiguous as to whether the Balance des Peintres is meant to be interpreted this way.

Note that combining ratings by addition implies that they are measured on an interval scale. It seems safe to conclude from the original explanation that de Piles intended a common scale for degrees of perfection across dimensions, but despite his reference to “degrees” of perfection and having “divided” the range from 0 to 20, it is not granted that these levels are meant to be equidistant. Assuming that the ratings are additive so that, say, an increase of two levels at the bottom end of the composition ratings compensates for a decrease of two levels in the mid ranges of the color ratings is therefore likely to be a stretch.

In summary, a ranking based on weighted degree centrality rests on two implicit assumptions; dimensions (skills) are substitutable for one another, and weights (ratings) are measured on an interval scale. Neither seems to be an intended property of de Piles’ ratings.

If we apply positional dominance and assume homogeneity of the skill dimensions, instead, the comparabilities shown in Figure 5 are obtained. Nodes are arranged vertically according to the number of others they dominate only to ensure that all dominance arrows are upward and thus ease the reading. It is open whether de Piles actually intended to suggest that Raphael and Rubens should be considered equally skilled painters. With respect to positional dominance under the weakest of assumptions, that is, only those that may be read into the table with certainty, they are incomparable. Both are attributed a high degree of perfection in three dimensions but not all four.

| Artist | Drawing | Composition | Color | Expression | Total |
|--------|---------|-------------|-------|------------|-------|
| Raphael | 18 | 17 | 12 | 18 | 65 |
| Carracci | 17 | 15 | 13 | 13 | 58 |
| Tintoretto | 14 | 15 | 16 | 4 | 49 |
| Rubens | 13 | 18 | 17 | 17 | 65 |
| Titian | 15 | 12 | 18 | 6 | 53 |

In fact, they are not simply the dominant painters according to de Piles’ rating, but the most highly accomplished in two different schools that are largely incomparable as groups because of their distinct foci mirroring the debate around relative importance of classic drawing (disegno) and vivid color (colore). This is reiterated in the coarsened positions obtained from partitioning the 54 painters with complete ratings into those dominated by either Rubens but not Raphael (19), Raphael but not Rubens (23), neither (5), or both (8).2

This example demonstrates that the one-dimensional view imposed by centrality indices may be inappropriate. As is the case here, substantively justified comparisons may not suffice to form a complete ranking but quantitative centrality scores always do.

From de Piles’ description of his ratings, it is clear only that a painter is considered more skilled if there is no dimension in which he receives a lower rating. As noted above, this leads to the incomparabilities that degree centrality always resolves by adding assumptions that are rarely argued for homogeneity (here, equal relevance of skills) and additivity (here, interval scale ratings).
Figure 4. Shortest path distance transform $\text{dist}(x)$ of the Medical Innovation data restricted to Bloomington. Dots represent infinite distances, that is, the target (column node) is not reachable from the source (row node). Note that the highlighted distances correspond exactly to ties in the contact network $x$. 
Positional dominance preorders, on the other hand, can be constructed with or without these assumptions and therefore allow for the possibility that a network is, for instance, multicentric or heterogeneous, or that weights are not on an interval scale.

As an example, Figure 6 represents the partial ranking obtained when assuming homogeneity, but not additivity. Compared to the case of distinguished dimensions in Figure 5, many more painters have become comparable. Observe that many dominance relationships are implied by transitivity but not shown to avoid excessive overlap. Although the ranking is more determined, it is still not complete. Most notably, Raphael and Rubens remain incomparable despite dominating everyone else, but Carracci who is dominated only by Rubens.

Similarly, six painters form the bottom of the hierarchy by not dominating anyone.

Weighted degree centrality imposes a complete ranking by introducing comparabilities between all pairs of artists that are incomparable even when it is assumed that mastery of one skill can compensate for the lack of another. The various reasons for which an artist might be considered excellent thus become indistinguishable and if the network of ratings is intended to explain another variable such as art market returns, all these reasons have to be equally good explanations.

The same strong assumptions are ingrained in other centrality indices, since most of them can be rewritten into additive representations of positions as indicated in Section “Centrality Indices.” Corresponding to the effect unveiled for the two-mode network of painter ratings, the aggregate nature of centrality indices hides the fact that nodes may be equally central for very different reasons.

**Medical innovation**

The literature on Example 2 is extensive but findings are conflicting (Kilduff and Oh, 2006; Valente, 2011). The original study concluded that the new drug was adopted earlier by physicians more fully integrated into the medical community (Coleman et al., 1966). Part of the reasoning is that opinion leaders provide guidance in ambiguous situations and therefore tend to make decisions that eventually become common. We focus on the use of centrality as an indicator for the integration or prominence of physicians in the medical community to demonstrate a second use case for positional comparison.

Valente and Foreman (1998) propose a novel centrality index and, among other things to demonstrate its utility, argue that it displays the strongest association between centrality and adoption in comparison to previous indices based on similar concepts. The index is a variant of the closeness centrality from Section “Centrality Indices,” which we redefine here to account for the directedness of contact nominations and the number of nodes in the network.
Definition 7 (Closeness Centrality). In a one-mode network \( x : S \rightarrow W \) for which the graph \( G(x) \) is strongly connected, the (incoming) closeness centrality \( c_C : N \rightarrow R \) is defined by

\[
c_C(i) = \left[ \sum_{s \in N \setminus \{i\}} \text{dist}(s,i) \right]^{-1}
\]

(Deauchamp, 1965)

Closeness centrality is thus the inverse of the average distance of a node from all others. The actions of a physician are assumed to be known and relevant to those who nominate him or her and to a lesser extent also to those nominating these. Since larger distances lead to larger distance sums, taking the inverse establishes a ranking in which higher visibility or influence results in larger centrality scores.

The restriction to strongly connected graphs is necessary to make the distance sums comparable. It appears that Valente and Foreman (1998) compare to a version of closeness that does not account for the fact that the contact networks are not strongly connected. We therefore study their proposed centrality index in comparison to two other variant indices that have been proposed to maintain the intuition of closeness in the general, not necessarily connected, case in which distances are considered only for reachable dyads.

Let \( r(i) = \left\| s \neq i : s \rightarrow i \right\| \) denote the number of nodes that can reach \( i \in N \) on a directed path, and let \( \text{diam}(G(x)) \) be the length of a longest directed path in the graph \( G(x) \) of the network

\[
c_G(i) = \frac{r(i)}{n-1} \left[ \sum_{s \in N \setminus \{i\}} \frac{\text{dist}(s,i)}{r(i)} \right]^{-1}
\]

\[
c_H(i) = \frac{1}{n-1} \sum_{s \in N \setminus \{i\}} \frac{1}{\text{dist}(s,i)}
\]

\[
c_I(i) = \frac{1}{n-1} \sum_{s \in N \setminus \{i\}} \left[ 1 - \frac{\text{dist}(s,i) - 1}{\text{diam}(G(x))} \right]
\]

We refer to the first variant index, \( c_G \), as generalized closeness (Lin, 1976). It weights the average of all finite distances to \( i \) by the proportion of nodes that can reach \( i \) and is thus a proper generalization coinciding with \( c_C \) if \( G(x) \) is strongly connected.

The second index, \( c_H \), is referred to as harmonic closeness (Gil-Mendieta and Schmidt, 1996) because it is the inverse of the harmonic (rather than arithmetic) mean of distances to \( i \). Note that the inverse approaches zero for growing distances.

The third, \( c_I \), is actually a rewritten version of what is introduced as (normalized) integration (Valente and Foreman, 1998). Instead of taking inverses, it reverses the
Figure 7. The variant closeness indices are highly correlated on the Medical Innovation data from Example 2, even when compared across the four city networks \((n = 246)\).

interval of attained values by replacing finite distances \(\text{dist}(s,i)\) with \(\text{diam}(G(x)) + 1 - \text{dist}(s,i)\) and infinite distances with zero.

Arguments for or against any variant index of centrality are most often based on applicability, functional form, or mathematical convenience. Figure 7 shows that there is very little difference between the above closeness variants on the Medical Innovation data. The curved shape in the scatterplots involving integration can be traced back to the different ways in which the ordering of values is reversed, linearly versus hyperbolically. The stronger curving of a larger interval also explains why nodes from the largest network, Peoria, are separated despite all indices being normalized for size and thus indicates that values should rather not be compared across the four networks.

As indicated in Figure 8, application of any of the three variant closeness indices yields a weak to moderate association between centrality and adoption date. Performance differences between indices are minor and although integration is favored in Valente and Foreman (1998), this may only be due to the particular adaptation of standard closeness that is used. Therefore, we want to address the question whether it is indeed the effect as such that is weak or whether a different index aggregating distances in another way could do better.

The common approach of trying out variant indices results in three difficulties:

- If the variant index performs well, it remains open whether yet another variant could lead to an even stronger result.
- If the variant index performs poorly, it is not clear how to go about finding a better one.
- Either way, it is not clear whether the performance of an index is due to the (mis)fit of the underlying rationale or whether it is coincidental on the particular data.

We strive for more general statements about any distance-based centrality index by studying positional dominance in the distance transform.

Consider therefore the network \(\text{dist}(x): V \times V \rightarrow W\) of shortest path distances. As the focus is on incoming relationships from nominations, we define positions
for all \( i \in \mathcal{N} \). A position thus consists of the distances from all physicians, that is, the columns in the matrix representation of \( \text{dist}(x) \) (cf. Figure 3). Unlike in the previous section, now the other actors make up the dimensions, and the position of a node in a one-mode network is characterized by relationships with these other actors. Whether the diagonal is included (as it is here) or not will not make a difference.

As before, we define two notions of dominance for positions in the distance network

\[
\mathbf{1} \leq \mathbf{1} \quad \text{if} \quad \text{dist}(s, i) \geq \text{dist}(s, j) \quad \text{for all } s \in \mathcal{N} \setminus \{i, j\}
\]

\[
\mathbf{1} \leq \mathbf{1} \quad \text{if} \quad \exists \text{a permutation } \pi: \mathcal{N} \rightarrow \mathcal{N} \quad \text{s.t. dist}(s, i) \geq \text{dist}(\pi(s), j) \quad \text{for all } s \in \mathcal{N}
\]

The first relation allows for heterogeneity of actors to influence the comparison of positions. As noted above, however, dominance under heterogeneity in the distance transform is the same as in the contact relation itself and therefore very sparse. Of 117 physicians in Peoria, for instance, 11 have no incoming ties and are therefore dominated by everyone. Of the remaining \( \binom{106}{2} = 5565 \) pairs, only 103 are actually comparable in \( \leq \).

Some degree of homogeneity must therefore be assumed for otherwise distances are not taken into account. Indeed, the assumption of total homogeneity is surprisingly universal. Very few centrality indices, and certainly none of the closeness variants above, distinguish between nodes. The second dominance relation \( \leq \) therefore extends the first by incorporating homogeneity.

As a side note applying the homogeneity assumption to the contact network \( x \), rather than the distance network \( \text{dist}(x) \), would yield the complete ranking that corresponds to indegree centrality.

The crucial observation is that in spite of the looser condition for dominance and thus stronger determinism of the partial ranking, positional dominance is necessarily preserved by centrality indices that are solely based on distances and improve if these are lowered. This is easily verified for the closeness variants above

\[
\mathbf{1} \leq \mathbf{1} \implies \begin{cases} c_G(i) \leq c_G(j) \\ c_H(i) \leq c_H(j) \\ c_I(i) \leq c_I(j) \end{cases}
\]

by observing that they are monotone in all pairwise distances, including infinite ones.

Essential information that can be extracted from homogeneous positional dominance is presented in the first columns of Figure 9. While only those 125 physicians with adoption data are listed, their positions are determined and compared in the overall networks. Even without any numerical operations on the distances, a high percentage of node pairs are comparable and few of them are equivalent.

To compare the partial rankings obtained from positional dominance with the order in which physicians adopted the new drug, we use rank correlation on the subset of comparable pairs as given by Kendall’s

\[
\tau_B = \frac{\text{concordant pairs} - \text{disconcordant pairs}}{\sqrt{\left(\text{pairs} - \text{mutually dominating}\right) \cdot \left(\text{pairs} - \text{same date}\right)}}
\]

Concordant and disconcordant pairs are those in which both positional dominance and the date variable rank one node strictly above the other, either in the same or in opposite order. Since distance-monotone centrality indices can only extend but not contradict positional dominance, an optimistic upper bound on the highest correlation achievable by any closeness variant is obtained by assuming that it ranks all incomparable pairs in (dis)accordance with adoption dates.

With the exception of the small network from Galesburg, none of the networks displays a strong association between position and time of adoption, and as we have just argued, no variant closeness index can ever yield a substantial improvement.
A negative result such as this cannot be obtained by trying out any number of newly invented indices. The finding does not rule out, however, that a smaller distance effect is moderated by other traits of physicians such as their specialization or exposure to marketing efforts (Van den Bulte and Lilien, 2001), or that there is a structural effect based on a different indirect relation.

Discussion

The methodological innovation put forward in this article is to make the explicit definition of network positions the pivotal step in network analytic pipelines.

By conceiving of the position of a node in a network as the entirety of its relevant relationships and attributes, we are separating the preparation of data from its analysis.

Treating derived relations as part of the definition of a position, rather than an essential ingredient of a particular centrality index, facilitates method modularization.

As noted in Borgatti (2005), for example, there are many flow processes in networks that are conceptually clear but have no associated centrality index. The positional perspective suggests that it is not the indices that are missing but the proper justification and validation of derived relations. Domain-specific theory should motivate the use of one indirect relation over another, and dyad-level testing of alternative derivations is a promising direction for future work.

A similarly interesting question is the accuracy of indirect relations derived from different observed relations. In general, indirect relations are determined computationally only because they are too difficult to obtain otherwise. How could subjects be aware, for example, of the potential influence of someone they do not even know? If the relation of interest is difficult to measure directly, which is the most appropriate combination of observable relations and transformations?

Once positions are defined in terms of the relations of interest, independent of whether they have been obtained directly or indirectly, the analysis can focus on the relative merit of a position without considering the process that led to it; just like ordinary variables are fully determined prior to their use as dimensions of social space.

It may appear that the network has vanished at this point. If we think about positions as feature vectors, dependencies that make up the social patterning are indeed encapsulated in the derived relations (Ziegler, 1987). While it is a fruitful perspective to think of positions as atomic entities amenable to conventional multivariate analysis and machine learning, there is a distinctive twist. Since dimensions of these feature vectors are defined by affiliations, which may even be the nodes themselves, coordinates may be compared more flexibly and combined across dimensions. Homogeneity assumptions on various levels inbetween those used in the examples above open up a huge space of options.

We have discussed positional dominance as the common aspect of all centrality indices based on the same, possibly derived, relation. While it yields (partial) rankings of nodes that those centrality indices necessarily respect and thus provides an alternative way of characterizing centrality in general, it also yields indifference classes of nodes that mutually dominate each other.

For the adjacency relation, dominance under heterogeneity yields indifference classes that correspond exactly to structural equivalence (Lorrain and White, 1971). With appropriate partial homogeneity assumptions, various regular equivalences (White and Reitz, 1983) can be obtained in the same way. It is beyond the scope of this introduction to show this formally but centrality rankings and role assignments thus turn out to be two aspects of the same concept, the comparison of positions.

Moreover, if positions are not only ordered by dominance but compared by appropriate indicators of degree of similarity, which is the guiding principle in Burt (1976), existing and novel methods for cohesion and community detection can be based on network positions as well.

We conclude that the approach via network positions has the potential to unify most of the network-analytic toolbox. Since existing methods can be reformulated to fit in the framework, it results in an extension rather than an alternative. The generic formulation facilitates more general results as exemplified by the multiscenred painter network and the simultaneous reasoning about an entire class of closeness centrality indices. It also points to the building blocks that can be adapted to domain-specific requirements, and, most importantly, it breaks down the process of analysis into steps with well-defined intermediate results. While these steps are more amenable to justification from theory, the intermediate results are easier to curate with qualitative data and to validate empirically. With the examples presented here, we have only scratched the surface.

Acknowledgments

The ideas in this article have been presented on various occasions over the past two years during which the author has benefitted greatly from comments and encouragement by valued colleagues. The author would like to thank in particular Russ Bernard, Ron Breiger, Martin Everett, Thomas Friemel, Betina Hollstein, Sven Kosub, Philip Leifeld, Peter Marsden, José Luis Molina, Volker Schneider, Tom Snijders, Thomas Weitin, and the anonymous reviewers.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This research is supported by Deutsche Forschungsgemeinschaft under grant Br 2158/6-1 and the European Research Council under grant no. 319209.

Notes

1. My version is from Lin Freeman’s collection of data sets.
2. The author is grateful to the reviewer who pointed this out.
References

Batagelj V (1994) Semirings for social network analysis. *Journal of Mathematical Sociology* 19(1): 53–68.
Beauchamp MA (1965) An improved index of centrality. *Behavioral Science* 10: 161–163.
Benzi M and Klymko C (2015) On the limiting behavior of parameter-dependent network centrality measures. *SIAM Journal on Matrix Analysis and Applications* 36(2): 686–706.
Blau PM (1977a) A macrosociological theory of social structure. *American Journal of Sociology* 83(1): 26–54.
Blau PM (1977b) Inequality and Heterogeneity: *A Primitive Theory of Social Structure*. New York: The Free Press.
Boldi P and Vigna S (2014) Axioms for centrality. *Internet Mathematics* 10(3–4): 222–262.
Bonacich P (1972) Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology* 2: 113–120.
Borgatti SP and Foster PC (2003) The network paradigm in organizational research: A review and typology. *Journal of Management* 29(6): 991–1013.
Borgatti SP and Halgin DS (2011) On network theory. *Organization Science* 22(5): 1359–1367.
Borgatti SP, Brass DJ and Halgin DS (2014) Social network research: Confusions, criticisms, and controversies. In: Brass DJ, Labianca G, Mehra A, et al. (eds) *Contemporary Perspectives on Organizational Social Networks* (Research in the Sociology of Organizations, vol. 40). Bingley: Emerald Publishing, pp. 1–29.
Borgatti SP, Everett MG and Johnson JC (2013) Analyzing *Social Networks*. London: SAGE.
Borgatti SP, Mehra A, Brass DJ, et al. (2009) Network analysis in the social sciences. *Science* 323(5916): 892–895.
Brandes U, Borgatti SP and Freeman LC (2016) Maintaining the duality of closeness and betweenness centrality. *Social Networks* 44: 153–159.
Brandes U, Freeman LC and Wagner D (2013) Social networks. In: Tamassia R (ed.) *Handbook of Graph Drawing and Visualization*. Boca Raton, FL: CRC Press, pp. 805–839.
Brandes U, Robins G, McCranie A, et al. (2013) What is network science? *Network Science* 1(1): 1–15.
Breiger RL and Pattison PE (1986) Cumulated social roles: The duality of persons and their algebras. *Social Networks* 8(3): 215–256.
Breiger RL., Boorman SA and Arabie P (1975) An algorithm for clustering relational data with applications to social network analysis and comparison with multidimensional scaling. *Journal of Mathematical Psychology* 12: 328–383.
Breiger RL, Schoon E, Melamed D, et al. (2014) Comparative configurational analysis as a two-mode network problem: A study of terrorist group engagement in the drug trade. *Social Networks* 36: 23–39.
Burt RS (1976) Positions in networks. *Social Forces* 55: 93–122.
Kilduff M and Oh H (2006) Deconstructing diffusion: An ethno-statistical examination of medical innovation network data analyses. *Organizational Research Methods* 9(4): 432–455.

Kishi G (1981) On centrality functions of a graph. In: Saito N and Nishizeki T (eds) *Graph Theory and Algorithms, Lecture Notes in Computer Science*. New York: Springer-Verlag, pp. 45–52.

Lin N (1976) *Foundations of Social Research*. New York: McGraw-Hill.

Lorrain F and White HC (1971) Structural equivalence of individuals in social networks. *Journal of Mathematical Sociology* 1(1): 49–80.

McPherson M (1983) An ecology of affiliation. *American Sociological Review* 48(4): 519–532.

Mayer K (2012) Objectifying social structures: Network visualization as means of social optimization. *Theory & Psychology* 22(2): 162–178.

Nadel SF (1957) *The Theory of Social Structure*. London: Cohen & West.

Opsahl T, Agneessens F and Skvoretz J (2010) Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks* 32(3): 245–251.

Ruhnau B (2000) Eigenvector-centrality—A node-centrality? *Social Networks* 22(4): 357–365.

Sabidussi G (1966) The centrality index of a graph. *Psychometrika* 31(4): 581–603.

Salancik GR (1995) WANTED: A good network theory of organization. *Administrative Science Quarterly* 40(2): 345–349.

Schaeffer SE (2007) Graph clustering. *Computer Science Review* 1(1): 27–64.

Scott J (2013) *Social Network Analysis: A Handbook* (3rd edn). London: SAGE.

Valente TW (2011) Medical innovation study. In: Barnett GA (ed.) *Encyclopedia of Social Networks*. London: SAGE, pp. 525–527.

Valente TW and Foreman RK (1998) Integration and radiality: Measuring the extent of an individual’s connectedness and reachability in a network. *Social Networks* 20(1): 89–105.

Van den Bulte C and Lilien GL (2001) Medical innovation revisited: Social contagion versus marketing effort. *American Journal of Sociology* 106(5): 1409–1435.

von Hayek FA (1952) *The Counter-Revolution of Science: Studies on the Abuse of Reason*. New York: The Free Press.

Wasserman S and Faust K (1994) *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

White DR and Reitz KP (1983) Graph and semigroup homomorphisms on networks of relations. *Social Networks* 5: 193–234.

Ziegler R (1987) Positionen in sozialen Räumen. In: Pappi FU (ed) *Methoden der Netzwerkanalyse*. München: R. Oldenbourg, pp. 64–100.

**Author biography**

Ulrik Brandes is a professor for algorithmics at the University of Konstanz. He is an area editor with Network Science, associate editor with Social Networks, and on the editorial board of the Journal of Graph Algorithms and Applications. Since 2007, he serves on the board of directors of the International Network for Social Network Analysis (INSNA).