

Centrality Measures in Networks*

Francis Bloch[†] Matthew O. Jackson[‡] Pietro Tebaldi[§]

August 2016

Abstract

We show that although the prominent centrality measures in network analysis make use of different information about nodes' positions, they all process that information in an identical way: they all spring from a common family that are characterized by the same simple axioms. In particular, they are all based on a monotonic and additively separable treatment of a statistic that captures a node's position in the network.

JEL CLASSIFICATION CODES: D85, D13, L14, O12, Z13, C65

KEYWORDS: Centrality, prestige, power, influence, networks, social networks, rankings, centrality measures

*We benefitted from comments by Jean-Jacques Herings, Gabrielle Demange, Alex Teytelboym, and seminar participants at Paris 1 and CTN Venice. We gratefully acknowledge financial support from the NSF grants SES-0961481 and SES-1155302, ARO MURI Award No. W911NF-12-1-0509, and from grant FA9550-12-1-0411 from the AFOSR and DARPA

[†]Université Paris 1 and Paris School of Economics, francis.bloch@univ-paris1.fr.

[‡]Department of Economics, Stanford University, external faculty of the Santa Fe Institute, and fellow of CIFAR, jacksonm@stanford.edu

[§]Stanford University, ptebaldi@stanford.edu

1 Introduction

The positions of individuals in a network drive a wide range of economic behaviors, from decisions concerning education and human capital (Hahn, Islam, Patacchini, and Zenou, 2015) to the identification of banks that are too-connected-to-fail (Gofman, 2015). Most importantly, there are many different ways to capture a person’s centrality, power, prestige, or influence. As such concepts depend heavily on context, which measure is most appropriate may vary with the application. Betweenness centrality is instrumental in explaining the rise of the Medici (Padgett and Ansell, 1993; Jackson, 2008), while Katz-Bonacich centrality is critical in understanding social multipliers in criminal behavior (Ballester, Calvó-Armengol, and Zenou, 2006), diffusion centrality is important in understanding many diffusion processes (Banerjee, Chandrasekhar, 2013), eigenvector centrality determines whether a society correctly aggregates information (Golub and Jackson, 2010), and degree centrality helps us to understand systematic biases in social norms (Jackson, 2016) and who is first hit in a contagion (Christakis and Fowler, 2010). For example, the power of an agent may depend upon the speed at which she receives or sends information to other agents, making agents with shortest distances to other agents most central. Or, the power of an agent may depend upon her ability to connect other agents and bridge between them, making agents who serve as intermediaries on the most paths connecting other agents most central. Or, the power of an agent may depend upon her connection to other powerful agents, making an agent highly central if she is surrounded by other central agents.

Despite the importance of network position in many settings, and the diversity of measures that have been proposed to capture different facets of centrality, very little is known about the properties that distinguish those measures. To address this, we axiomatize the standard centrality measures within a simple unified framework. Perhaps unexpectedly, we find that all of the standard measures of centrality are characterized by a simple and common set of axioms. They are all characterized by the same monotonicity, symmetry, and additivity axioms. This comes from the fact that they can all be written as an additively separable weighted average of a vector of statistics (or as a limit of such an expression). They differ solely in terms of the vectors of statistics that they process and not the manner in which they process that information.

In particular, we first note that standard centrality measures can be written as functions of what we call ‘nodal statistics’: vectors of data describing the position of a node in a social network. For instance, the ‘neighborhood’ statistic measures the number of nodes at each possible distance in the network from a given node - so lists how many neighbors it has, how many nodes at path distance two, three, etc. The ‘walk’ statistic measures the number of walks of different lengths originating from a given node; and the ‘intermediary’ statistic measures the number of shortest paths connecting other nodes in the network which

pass through a given node. Standard centrality measures are actually each based on one of just a few variations on nodal statistics. A centrality measure processes this information and produces a score for each node. In principle, one can imagine many different ways of processing such information. We show, however, that all standard centrality measures are characterized by axioms of monotonicity (higher statistics lead to higher centrality), symmetry (nodes' centralities only depend on their statistics and not their labels), and additivity (statistics are processed in an additively separable manner). Monotonicity and symmetry are reasonably weak axioms. However, additivity is a strong and narrow axiom, and yet it is one that has always been implicitly used when defining centrality measures.

With this perspective, the critical differences between the standard centrality measures boils down entirely to which nodal statistics they incorporate and not in the way in which they process that information. This suggests that there is ample room for new measures that process information about nodes' positions in novel ways - moving beyond the narrow class of additively separable measures.

Finally, in order to get a better grasp of the differences of centrality measures with respect to the network, we identify networks for which different centrality measures coincide, and networks for which they differ. We also compare network statistics on some simulated networks.

We discuss the associated literature that has characterized particular centrality measures later in the paper, once it becomes relevant and then can be compared to our axioms. The short summary is that our main contributions relative to the previous literature is to provide a comprehensive axiomatization of centrality measures via a common set of axioms, along with the introduction of the concept of nodal statistics; and to show that standard centrality measures differ only in terms of the network information that they take into account, and not how that information is processed.

2 Centrality measures in social networks

2.1 Background Definitions and Notation

We consider a network on n nodes indexed by $i \in \{1, 2, \dots, n\}$. A *network* is a graph, represented by its adjacency matrix $\mathbf{g} \in \{0, 1\}^{n \times n}$, where $g_{ij} = 1$ indicates the existence of an edge between nodes i and j and $g_{ij} = 0$ indicates the absence of an edge between the two nodes.

Our characterization results apply to both directed and undirected versions of networks, and also allow for weighted networks. To keep the presentation uncluttered, we discuss the definitions for the undirected case (and define basic graph terminology for that case), but

the results extend directly. We generally presume that $g_{ii} = 0$ for all i .

Let $G(n)$ denote the set of all admissible networks on n nodes.

The *degree* of a node i in a network g , denoted $d_i(\mathbf{g}) = \sum_j g_{ij}$, is the number of edges involving node i .

A *walk* between i and j is a succession of (not necessarily distinct) nodes $i = i^0, i^1, \dots, i^M = j$ such that $g_{i^m i^{m+1}} = 1$ for all $m = 0, \dots, M-1$. A *path* in g between two nodes i and j is a succession of distinct nodes $i = i^0, i^1, \dots, i^M = j$ such that $g_{i^m i^{m+1}} = 1$ for all $m = 0, \dots, M-1$. Two nodes i and j are connected (or path-connected) if there exists a path between them.

A *geodesic* (shortest path) between nodes i and j is a path such that no other path between them involves a smaller number of edges. The *distance* between nodes i and j , $\rho_{\mathbf{g}}(i, j)$ is the number of edges involved in a geodesic between i and j , which is defined only for pairs of nodes that have a path between them and may be taken to be ∞ otherwise. The number of geodesics between i and j is denoted $\nu_{\mathbf{g}}(i, j)$. We let $\nu_{\mathbf{g}}(k : i, j)$ denote the number of geodesics between i and j involving node k .

It is useful to note that the elements of the ℓ -th power of \mathbf{g} , denoted \mathbf{g}^ℓ , have a simple interpretation: g_{ij}^ℓ counts the number of walks of length ℓ between nodes i and j .

We let $n_i^\ell(\mathbf{g})$ denote the number of nodes at distance ℓ from i in network g : $n_i^\ell(\mathbf{g}) = |\{j : \rho_{\mathbf{g}}(i, j) = \ell\}|$.

A *tree* is a graph such that for any two nodes i, j there is a unique path between i and j . A tree can be oriented by selecting one node i^0 (the root) and constructing a binary relation \succ^d as follows: For all nodes such that $g_{i^0 i} = 1$, set $i^0 \succ^d i$. Next, for each pair of nodes i and j that are distinct from i^0 , say that $i \succ^d j$ if that $g_{ij} = 1$ and the geodesic from i to i^0 is shorter than the geodesic from j to i^0 . If $i \succ^d j$, then i is called the *direct predecessor* of j and j is called a *direct successor* of i . The transitive closure of the binary relation \succ^d defines a partial order \succ , where if $i \succ j$ then we say that i is a predecessor of j and j a successor of i , in the oriented tree.

Let $\lambda^{\max}(\mathbf{g})$ denote the largest right-hand-side eigenvalue of \mathbf{g} .

2.2 Centrality measures

A centrality measure is a function $\mathbf{c} : G(n) \rightarrow \mathbb{R}^n$, where $c_i(\mathbf{g})$ is the centrality of node i in the social network \mathbf{g} . Here are some of the key centrality measures from the literature.¹

Degree centrality Degree centrality measures the number of edges attached to node i , $d_i(\mathbf{g})$. We can also normalize by the maximal possible degree, $n - 1$, to obtain a number

¹For more background on centrality measures see Borgatti (2005), Wasserman and Faust (1994, Chapter 4) and Jackson (2008, Chapter 2.2).

between 0 and 1:

$$c_i^{deg}(\mathbf{g}) = \frac{d_i(\mathbf{g})}{n-1}.$$

Degree centrality is an obvious centrality measure, and gives some insight into the connectivity or ‘popularity’ of node i , but misses potentially important aspects of the architecture of the network and a node’s position in the network.

Closeness centrality Closeness centrality is based on the network distance between a node and each other node. It extends degree centrality by looking at neighborhoods of all radii. The input into measures of closeness centrality is the list of distances between node i and other nodes j in the network, $\rho_{\mathbf{g}}(i, j)$. There are different variations of closeness centrality based on different functional forms. The measure proposed by Bavelas (1950) and Sabidussi (1966), is based on distances between node i and all other nodes, $\sum_j \rho_{\mathbf{g}}(i, j)$. In that measure a higher score indicates a lower centrality. To deal with this inversion, and also to deal with the fact that this distance becomes infinite if nodes belong to two different components, Sabidussi (1966) proposed to a centrality measure of $\frac{1}{\sum_j \rho_{\mathbf{g}}(i, j)}$. One can also normalize that measure so that the highest possible centrality measure is equal to 1, to obtain the closeness centrality measure,

$$c_i^{cls}(\mathbf{g}) = \frac{n-1}{\sum_{j \neq i} \rho_{\mathbf{g}}(i, j)}.$$

An alternative measure of closeness centrality (e.g., see Rochat (2009); Garg (2009)) , aggregates distances differently. It aggregates the sum of all inverses of distances, $\sum_j \frac{1}{\rho_{\mathbf{g}}(i, j)}$. This avoids having a few nodes for which there is a large or infinite distance drive the measurement. This measure can also be normalized so that spans from 0 and 1, and one obtains

$$c_i^{cl}(\mathbf{g}) = \frac{\sum_{\ell} \frac{1}{\ell} |\{j : \rho_{\mathbf{g}}(i, j) = \ell\}|}{n-1} = \frac{1}{n-1} \sum_{j \neq i} \frac{1}{\rho_{\mathbf{g}}(i, j)}.$$

Decay centrality Decay centrality proposed by Jackson (2008) is a measure of distance that takes into account the decay in traveling along shortest paths in the network. It reflects the fact that information traveling along paths in the network may be transmitted stochastically, or that other values or effects transmitted along paths in the network may decay, according to a parameter δ . Decay centrality is defined as

$$c_i^{\delta}(\mathbf{g}) = \sum_{\ell \leq n-1} \delta^{\ell} n_i^{\ell}(\mathbf{g}).$$

As δ goes to 1, decay centrality measures the size of the component in which node i lies. As δ goes to 0, decay centrality becomes proportional to degree centrality.

Katz-Bonacich centrality Katz (1953) and Bonacich (1972, 1987) proposed a measure of prestige or centrality based on the number of walks emanating from a node i . Because the length of walks in a graph is unbounded, Katz-Bonacich centrality requires a discount factor – a factor δ between 0 and 1 – to compute the discounted sum of walks emanating from the node. Walks of shorter length are evaluated at an exponentially higher value than walks of longer length. In particular, the centrality score for node i is based on counting the total number of walks from it to other nodes, each exponentially discounted based on their length:

$$c_i^{KB}(\mathbf{g}, \delta) = \sum_{\ell} \delta^{\ell} \sum_j g_{ij}^{\ell}.$$

In matrix terms (when $\mathbf{I} - \delta\mathbf{g}$ inverts):^{2,3} $\mathbf{c}^{KB}(\mathbf{g}, \delta) = \sum_{\ell=1}^{\infty} \delta^{\ell} \mathbf{g}^{\ell} \mathbf{1} = (\mathbf{I} - \delta\mathbf{g})^{-1} \delta\mathbf{g} \mathbf{1}$.

Eigenvector centrality Eigenvector centrality, proposed by Bonacich (1972), is a related measure of prestige. It relies on the idea that the prestige of node i is related to the prestige of her neighbors. Eigenvector centrality is computed by assuming that the centrality of node i is proportional to the sum of centrality of node i 's neighbors: $\lambda c_i = \sum_j g_{ij} c_j$, where λ is a positive proportionality factor. In matrix terms, $\lambda \mathbf{c} = \mathbf{g} \mathbf{c}$. The vector $c_i^{eig}(\mathbf{g})$ is thus the right-hand-side eigenvector of \mathbf{g} associated with the eigenvalue $\lambda^{\max}(\mathbf{g})$.⁴

The eigenvector centrality of a node is thus self-referential, but has a well-defined fixed point. This notion of centrality is closely related to ways in which scientific journals are ranked based on citations, and also relates to influence in social learning.

Diffusion centrality Diffusion centrality, proposed by Banerjee, Chandrasekhar, Duflo, and Jackson (2013),⁵ a dynamic contagion process starting at node i . In period 1, every neighbor of i

² $\mathbf{1}$ denotes the n -dimensional vector of 1s, and \mathbf{I} is the identity matrix. Invertibility holds for small enough δ .

³ In a variation proposed by Bonacich there is a second parameter β that rescales: $\mathbf{c}^{KB}(\mathbf{g}, \delta, \eta) = (\mathbf{I} - \delta\mathbf{g})^{-1} \beta \mathbf{g} \mathbf{1}$. Since the scaling is inconsequential, we ignore it.

⁴ $\lambda^{\max}(\mathbf{g})$ is positive when \mathbf{g} is nonzero (recalling that it is a nonnegative matrix), the associated vector is nonnegative, and for a connected network the associated eigenvector is positive and unique up to a rescaling (e.g., by the Perron-Frobenius Theorem).

⁵ This is related in spirit to basic epidemiological models (e.g. see Bailey (1975)), as well as the cascade model of Kempe, Kleinberg, and Tardos (2003) that allowed for thresholds of adoption (so that an agent cares about how many neighbors have adopted). The cascade model leads to a centrality measure introduced by Lim, Ozdaglar, and Teytelboym (2015) called cascade centrality. As Kempe et al. (2003) show, their model with thresholds is equivalent to a model without thresholds, provided that probabilities of transmission can depend on the target node and on its degree. Thus, that class of centralities nest what was defined as communication centrality by Banerjee et al. (2013) and is closely related to the decay centrality of Jackson (2008). Diffusion centrality differs from these other measures in that it is based on walks rather than paths, which make it easier to relate to Katz-Bonacich centrality and eigenvector centrality as discussed in Banerjee et al. (2013) and formally shown in Banerjee, Chandrasekhar, Duflo, and Jackson

is contacted with independent probability δ . In period $\ell = 2$, neighbors of nodes contacted at period $\ell = 1$ are contacted with independent probability δ . In any arbitrary period ℓ , neighbors of nodes contacted at period $\ell - 1$ are contacted with independent probability p . At period L , the expected number of times that agents have been contacted is computed using the number of walks

$$c_i^{dif}(\mathbf{g}, \delta, L) = \sum_{\ell=1}^L \sum_j \delta^\ell g_{ij}^\ell.$$

In matrix terms, $\mathbf{c}^{dif}(\mathbf{g}, \delta, L) = \sum_{\ell=1}^L \delta^\ell \mathbf{g}^\ell \mathbf{1}$.

If $L = 1$, diffusion centrality is proportional to degree centrality. As $L \rightarrow \infty$, c_i^{dif} converges to Katz-Bonacich centrality whenever δ is smaller than the inverse of the largest eigenvalue, $1/\lambda^{\max}(\mathbf{g})$. Banerjee, Chandrasekhar, Duflo, and Jackson (2013, 2014) show that diffusion centrality converges to eigenvector centrality as L grows whenever δ is larger than the inverse of the largest eigenvalue, $1/\lambda^{\max}(\mathbf{g})$.

Betweenness centrality Freeman’s betweenness centrality measures the importance of a node in connecting other nodes in the network. It considers all geodesics between two nodes j, k different from i which pass through i . Betweenness centrality thus captures the role of an agent as an intermediary in the transmission of information or resources between other agents in the network. As there may be multiple geodesics connecting j and k , we need to keep track of the fraction of geodesic paths passing through i , $\frac{\nu_{\mathbf{g}}(i:j,k)}{\nu_{\mathbf{g}}(j,k)}$. The betweenness centrality measure proposed by Freeman (1977) is

$$c_i^{bet}(\mathbf{g}) = \frac{2}{(n-1)(n-2)} \sum_{(j,k), j \neq i, k \neq i} \frac{\nu_{\mathbf{g}}(i:j,k)}{\nu_{\mathbf{g}}(j,k)}.$$

There are other variations that one can consider. For example, in a setting where intermediaries connect buyers and sellers in a network, the number of intermediaries on a geodesic matters, as intermediaries must share surplus along the path. In that case, it is useful to consider a variation on betweenness centrality where the length of the geodesic paths between any two nodes j and k is taken into account.

(2014). Nonetheless, diffusion centrality can be thought of as a representative of a whole class of related measures that are built on the same premise of seeing how much diffusion one gets from various nodes, but with variations in specifics of how the process is modeled. These sorts of measures, along with others, are used as inputs into other measures such as that of Kermani et al. (2015), which then combine information from a variety of centrality measures.

3 Axiomatization of centrality measures

In this section, we provide axiomatizations of centrality measures based on a notion of ‘nodal statistics’ - vectors of data capturing a facet of the position of a node in the network – as well as an aggregator that transforms those vectors of data into scalars. We first introduce nodal statistics, and then discuss axioms to characterize centrality measures.

3.1 Nodal statistics

A *nodal statistic*, $s_i(\mathbf{g})$, is a vector of data describing the position of node i in the network \mathbf{g} . These lie in some Euclidean space, \mathbb{R}^L , where L may be finite or infinite. We presume that the vector of all 0’s (usually an isolated node, or an empty network) is a feasible statistic.

Because networks are complex objects, nodal statistics are useful, as they allow an analyst to reduce the complexity of a network into a (small) vector of data. Different nodal statistics capture different aspects of a node’s position in a network.

Standard centrality measures use nodal statistics that pay attention only to the network and not on the identity of the nodes, as captured in the following property.

For a permutation π of $\{1, \dots, n\}$, let $\mathbf{g} \circ \pi$ be defined by $(\mathbf{g} \circ \pi)_{ij} = g_{\pi(i)\pi(j)}$

DEFINITION 1 *A nodal statistic is symmetric if for any permutation π of $\{1, \dots, n\}$, $s_i(\mathbf{g}) = s_{\pi(i)}(\mathbf{g} \circ \pi)$.*

Several nodal statistics are fundamental.

The neighborhood statistic, $n_i(\mathbf{g}) = (n_i^1(\mathbf{g}), \dots, n_i^\ell(\mathbf{g}), \dots, n_i^{n-1}(\mathbf{g}))$, is a vector counting the number of nodes at path-distance $\ell = 1, 2, \dots, n - 1$ from a given node i .

The neighborhood statistic measures how quickly (in terms of path length) node i can reach the other nodes in the network.

The degree statistic, $d_i(\mathbf{g}) = n_i^1(\mathbf{g})$, counts the connections of a given node i .

This is a truncated version of the neighborhood statistic.

The closeness statistic, $cl_i(\mathbf{g}) = (cl_i^1(\mathbf{g}), \dots, cl_i^\ell(\mathbf{g}), \dots, cl_i^{n-1}(\mathbf{g}))$, is the vector such that $cl_i^\ell(\mathbf{g}) = \frac{n_i^\ell(\mathbf{g})}{\ell}$ for each $\ell = 1, 2, \dots, n - 1$, tracking nodes at different distances from a given node i , weighted by the inverse of those distances.

The walk statistic, $w_i(\mathbf{g}) = (w_i^1(\mathbf{g}), \dots, w_i^\ell(\mathbf{g}), \dots)$, is an infinite vector counting the number of walks of length $\ell = 1, 2, \dots$ emanating from a given node i . Using the connection between number of walks and iterates of the adjacency matrix, $w_i^\ell(\mathbf{g}) = \sum_j (g^\ell)_{i,j}$.

The main difference from the distance statistic is that it keeps track of multiplicities of routes between nodes and not just shortest paths, and thus is useful in capturing processes that may involve random transmission.

The intermediary statistic, $I_i(\mathbf{g}) = (I_i^1(\mathbf{g}), \dots, I_i^\ell(\mathbf{g}), \dots, I_i^{n-1}(\mathbf{g}))$, is a vector counting the normalized number of geodesics of length $\ell = 1, 2, \dots$ which contain node i . For any pair j, k of nodes different from i , the normalized number of geodesic paths between i and j containing i is given by the proportion of geodesics passing through i , $\frac{\nu_{\mathbf{g}}(i:j,k)}{\nu_{\mathbf{g}}(j,k)}$. Summing across over all pairs of nodes j, k different from i who are at distance ℓ from each other:

$$I_i^\ell = \sum_{j,k: \rho_{\mathbf{g}}(j,k)=\ell, j \neq i, k \neq i} \frac{\nu_{\mathbf{g}}(i:j,k)}{\nu_{\mathbf{g}}(j,k)}.$$

The intermediary statistic measures how important node i is in connecting other agents in the network.

It is often useful to compare nodal statistics. There are various ways to compare them depending on the application, but given that the applications all involve vectors in some Euclidean space, the default is to use the Euclidean partial order, so that $s \geq s'$ if it is at least as large in every entry. We discuss another method of comparison based on stochastic dominance below.

3.2 Axioms and Characterizations

We now present the axioms that characterize all of the families of centrality measures that we have discussed above - i.e., the canonical measures from the literature. Thus, if a researcher finds one of these properties objectionable, it suggests that new definitions of centrality are needed.

We first note that two elementary axioms guarantee that the centrality of a node i depends only on the information contained in the nodal statistic s_i .

AXIOM 1 (Monotonicity) *c is monotonic relative to a nodal statistic s_i if $s_i(\mathbf{g}) \geq s_i(\mathbf{g}')$ implies that $c_i(\mathbf{g}) \geq c_i(\mathbf{g}')$.*

Monotonicity connects a centrality measure with a nodal statistic: a node i is at least as central in social network g than in social network \mathbf{g}' if the nodal statistic of i at g is at least as large as the nodal statistic of i at \mathbf{g}' .

AXIOM 2 (Symmetry) *c is symmetric if for any bijection π on $\{1, \dots, n\}$ and all i : $c_i(\mathbf{g}) = c_{\pi(i)}(\mathbf{g} \circ \pi)$.*

Symmetry guarantees that a centrality measure does not depend on the identity of a node.

Both of these axioms are satisfied by all standard centrality measures, and so are useful in characterizing them.

Under monotonicity and symmetry, the centrality of node i only depends only on a statistic s_i , as shown by the following lemma.

Let us say that a function $\mathcal{C} : \mathbb{R}^L \rightarrow \mathbb{R}$ is monotone if $\mathcal{C}(s) \geq \mathcal{C}(s')$ whenever $s \geq s'$.

We say that a centrality measure is *representable* relative to a nodal statistic s of dimension L if there exists a function $\mathcal{C} : \mathbb{R}^L \rightarrow \mathbb{R}$, for which $c_i(\mathbf{g}) = \mathcal{C}(s_i(\mathbf{g}))$ for all i and \mathbf{g} .

LEMMA 1 *A centrality measure c is symmetric and satisfies monotonicity relative to some nodal statistic s if, and only if, s is symmetric and c is representable relative to s by a monotone \mathcal{C} .*

Next, we show that the following axioms then result in a comprehensive axiomatization of the wide class of centrality measures considered here.

In what follows, we normalize centrality measures so that $c_i(\mathbf{g}) = \mathcal{C}(s_i(\mathbf{g})) = 0$ if $s_i(\mathbf{g}) = \mathbf{0}$. This is without loss of generality, as any centrality measure can be so normalized simply by subtracting off this value everywhere.

Given that standard centrality measures all satisfy symmetry and monotonicity, we can work with centrality measures as a function of the statistics.

Before moving to our main characterization, we build up the characterization by showing the classes of centrality measures that satisfy increasingly strong axioms on how they aggregate nodal statistics.

Such an aggregation problem has some parallels to constructing a utility index for a stream of intertemporal consumptions, as studied by Koopmans (1960) and Debreu (1960). Some axioms from that literature have parallels here.

AXIOM 3 (Independence) *A function $\mathcal{C} : \mathbb{R}^L \rightarrow \mathbb{R}$ satisfies independence if*

$$\mathcal{C}(s_i) - \mathcal{C}(s'_i) = \mathcal{C}(s''_i) - \mathcal{C}(s'''_i).$$

for any s_i, s'_i, s''_i , and s'''_i - all in \mathbb{R}^L - such that $s_{i,\ell} = s'_{i,\ell} = a$ for some ℓ , and we equally change that entry for both statistics without changing any other entries: so $s''_{i,\ell} = s'''_{i,\ell} = b$, but $s''_{i,k} = s_{i,k}$ and $s'''_{i,k} = s'_{i,k}$ for all $k \neq \ell$.

Independence requires that, whenever a component of two nodal statistics are equal, the difference in centrality across those two nodal statistics does not depend on the level of that component. Independence is weaker than additivity and so satisfied by all standard centrality measures and implies that a centrality measure is an additively separable function of the elements of the nodal statistic.

THEOREM 1 *A centrality measure c is representable relative to a nodal statistic s by a monotone \mathcal{C} (Lemma 1) that satisfies independence, if, and only if, s is symmetric and there exist a set of monotone functions $F^\ell : \mathbb{N} \rightarrow \mathbb{R}$ with $F^\ell(0) = 0$, such that*

$$c_i(g) = \mathcal{C}(s_i(g)) = \sum_{\ell=1}^L F^\ell(s_i^\ell(g)). \quad (1)$$

Although Theorem 1 shows that independence (together with monotonicity and symmetry) implies that a central measure is additively separable, it provides no information on the specific shape of the functions F^ℓ . Additional axioms tie down the functional forms.

A recursive axiom, in the spirit of an axiom from Koopmans (1960), implies that a centrality measure has an exponential decay aspect to it.

AXIOM 4 (Recursivity) *A function $\mathcal{C} : \mathbb{R}^L \rightarrow \mathbb{R}$ is recursive if for any s_i, s'_i (which are L dimensional⁶)*

$$\frac{\mathcal{C}(s_i) - \mathcal{C}(s_{i,1}, \dots, s_{i,\ell}, 0, \dots, 0)}{\mathcal{C}(s'_i) - \mathcal{C}(s'_{i,1}, \dots, s'_{i,\ell}, 0, \dots, 0)} = \frac{\mathcal{C}(s_{i,\ell+1}, \dots, s_{i,L}, 0, \dots, 0)}{\mathcal{C}(s'_{i,\ell+1}, \dots, s'_{i,L}, 0, \dots, 0)}.$$

Recursivity requires that the calculation being done based on later stages of the nodal statistic look ‘similar’ (in a ratio sense) to those done earlier in the nodal statistic. Again, standard centrality measures are recursive. Recursivity implies that the functions F^ℓ have a decay structure $F^\ell = \delta^{\ell-1}f$, for some increasing f and $\delta > 0$:

THEOREM 2 *A centrality measure c is representable relative to a nodal statistic s by a monotone $\mathcal{C} : \mathbb{R}^L \rightarrow \mathbb{R}$ (Lemma 1) that is recursive and satisfies independence, if, and only if, s is symmetric and there exist an increasing function $f : \mathbb{N} \rightarrow \mathbb{R}$ with $f(0) = 0$, and $\delta \geq 0$ such that*

$$c_i(g) = \mathcal{C}(s_i(g)) = \sum_{\ell=1}^L \delta^{\ell-1} f(s_i^\ell(g)). \quad (2)$$

Recursivity ties down that each dimension of the nodal statistic is processed according to the same monotone function f , and the only difference in how they enter the centrality measure is in how they are weighted - which is according to an exponential decay function.

Next, we show that an additivity axiom provides a complete characterization that encompasses all the standard centrality measures, and results in our main characterization.

AXIOM 5 (Additivity) *A function $\mathcal{C} : \mathbb{R}^L \rightarrow \mathbb{R}$ is additive if for any s_i and s'_i in \mathbb{R}^L :*

$$\mathcal{C}(s_i + s'_i) = \mathcal{C}(s_i) + \mathcal{C}(s'_i).$$

⁶If $L = \infty$, then when writing $s_{i,\ell+1}, \dots, s_{i,L}, 0, \dots, 0$ below simply ignore the trailing 0’s.

Additivity is another axiom that is generally satisfied by centrality measures. It clearly implies independence,⁷ and results in the following characterization.

THEOREM 3 *A centrality measure c is representable relative to a nodal statistic s by a monotone $\mathcal{C} : \mathbb{R}^L \rightarrow \mathbb{R}$ (Lemma 1) that is recursive and additive, if, and only if, s is symmetric and there exists $\delta \geq 0$ such that*

$$c_i(g) = \mathcal{C}(s_i(g)) = \sum_{\ell=1}^L \delta^{\ell-1} s_i^\ell. \quad (3)$$

Theorem 3 shows that monotonicity, symmetry, recursivity, and additivity, completely tie down that a centrality measure must be the discounted sum of successive elements of the nodal statistic. As a corollary, all centrality measures which can be expressed as discounted sums of some nodal statistic can be characterized by the three axioms of monotonicity, symmetry, recursivity and additivity.

COROLLARY 1 *Consider a centrality measure c that is representable relative to a nodal statistic by a monotone $\mathcal{C} : \mathbb{R}^L \rightarrow \mathbb{R}$ (Lemma 1) that is recursive and additive.*

- *If the nodal statistic is the degree statistic, d , then the centrality measure is degree centrality.*
- *If the nodal statistic is the neighborhood statistic, n_i , then the centrality measure is decay centrality.*
- *If the nodal statistic is the (infinite) walk statistic, w_i , then the centrality measure is Katz-Bonacich centrality.*
- *If the nodal statistic is the walk statistic, $(w_i^\ell)_{\ell \leq L}$, restricted to the first $L < \infty$ elements, then the centrality measure is diffusion centrality (which is proportional to degree centrality if $L = 1$).*
- *If the nodal statistic is the intermediary statistic, I_i , then the centrality measure is betweenness centrality.*
- *If the nodal statistic is the closeness statistic, cl_i , then the centrality measure is closeness centrality.*

⁷It also implies monotonicity, but since we use monotonicity to establish the aggregator function on which additivity is stated, we maintain it as a separate condition in the statement of the theorem.

3.3 Eigenvector Centrality

Although eigenvector centrality is the one prominent centrality measure not in the corollary, it is still captured as a limiting case. In particular, note that if the discount factor δ is larger than the inverse of the largest eigenvector of the adjacency matrix, then there exists \bar{L} large enough such that the ranking generated by diffusion centrality will remain the same for any $L \geq \bar{L}$. In fact, for large enough L , diffusion centrality converges to eigenvector centrality if the discount factor δ is larger than the inverse of the largest eigenvector of the adjacency matrix (see Banerjee et al. (2013, 2014)). Thus, eigenvector centrality is simply the limit of a centrality measure that satisfies the axioms and has a representation of the form in Theorem 3.

3.4 Ordering Neighborhood Statistics

The ordering that we used to compare nodal statistics in defining monotonicity conditions above was based on the Euclidean partial order. That was all that was needed to derive the characterizations above.

In some cases it is useful to also compare nodal statistics using other partial orders that may make orderings in cases in which the Euclidean ordering does not. For example, because the total number of nodes in a connected network is fixed, and the neighborhood statistic measures the *distribution* of nodes at different distances in the network, the statistics of two different nodes in a network according to the neighborhood statistic will not be comparable via the Euclidean partial ordering unless they are equal.

Thus, a natural partial order to associate with the neighborhood statistic is based on *first order stochastic dominance*.⁸ We say that $s_i \succeq s'_i$ if for all t , $\sum_{\ell=1}^t s_i^\ell \geq \sum_{\ell=1}^t s'^\ell_i$. This induces a strict version: $s_i \succ s'_i$ if $s_i \succeq s'_i$ and $s'_i \not\succeq s_i$. In other words, a statistic s_i dominates s'_i if, for any distance t , the number of nodes at distance less than t under s_i is at least the number of nodes at distance less than t in s'_i .

It is then easy to check, that if $s_i \succ s'_i$ and $\delta < 1$ in an additive, symmetric and monotone centrality measure, then it also follows that $C(s_i) > C(s'_i)$. This applies to many of the measures we have defined, as they are additive, symmetric, and monotone.

3.5 Some Related Literature

A main contribution of our work is to show that standard centrality measures can be viewed as being based on the same logical structure, varying only which statistics are paid attention

⁸The entries of the s 's may not sum to one, so this is not always a form of *stochastic dominance*, but it is defined analogously.

to when aggregating. This distinguishes our work from the previous literature that has focused on specific centrality measures. Let us discuss some of the previous characterizations of various measures.

Centrality measures are related to other ranking problems. For example, ranking problems have been considered in the contexts of tournaments (Laslier, 1997), citations across journals (Palacios-Huerta and Volij, 2004), and hyperlinks between webpages (Page, Brin, Motwani, and W 1998). There is a literature in social choice devoted to the axiomatization of ranking methods. For example, the Copeland rule (by which agents are ranked according to their count of wins in a tournament), – the equivalent of degree centrality in our setting – has been axiomatized by Rubinstein (1980), Henriot (1985), and van den Brink and Gilles (2003). Palacios-Huerta and Volij (2004) axiomatize the invariant solution – an eigenvector-based measure on a modified matrix normalized by the number of citations. Their axiomatization relies on global properties – anonymity, invariance with respect to citations intensity. It then introduces an axiom characterizing the solution for 2×2 matrices (weak homogeneity) and a specific definition of reduced games, which together with a consistency axiom, allows to extend the solution in 2×2 games to general matrices. Slutzki and Volij (2006) propose an alternative axiomatization of the invariant solution, replacing weak homogeneity and consistency by a weak additivity axiom. They characterize the invariant solution as the only solution satisfying weak additivity, uniformity and invariance with respect to citations intensity. Slutzki and Volij (2006) axiomatize a different eigenvector centrality measure – the fair bets solution. The fair bets solution is the only solution satisfying uniformity, inverse proportionality to losses and neutrality.

Dequiedt and Zenou (2014) recently proposed an axiomatization of prestige network centrality measures, departing from the axioms of Palacios-Huerta and Volij (2004) in several directions. As in Palacios-Huerta and Volij (2004), their axiomatization relies on the characterization of the solution in simple situations (in this case stars) and the definition of a reduced problem such that consistency extends the solution from the simple situation to the entire class of problems. The reduced game is defined using the concept of an “embedded network”: a collection of nodes partitioned into two groups - one group where a value is already attached to the node (terminal nodes) and one group where values still have to be determined (regular nodes). One axiom used is a normalization axiom. Two axioms are used to determine the solution in the star – the linearity and additivity axioms. Consistency is then applied to generate a unique solution – the Katz Bonacich centrality measure with an arbitrary parameter a . Replacing linearity and additivity by invariance, Dequiedt and Zenou (2014) obtain a different solution in the star network, which extends by consistency to degree centrality for general situations. Eigenvector centrality can also be axiomatize using a different set of axioms on the star network, and adding a converse consistency axiom.

Garg (2009)⁹ proposed different sets of axioms to characterize each of degree, decay and closeness centralities. To axiomatize degree centrality, he uses an additivity axiom across subgraphs – a much stronger requirement than that discussed here, which makes the measure independent of the structure of neighborhoods at distance greater than one. In order to axiomatize decay and degree centrality, Garg uses an axiom which amounts to assuming that the only relevant information in the network is the distance statistic. The ”breadth first search” axiom assumes that centrality measures are identical whenever two graphs generate the same reach statistics for all nodes. A specific axiom of closeness pins down the functional form of the additively separable functions so that the closeness centrality measure is obtained. In order to characterize decay centrality, Garg uses another axiom which pins down a specific functional form, termed the up-closure axiom.

4 Comparisons of Centrality Measures

In this section, we compare centrality measures both theoretically and via some simulations.

4.1 Comparing Centrality Measures on Trees

We first focus attention on trees and characterize the class of trees for which all neighborhood-based centrality measures coincide, and all neighborhood-, intermediary- and walk- based centrality measures coincide.¹⁰

4.1.1 Monotone Hierarchies

We define a class of trees that we call *monotone hierarchies*. A tree \mathbf{g} is a monotone hierarchy if there exists a node i_0 (the root) such that the oriented tree starting at i_0 satisfies the following conditions:

- For any two nodes i, j , if the distance between the root and i , $\rho(i)$, is smaller than the distance between the root and j , $\rho(j)$, then $d_i \geq d_j$
- For any two nodes i, j such that $\rho(i) = \rho(j)$, if $d_i > d_j$, then $d_k \geq d_l$ for any successor k of i and any successor l of j such that $\rho(k) = \rho(l)$.

⁹Garg’s paper was never completed, and so the axiomatizations are not full characterizations and/or are without proof. Nonetheless some of the axioms in his paper are of interest.

¹⁰ For networks that are *not* trees, the characterization of all networks for which centrality measures coincide is an open problem. König, Tessone, and Zenou (2014) prove that degree, closeness, betweenness and eigenvector centrality generate the same ranking on nodes for nested-split graphs, which are a very structured hierarchical form of network.

In a monotone hierarchy, nodes further from the root have a weakly smaller number of successors. Because leaf nodes have the smallest degree $d = 1$, all leaf nodes must be at the same distance from the root, and hence the depth of the hierarchy, denoted h , is the same following any path from the root. In a monotone hierarchy, different subgraphs may have different numbers of nodes. However, if at some point, a node i has a larger number of successors than a node j at the same level of the hierarchy, in the sub-tree starting from i , all nodes must have a (weakly) larger degree than nodes at the same level of the hierarchy in the sub-tree starting from j .

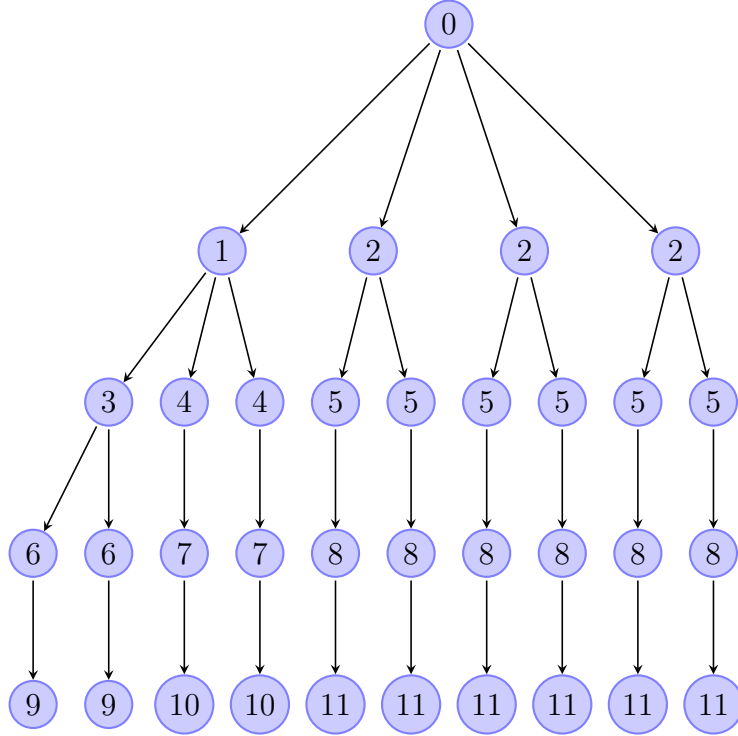


Figure 1: A monotone hierarchy

Figure 1 displays a monotone hierarchy, with numbers corresponding to the centrality ranking of nodes according to the neighborhood statistic. The most central node is node 0, followed by node 1, node 2, etc., with any node at level t being more central than a node at level $t + 1$. For two nodes at the same level, one is more central than another if it has larger degree, or if a predecessor or successor has larger degree.¹¹ More generally, define the following ranking on nodes in a monotone hierarchy:

1. If $\rho(i) < \rho(j)$, then $i \succ j$

¹¹Alternatively, we could consider all subtrees starting at direct successors of the root, and rank them according to their number of nodes. At any level t of the hierarchy, nodes would be ranked according to the ranking of subtrees.

2. If $\rho(i) = \rho(j)$ and $d(i) > d(j)$ then $i \succ j$
3. If $\rho(i) = \rho(j)$, $d(i) = d(j)$ and there exist two successors k, l of i and j such that $\rho(k) = \rho(l)$ and $d(k) > d(l)$ then $i \succ j$.
4. If $\rho(i) = \rho(j)$, $d(i) = d(j)$ and there exist two distinct predecessors k, l of i, j such that $\rho(k) = \rho(l)$ and $d(k) > d(l)$ then $i \succ j$.

Note that, by the definition of a monotone hierarchy, if there exist two predecessors k and l of i and j at the same level such that $d(k) > d(l)$, we cannot have $d(k') > d(l')$ for two successors k' and l' of i and j at the same level. Hence, for two nodes i and j at the same level of the hierarchy, the ranking \succ is well defined and allows for a complete ranking of all nodes (two nodes at the same level for which none of the conditions (2), (3), (4) hold have the same centrality). To check that the ranking is transitive, notice that if $d(i) > d(j)$ and $d(j) \geq d(k)$ then $d(i) > d(k)$. A repeated use of this argument guarantees that if $i \succ j$ and $j \succ k$, then $i \succ k$. For two nodes at different levels of the hierarchy, condition (1) guarantees that the ranking is well defined. It is also clearly transitive for three nodes taken at different levels of the hierarchy.

Monotone hierarchies are the only trees for which all centrality measures defined by the neighborhood statistic coincide.

PROPOSITION 1 *In a monotone hierarchy, for any two nodes i, j , $i \succ j$ if and only if $n_i \succ n_j$ and $i \doteq j$ if and only if $n_i = n_j$. Conversely, if a tree with even diameter is not a monotone hierarchy, there exist two nodes i and j such that neither $n_i \succeq n_j$ nor $n_j \succeq n_i$.*

4.1.2 Regular Monotone Hierarchies

For all centrality measures based on other nodal statistics to coincide too, we need to consider a more restrictive class of trees, which we refer to as *regular monotone hierarchies*. A monotone hierarchy is a regular monotone hierarchy if there exists a root i_0 such that (i) all nodes at the same distance from the root have the same degree and (ii) the degree of nodes is decreasing, as the distance to the root increases. Formally, $d_i = d_j$ if $\rho(i) = \rho(j)$ and $d_i \geq d_j$ if $\rho(i) < \rho(j)$.

Cayley trees are regular monotone hierarchies. Stars and lines are regular monotone hierarchies. In a regular monotone hierarchy, all nodes at the same distance from the root are symmetric and hence have the same centrality. Centrality is highest for the root i_0 and decreases with the levels of the hierarchy. Figure 2 illustrates a regular monotone hierarchy.

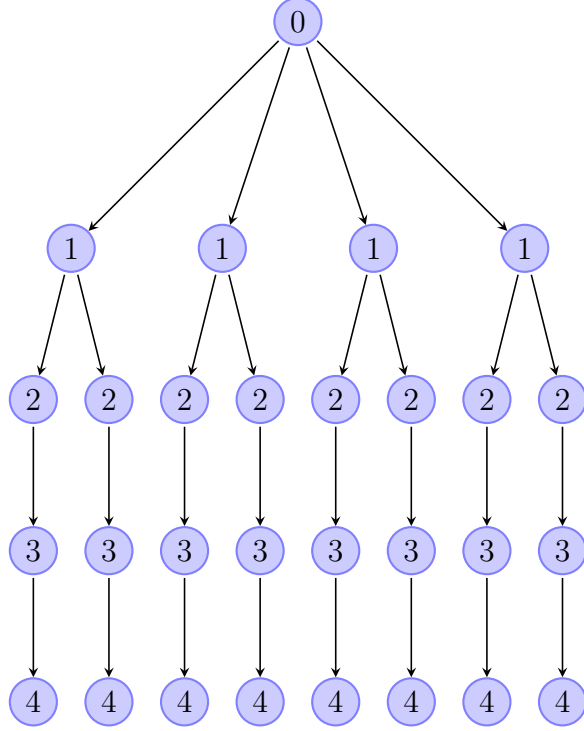


Figure 2: A regular monotone hierarchy

We define the centrality ranking : $i \succ j$ if and only if $\rho(i) < \rho(j)$ and $i \doteq j$ if and only if $\rho(i) = \rho(j)$. We now have

PROPOSITION 2 *In a regular monotone hierarchy, $i \succ j$ if and only if $n_i \succ n_j, I_i \succ I_j$ and $w_i \succ w_j$ and $i \doteq j$ if and only if $n_i = n_j, I_i = I_j$ and $w_i = w_j$. For any tree which is not a regular monotone hierarchy, there exist two nodes i and j and two statistics $s, s' \in \{n, I, w\}$ such that $s_i \succeq s_j$ and $s'_j \succ s'_i$.*

Proposition 2 shows that, in a regular monotone hierarchy, *all* centrality measures based on neighborhood, intermediary, and walk statistics rank nodes in the same order: based on their distance from the root. This is also true for any other statistic for which distance from the root orders nodal statistics according to \succeq (and it is hard to think of any natural statistic that would not do this in such a network). Conversely, if the social network is a tree which is not a regular monotone hierarchy, then the centrality measures will not coincide. The intuition underlying Proposition 2 is as follows. In a regular monotone hierarchy, agents who are more distant from the root have longer distances to travel to other nodes, are less likely to lie on paths between other nodes, and have a smaller number of walks emanating from them. Next consider the leaves of the tree. By definition, they do not sit on any path connecting other agents and have a intermediary statistic equal to $I = (0, 0, \dots, 0)$. Hence, if

centrality measures are to coincide, all leaves of the tree must have the same neighborhood statistic, a condition which can only be satisfied in a regular monotone hierarchy. This last argument shows that centrality rankings based on the intermediary and neighborhood statistic can only coincide in regular monotone hierarchies.

4.2 Simulations: Differences in Centrality Measures by Network Type

Another way to see how centrality measures compare, is to examine how differently they rank nodes on various random networks. To perform this exercise, we simulate networks on 40 nodes. We vary the type of network to have three different basic structures. The first is an Erdos-Renyi random graph in which all links are formed independently. The second is a network that has some homophily: there are two types of nodes and we connect nodes of the same type with a different probability than nodes of different types. The third is a variant of a homophilistic network in which some nodes are ‘bridge nodes’ that connect to other nodes with a uniform probability, thus putting them as connector nodes between the two homophilistic groups. We vary the overall average degrees of the networks to be either 2, 5 or 10. In the cases of the homophily and homophily bridge nodes, there are also relative within and across group link probabilities that vary. Given all of these dimensions, we end up with many different networks on which to compare centrality measures, and so many of the results appear in the appendix, and a few representative results are presented here.

We then compare 5 different centrality measures on these networks: degree, decay, closeness, diffusion, and Katz-Bonacich. Decay, diffusion and Katz-Bonacich all depend on a parameter that we call the decay parameter, and we vary that as well.¹²

The details on the three network types we perform the simulation on are:

- **ER random graphs:** Each possible link is formed independently with probability $p = \bar{d}/(n-1)$.

- **Homophily:**

There are two equally-sized groups of 20 nodes. Links between pairs of nodes in the same group are formed with probability p_{same} and between pairs of nodes in different groups are formed with probability p_{diff} , all independently.

Letting $p_{same} = H \times p_{diff}$, average degree is:

$$\bar{d} = \left(\frac{n}{2} - 1\right) p_{same} + \frac{n}{2} p_{diff}$$

¹²In addition, diffusion centrality has $T = 5$ in all of the simulations.

- **Homophily with Bridge Nodes:**

There are L bridge nodes and two equal-sized groups of $\frac{nN-L}{2}$ non-bridge nodes. Each bridge node connects to any other node with probability p_b . Non-bridge nodes connect to other same group nodes with probability p_{same} and different group nodes with probability p_{diff} .

Letting $p_{same} = H \times p_{diff}$, we set $p_b = \bar{d} / (n - 1)$ where \bar{d} is defined as the average degree

$$\bar{d} = \left(\frac{n-L}{2} - 1 \right) p_{same} + \left(\frac{n-L}{2} \right) p_{diff} + Lp_b.$$

A first thing to note about the simulations (see the tables below) is that the correlation in rankings of the various centrality measures is very high across all of the simulations and measures, often above .9, and usually in the .8 to 1 range. This is in part reflective of what we have seen from our characterizations: all of these measures operate in a similar manner and are based on nodal statistics that often move in similar ways: nodes with higher degree tend to be closer to other nodes and have more walks to other nodes, and so forth. In terms of differences between measures, closeness and betweenness are more distinguished from the others in terms of correlation, while the other measures all correlate above .98 in Table 1.

These extreme correlations are higher than those found in Valente, Coronges, Lakon, and Costenbader (2008), who also find high correlations, but lower in magnitude, when looking at a series of real data sets. The artificial nature of the Erdos-Renyi networks serves as a benchmark from which we can jump off as it results in less differentiation between nodes than one finds in many real-world networks, but also allows us to know that differences among nodes are coming from random variations. As we add homophily in Table 3, and then bridge nodes in Table 4, we see the correlations drop significantly, especially comparing betweenness centrality to the others, and this then has an intuitive interpretation as bridge nodes naturally have high betweenness centrality, but may not stand out according to other measures.

Correlation is a very crude measure, and it does not capture whether nodes are switching ranking or by how much. Some nodes could have dramatically different rankings and yet the correlation could be relatively high overall. Thus, we also look at how many nodes switch rankings between two measures, as well as how the maximal extent to which some node changes rankings. There, we see more substantial differences across centrality measures, and with most measures being more highly distinguished from each other.

As we increase the decay factor (e.g., from Table 1 to 2), we see greater differences between the measures, as the correlations drop and we see more changes in the rankings. With a very low decay factor, decay, diffusion, Katz-Bonacich are all very close to degree, while for higher decay parameters they begin to differentiate themselves. This makes sense as it allows the

measures to incorporate information that depends on more of the network, and that is less tied to immediate neighborhoods.

5 Concluding Remarks: Potential for New Measures

Given that our results show that all standard centrality measures are based on the same method of aggregation, there seems to be ample room for the development of new measures. We close with thoughts on such classes of measures.

The first class of measures that may be worth exploring in greater detail are those based on power indices from cooperative game theory, with the Shapley value being a prime example. Myerson (1977) adapted the Shapley value to allow for communication structures, and Jackson and Wolinsky (1996) later adapted the Myerson value to more general network settings. The Myerson value defined in Jackson and Wolinsky (1996) provides a whole family of centrality measures, as once one ties down how value is generated by the network, it then indicates how much of that value is allocated - or ‘due’ - to each node. Some specific instances of these measures have popped up in the later literature Gomez, Gonzalez-Aranguena, Manuel, Owen, Pozo, and Tejada (2003); Michalak, Aadithya, Szczepanski, Ravindran, and Jennings (2013); Molinero, Riquelme, and Serna (2013). These can be difficult to compute, and in some cases still satisfy variations on the additivity axiom. It appears that whether or not the additivity axiom would be violated depends on the choice of the value function.¹³ The choice of value function would be tied to the application.

One other idea, which is a more direct variation on standard centrality measures, is to look at things like the probability of infecting a whole population starting from some node and a diffusion/contagion process, rather than the expected number of infected nodes, as embodied in the notion of contagion centrality defined by Lim, Ozdaglar, and Teytelboym (2015). One could also ask questions for other fractions of the population, or getting nontrivial diffusion, etc. Although such measures build on the same sorts of models as diffusion and related centralities, they clearly violate the additivity axiom, and so would move outside of the standard classes. The differences that they exhibit compared to standard measures would be interesting to explore.

Another new class of measures that may be worth exploring involve a multiplicative formulation instead of an additive one. This would reflect strong complementarities among different elements of the nodal statistics, for instance nodes at various distances. Given scalars α_ℓ and β_ℓ that capture the relative importance of the different dimensions of the nodal statistics,

¹³Even though the Shapley value satisfies an additivity axiom, it is an additivity across value functions and not across nodal statistics; and so does not translate here.

for instance the role of nodes at various distances from the node in question, we define a new family of centrality measures as follows:

$$c_i(g) = \mathcal{C}(s_i(g)) = \times_{\ell=1}^L (\alpha_\ell + s_i^\ell)^{\beta_\ell}. \quad (4)$$

These are a form of multiplicative measures that parallel the form of some production functions and would capture the idea, for instance, that nodes at various distances are complementary inputs into a production process for a given node.¹⁴ This class of measures could produce different rankings of nodes compared to standard centrality measures, and would capture ideas such as nodes that are well-balanced in terms of how many other nodes are at various distances, for instance.¹⁵

References

- BAILEY, N. (1975): *The Mathematical Theory of Infectious Diseases*, London: Griffin.
- BALLESTER, C., A. CALVÓ-ARMENGOL, AND Y. ZENOU (2006): “Who’s who in networks, wanted: the key player,” *Econometrica*, 74, 1403–1417.
- BANERJEE, A., A. CHANDRASEKHAR, E. DUFLO, AND M. O. JACKSON (2013): “Diffusion of Microfinance,” *Science*, 341, DOI: 10.1126/science.1236498, July 26 2013.
- (2014): “Gossip and Identifying Central Individuals in a Social Network,” *working paper*: <http://ssrn.com/abstract=2425379>.
- BAVELAS, A. (1950): “Communication patterns in task-oriented groups,” *Journal of the acoustical society of America*, 22, 725–730.
- BONACICH, P. (1972): “Factoring and weighting approaches to status scores and clique identification,” *Journal of Mathematical Sociology*, 2(1), 113–120.
- (1987): “Power and centrality : a family of measures,” *American Journal of Sociology*, 92, 1170–1182.
- BORGATTI, S. P. (2005): “Centrality and network flow,” *Social Networks*, 27, 55–71.

¹⁴It would generally make sense to have the β_ℓ be a non-increasing function of ℓ . The presence of the α_ℓ s ensures that there is no excessive penalty for having $s_i^\ell = 0$ for some ℓ .

¹⁵Note that even the ordering produced by this class of measures is equivalent to ordering nodes according to $\sum_{\ell=1}^L \beta_\ell \log(\alpha_\ell + s_i^\ell)$. This is an additive form, with nodal statistics $\beta_\ell \log(\alpha_\ell + s_i^\ell)$. This shows that it can be challenging to escape the additive family. Nonetheless, this is a new and potentially interesting family prompted by our analysis.

- CHRISTAKIS, N. AND J. FOWLER (2010): “Social Network Sensors for Early Detection of Contagious Outbreaks,” *PLOS One*: *e12948*. doi:10.1371/journal.pone.0012948, 5(9).
- DEBREU, G. (1960): “Topological Methods in Cardinal Utility Theory,” *Mathematical Methods in the Social Sciences*, Stanford University Press, 16–26.
- DEQUIEDT, V. AND Y. ZENOU (2014): “Local and Consistent Centrality Measures in Networks,” *unpublished*.
- FREEMAN, L. (1977): “A set of measures of centrality based on betweenness,” *Sociometry*, 40(1), 35–41.
- GARG, M. (2009): “Axiomatic Foundations of Centrality in Networks,” *unpublished*.
- GOFMAN, M. (2015): “Efficiency and Stability of a Financial Architecture with Too-Interconnected-To-Fail Institutions,” *Forthcoming at the Journal of Financial Economics*.
- GOLUB, B. AND M. O. JACKSON (2010): “Naive Learning in Social Networks and the Wisdom of Crowds,” *American Economic Journal: Microeconomics*, 2, 112–149.
- GOMEZ, D., E. GONZALEZ-ARANGUENA, C. MANUEL, G. OWEN, M. D. POZO, AND J. TEJADA (2003): “Centrality and power in social networks: A game theoretic approach,” *Mathematical Social Sciences*, 46 (1), 27–54.
- HAHN, Y., A. ISLAM, E. PATACCHINI, AND Y. ZENOU (2015): “Teams, organization and education outcomes: Evidence from a field experiment in Bangladesh,” *CEPR Discussion Paper No. 10631*.
- HENRIET, D. (1985): “The Copeland choice function: an axiomatic characterization,” *Social Choice and Welfare*, 2, 49–63.
- JACKSON, M. O. (2008): *Social and economic networks*, Princeton: Princeton University Press.
- (2016): “The Friendship Paradox and Systematic Biases in Perceptions and Social Norms,” <http://ssrn.com/abstract=2780003>.
- JACKSON, M. O. AND A. WOLINSKY (1996): “A Strategic Model of Social and Economic Networks,” *Journal of Economic Theory*, 71, 44–74.
- KATZ, L. (1953): “A new status index derived from sociometric analysis,” *Psychometrika*, 18, 39–43.

- KEMPE, D., J. KLEINBERG, AND E. TARDOS (2003): “Maximizing the Spread of Influence through a Social Network,” *Proc. 9th Intl. Conf. on Knowledge Discovery and Data Mining*, 137146.
- KERMANI, M. A. M. A., A. BADIEE, A. ALIAHMADI, M. GHAZANFARI, AND H. KALANTARI (2015): “Introducing a procedure for developing a novel centrality measure (Sociability Centrality) for social networks using TOPSIS method and genetic algorithm,” *in press*.
- KÖNIG, M., C. TESSONE, AND Y. ZENOU (2014): “Nestedness in networks: A theoretical model and some applications,” *Theoretical Economics*, *forthcoming*.
- KOOPMANS, T. C. (1960): “Stationary ordinal utility and impatience,” *Econometrica*, 28, 287–309.
- LASLIER, J.-F. (1997): *Tournament Solutions and Majority Voting*, Springer-Verlag, Berlin.
- LIM, Y., A. OZDAGLAR, AND A. TEYTELBOYM (2015): “A Simple Model of Cascades in Networks,” *mimeo*.
- MICHALAK, T. P., K. V. AADITHYA, P. L. SZCZEPANSKI, B. RAVINDRAN, AND N. R. JENNINGS (2013): “Efficient Computation of the Shapley Value for Game-Theoretic Network Centrality,” *Journal of Artificial Intelligence Research*, 46, 607–650.
- MOLINERO, X., F. RIQUELME, AND M. SERNA (2013): “Power indices of influence games and new centrality measures for social networks,” *Arxiv*.
- MYERSON, R. B. (1977): “Graphs and cooperation in games,” *Mathematics of Operation Research*, 2, 225–229.
- PADGETT, J. AND C. ANSELL (1993): “Robust action and the rise of the Medici 1400-1434,” *American Journal of Sociology*, 98(6), 1259–1319.
- PAGE, L., S. BRIN, R. MOTWANI, AND T. WINOGRAD (1998): “The PageRank Citation Ranking: Bringing Order to the Web,” *Technical Report: Stanford University*.
- PALACIOS-HUERTA, I. AND O. VOLIJ (2004): “The measurement of intellectual influence,” *Econometrica*, 72, 963–977.
- ROCHAT, Y. (2009): “Closeness centrality extended to unconnected graphs: The harmonic centrality index,” *Applications of Social Network Analysis*.
- RUBINSTEIN, A. (1980): “Ranking the participants in a tournament,” *SIAM Journal of Applied Mathematics*, 38, 108–111.

- SABIDUSSI, G. (1966): “The centrality index of a graph,” *Psychometrika*, 31, 581–603.
- SLUTZKI, G. AND O. VOLIJ (2006): “Scoring of web pages and tournaments-axiomatizations,” *Social Choice and Welfare*, 26, 75–92.
- VALENTE, T. W., K. CORONGES, C. LAKON, AND E. COSTENBADER (2008): “How Correlated Are Network Centrality Measures?” *Connect (Tor)*, 28, 16–26.
- VAN DEN BRINK, R. AND R. P. GILLES (2003): “Ranking by outdegree for directed graphs,” *Discrete Mathematics*, 271, 261–270.
- WASSERMAN, S. AND K. FAUST (1994): *Social Network Analysis*, Cambridge: Cambridge University Press.

Appendix: Proofs

Proof of Lemma 1: Suppose that $s(\mathbf{g}) = s(\mathbf{g}')$ but $c_i(\mathbf{g}) \neq c_i(\mathbf{g}')$. Without loss of generality let $c_i(\mathbf{g}) > c_i(\mathbf{g}')$. By monotonicity, since $c_i(\mathbf{g}) > c_i(\mathbf{g}')$ it must be that $s_i(\mathbf{g}') \not\leq s_i(\mathbf{g})$. However, this contradicts the fact that $s_i(\mathbf{g}) = s_i(\mathbf{g}')$ which implies that $s_i(\mathbf{g}) \sim s_i(\mathbf{g}')$ (by reflexivity of a partial order). Let S denote the range of $s_i(g)$ (which is the same for all i by symmetry). Thus, there exists $\mathcal{C}_i : S \rightarrow \mathbb{R}$ for which $c_i(\mathbf{g}) = \mathcal{C}_i(s_i(\mathbf{g}))$, and moreover this must be a monotone function on the range of $s_i(g)$, denoted S , given the monotonicity of c_i . We extend the function \mathcal{C}_i to be monotone on all of \mathbb{R}^L as follows. Let S_1 be the set of $s \notin S$ such that there exists some $s' \in S$ for which $s' \geq s$. For any $s \in S_1$ $\mathcal{C}_i(s) = \inf_{s' \in S, s' \geq s} \mathcal{C}_i(s')$. Next, let S_2 be the set of $s \notin S \cup S_1$. For any $s \in S_2$ let $\mathcal{C}_i(s) = \sup_{s' \in S \cup S_1, s' \leq s} \mathcal{C}_i(s')$ (and note that this is well defined for all $s \in S_2$ since there is always some $s' \in S \cup S_1, s' \leq s$). This is also monotone, by construction.

Next, suppose that there are two nodes i and j for which $s_i(\mathbf{g}) = s_j(\mathbf{g}')$ but $c_i(\mathbf{g}) \neq c_j(\mathbf{g}')$. Again, without loss of generality let $c_i(\mathbf{g}) > c_j(\mathbf{g}')$. Consider a permutation π such that $\pi(j) = i$ and $\pi(i) = j$. Because the statistic is symmetric, $s_i(\pi\mathbf{g}') = s_j(\mathbf{g}')$, and by symmetry of the centrality measure $c_i(\mathbf{g}' \circ \pi) = c_j(\mathbf{g}')$. Hence, $c_i(\mathbf{g}) > c_i(\mathbf{g}' \circ \pi)$ and $s_i(\mathbf{g}) = s_i(\mathbf{g}' \circ \pi)$, yielding a contradiction of our first claim. Thus, it must be that $\mathcal{C}_i = \mathcal{C}_j$. \square

Proof of Theorems 1-3

IF part:

It is easily checked that if \mathcal{C} can be expressed as in equation (1), then independence holds. Similarly, if the representation is as in (2), then recursivity also holds, as does additivity if (3) is satisfied.

ONLY IF part:

Let \mathbf{e}^ℓ denote the vector in \mathbb{N}^L with every $\mathbf{e}_\ell^\ell = 1$ and $\mathbf{e}_j^\ell = 0$ for all $j \neq \ell$. Define $F_\ell : \mathbb{R} \rightarrow \mathbb{R}_+$ as

$$F_\ell(x) = \mathcal{C}(x\mathbf{e}^\ell). \quad (5)$$

Iterated applications of independence imply that

$$\mathcal{C}(s_i) = \sum_{\ell=1}^L F_\ell(s_{i,\ell}). \quad (6)$$

To see this, note that for $s_i = (x, 0, \dots, 0, \dots)$, $s'_i = (0, y, 0, \dots, 0, \dots)$, and $s''_i = (x, y, 0, \dots, 0, \dots)$, independence requires

$$\begin{aligned} \mathcal{C}(s''_i) - F_2(y) &= F_1(x) - 0, \\ \mathcal{C}(s''_i) &= F_1(x) + F_2(y). \end{aligned}$$

Doing this again for $s_i = (x, y, 0, 0, \dots, 0, \dots)$, $s'_i = (x, 0, z, 0, \dots, 0, \dots)$, and $s''_i = (x, y, z, 0, \dots, 0, \dots)$, independence requires

$$\mathcal{C}(s''_i) = F_1(x) + F_2(y) + F_3(z).$$

By induction, this holds for arbitrary vectors. Monotonicity implies that F_ℓ is increasing and $F_\ell(0) = 0$ for all ℓ .

By recursivity, for all $\ell \leq L$ and all x in \mathbb{N} :

$$\frac{F_{\ell+1}(x)}{F_\ell(x)} = \frac{F_2(x)}{F_1(x)} = \delta(x). \quad (7)$$

Moreover, recursivity also implies that for any two x, x' in \mathbb{R} and any ℓ (provided the denominators are not 0):

$$\frac{F_\ell(x')}{F_\ell(x)} = \frac{F_1(x')}{F_1(x)}. \quad (8)$$

From (7) this is true only if $\delta(x) \equiv \delta$ is constant.

Next, note that (8), together with the fact $F_\ell(0) = 0$ for all ℓ , imply that $F_\ell(x) = \delta^\ell f(x)$ for a common f for which $f(0) = 0$. This implies that \mathcal{C} can be written as in equation (2).

Finally, additivity (which clearly implies independence) then implies that f is linear (with a slope of 1), and given that it must be that $f(0) = 0$, the final characterization follows. \square

Proof of Proposition 1: [IF] We first provide a formula to compute the number of nodes at distance less than or equal to d from node i for a monotone hierarchy, $Q(i, d)$. Let $\rho(i)$ denote the distance from the root and $i + 0, i_1, \dots, i_k, \dots, i_{\rho(i)} = i$ the unique path between the root and node i . Let $p(i, \ell)$ denote the number of successors of node i at distance ℓ . If $d \geq \rho(i)$, we compute the number of nodes at distance less than or equal to d as

$$\begin{aligned}
Q(i, d) &= p(i_0, 0) + p(i_0, 1) + \dots + p(i_0, d - \rho(i)) \\
&+ p(i_1, d - \rho(i)) + p(i_1, d - \rho(i) + 1) + p(i_2, d - \rho(i) + 1) + p(i_2, d - \rho(i) + 2) \\
&+ \dots p(i_{\rho(i)-1}, d - 2) + p(i_{\rho(i)-1}, d - 1) + p(i, d - 1) + p(i, d).
\end{aligned}$$

To understand this computation, notice that all nodes which are at distance less than or equal to $d - \rho(i)$ from the root are at a distance less than d from node i . Other nodes at a distance less than d from node i are computed considering the path between i_0 and i . Fix i_1 . There are successor nodes which are at distance $d - \rho(i)$ from node i_1 (and hence at a distance $d - 1$ from i) and were not counted earlier because they are at a distance of $d - \rho(i) + 1$ from the root, and successor nodes which are at a distance $d - \rho(i) + 1$ from node 1 (and hence at a distance d from i) and were not counted earlier because they are at a distance $d - \rho(i) + 2$ from the root. Continuing along the path, for any node i_k we count successor nodes at a distance $d - \rho(i) + k - 1$ and $d - \rho(i) + k$ from node i_k which are at a distance $d - 1$ and d from node i and were not counted earlier, and finally obtain the total number of nodes at a distance less or equal to d from node i .

Next suppose that $d \leq \rho(i)$. In that case, no node beyond i_0 who does not belong to the subtree starting at i_1 can be at a distance smaller than d . The expression for the number of nodes at a distance less than or equal to d simplifies to

$$\begin{aligned}
Q(i, d) &= p(i_{\rho(i)-d}, 0) + p(i_{\rho(i)-d+1}, 0) + p(i_{\rho(i)-d+1}, 1) \\
&+ \dots p(i_{\rho(i)-1}, d - 2) + p(i_{\rho(i)-1}, d - 1) + p(i, d - 1) + p(i, d).
\end{aligned}$$

Next we prove the following claim:

CLAIM 1 *In a monotone hierarchy, for any i, j such that $\rho(j) = \rho(i) + 1$, for any t , $p(i, \ell) \geq p(j, \ell)$*

Proof of the Claim: The proof is by induction on d . For $d = 1$, the statement is true as $p(i, 1) \equiv d(i) \geq d(j) \equiv p(j, 1)$. Suppose that the statement is true for all $d' < d$. Let i_1, \dots, i_I be the direct successors of i and j_1, \dots, j_J the direct successors of j , with $J < I$. Then

$$\begin{aligned}
p(i, d) &= \sum_{r=1}^I p(i_r, d - 1), \\
&\geq \sum_{r=1}^J p(i_r, d - 1), \\
&\geq \sum_{r=1}^J p(j_r, d - 1) \\
&= p(j, d).
\end{aligned}$$

where the first inequality is due to the fact that $I \geq J$ and the second that, by the induction hypothesis, as $\rho(i_r) = \rho(j_r) - 1$ for all r , $p(i_r, d' - 1) \geq p(j_r, d' - 1)$.

Consider a monotone hierarchy and two nodes i, j such that $\rho(j) = \rho(i) + 1$. Let $d \geq \rho(i) + 1$ and $i_0, i_1, \dots, i_{\rho(i)}, i_0, j_1, \dots, j_r, j_{\rho(i)+1}$ the paths linking i and j to the root. Then

$$\begin{aligned} Q(i, d) - Q(j, d) &= p(i_0, d - \rho(i)) - p(j_1, d - \rho(i) - 1) - p(j_1, d - \rho(i)) \\ &+ [p(i_1, d - \rho(i)) + p(i_1, d - \rho(i) + 1) - p(j_2, d - \rho(i)) - p(j_2, d - \rho(i) + 1)] \\ &+ \dots [p(i_r, d - \rho(i) + r - 1) + p(i_r, d - \rho(i) + r) - p(j_{r+1}, d - \rho(i) + r - 1) - p(j_{r+1}, d - \rho(i) + r)] \\ &+ \dots [p(i, d - 1) + p(i, d) - p(j, d - 1) - p(j, d)] \end{aligned}$$

Notice that $p(i_0, d - \rho(i)) = p(j_1, d - \rho(i) - 1) + \sum_{k \neq j_1, \rho(k)=1} p(k, d - \rho(i) - 1)$ and that $p(j_1, d - \rho(i)) = \sum_{l | \rho(l)=2, \rho(j_1, l)=1} p(l, d - \rho(i) - 1)$. By Claim 1, as $\rho(l) = \rho(k + 1)$, $p(l, d - \rho(i) - 1) \leq p(k, d - \rho(i) - 1)$ and as $d(j_1) \leq d(i_0) - 1$, $\sum_{k \neq j_1, \rho(k)=1} p(k, d - \rho(i) - 1) \geq \sum_{l | \rho(l)=2, \rho(j_1, l)=1} p(l, d - \rho(i) - 1)$. Furthermore, by Claim 1, for all r and all d , $p(i_r, d) \geq p(j_{r+1}, d)$, so that $Q(i, d) - Q(j, d) \geq 0$.

Next suppose that $d \leq \rho(i) < \rho(i) + 1 + 1$. Then

$$\begin{aligned} Q(i, d) - Q(j, d) &= [p(i_{\rho(i)-d}, 0) + p(i_{\rho(i)-d+1}, 0) - p(j_{\rho(i)-d+1}, 0) - p(j_{\rho(i)-d+2}, 0)] \\ &+ \dots [p(i_{\rho(i)-1}, d - 1) + p(i, d - 1) - p(j_{\rho(i)}, d - 1) - p(j, d - 1)] \\ &+ [p(i, d) - p(j, d)] \end{aligned}$$

and by a direct application of Claim 1, $Q(i, d) - Q(j, d) \geq 0$. We finally observe that there exists a distance d such that $Q(i, d) > Q(j, d)$. Let h be the total number of levels in the hierarchy. Consider a distance d such that $h = d + \rho(i)$. Then there are successor nodes at distance d from i but no successor nodes at distance d from j . Hence $p(i, d) > 0 = p(j, d)$. This establishes that $Q(i, d) > Q(j, d)$ and hence $n_i(\mathbf{g}) \succ n_j(\mathbf{g})$. By a repeated application of the same argument, for any i, j such that $\rho(i) < \rho(j)$, for any i, j such that $\rho(i, i_0) < \rho(j, i_0)$, $n_i(\mathbf{g}) \succ n_j(\mathbf{g})$.

Now consider two nodes i, j at the same level of the hierarchy but such that $i \succ j$. Let $i_0, i_1, \dots, i_{\rho(i)} = i$ and $i_0, j_1, \dots, j_{\rho(i)} = j$ be the nodes connecting i and j to the root. Consider the two subtrees starting at i_1 and j_1, g_{i_1} and g_{j_1} . By construction of the monotone hierarchy, if $i \succ j$, then for any two nodes k, l such that k and l are at the same level of the hierarchy, $k \in g_{i_1}, l \in g_{j_1}$, $d(k) \geq d(l)$ with strict inequality for some pair of nodes. But this implies that $p(i_r, d) \geq p(j_r, d)$ for all $r = 1, 2, \dots, \rho(i)$ and all d with strict inequality for some pair (i_r, j_r) and some distance d . Now compute, for $d \geq \rho(i)$,

$$\begin{aligned}
Q(i, d) - Q(j, d) &= \sum_{r=1}^{\rho(i)} [p(i_r, d - \rho(i) + r - 1) + p(i_r, d - \rho(i) + r) \\
&\quad - p(j_r, d - \rho(i) + r - 1) - p(j_r, d - \rho(i) + r)] \\
&\geq 0,
\end{aligned}$$

and for $d \leq \rho(i)$,

$$\begin{aligned}
Q(i, d) - Q(j, d) &= \sum_{r=\rho(i)-d+1}^{\rho(i)} [p(i_r, d - \rho(i) + r - 1) + p(i_r, d - \rho(i) + r) \\
&\quad - p(j_r, d - \rho(i) + r - 1) - p(j_r, d - \rho(i) + r)] \\
&\geq 0,
\end{aligned}$$

To show that there exists a distance d such that $Q(i, d) - Q(j, d) > 0$ pick $\rho(i) + d = h$. Then all nodes in the subtrees g_{i_1} and g_{j_1} are counted, and there must exist a pair (i_r, j_r) such that $p(i_r, d - \rho(i) + r) - p(j_r, d - \rho(i) + r) > 0$. We thus have shown that for any i, j at the same level of the hierarchy such that $i \succ j$, $n_i(\mathbf{g}) \succ n_j(\mathbf{g})$.

Pick two nodes at the same level of the hierarchy such that $i \doteq j$. It must be that the two subgraphs starting at i_1 and j_1 are symmetric, with the same number of nodes at each level of the hierarchy. Hence i and j are symmetric, and $n_i(\mathbf{g}) = n_j(\mathbf{g})$.

[ONLY IF]: Suppose that the tree g is not a monotone hierarchy and has an even diameter. Consider a line in the tree which has the same length as the diameter of the tree. Pick as a root the unique middle node in the line and let h be the maximal distance between the root and a terminal node. First assume that there exist two nodes i and j such that $\rho(j) = \rho(i) + 1$ but $d_j > d_i$. Then clearly $Q(j, 1) > Q(i, 1)$. Notice that all nodes are at a distance less than or equal to $d = h + \rho(i)$ from node i whereas there exist nodes which are at a distance $h + \rho(i) + 1$ from node j , and hence $Q(j, h + \rho(i)) < Q(i, h + \rho(i))$ so that neither $n_i \succeq n_j$ nor $n_j \succeq n_i$.

Next suppose that for all nodes i, j such that $\rho(j) = \rho(i) + 1$, $d_j \leq d_i$, but that there exists two nodes i, j at the same level of the hierarchy such that $d_i > d_j$ and two successors of i and j , k and l , at the same level of the hierarchy such that $d_k < d_l$. Because $d_i > d_j$, $Q(i, 1) > Q(j, 1)$. Suppose that $n_i \succ n_j$. Then $Q(i, d) > Q(j, d)$ for all $d = 1, 2, \dots, h + \rho(i)$. Now consider the two successors k and l of i and j . As $d_k < d_l$, $Q(k, 1) < Q(l, 1)$. Now count all the nodes which are at a distance less than $h + \rho(k)$ from k , $Q(k, h + \rho(k) - 1)$. As k is a successor of i , $Q(k, h + \rho(k) - 1) = Q(i, h + \rho(i) - 1)$. By assumption $Q(i, h + \rho(i) - 1) > Q(j, h + \rho(j) - 1) = Q(l, h + \rho(l) - 1)$, showing that neither $n_k \succ n_l$ nor $n_l \succ n_k$, completing the proof of the Proposition. \square

Proof of Proposition 2: [IF] Because a regular monotone hierarchy is a monotone hierarchy, we know by Proposition 1 that if $i \succ j$, $n_i \succ n_j$.

Next we show that the number of geodesic paths of any length d between two nodes is smaller for a node further away from the root. To this end, consider two nodes i and j such that j is a direct successor of i . For any d , if a geodesic path contains j but not i , then i must be an endpoint of the path. Hence, the total number of geodesic paths of length d going through j but not through i is $2p(j, d-1)$. If $d(i) \geq 3$, pick a direct successor $k \neq j$ of i , and consider paths of length d connecting successors of k to j . All these paths must go through i and there are $2p(k, d-1) = 2p(j, d-1)$ such paths. If $d(i) = 2$, then $d(j) \leq 2$ so that $2p(j, d-1) = 0$ or $2p(j, d-1) = 2$. If $2p(j, d-1) = 2$, then d is small enough so that there exists at least two paths of length d connecting a node in the network to j through i . Furthermore, if $d = h - \rho(i, i_0) + 1$ where h is the number of levels of the hierarchy, there is no path of length d connecting i to a node through j whereas there exist paths of length d connecting a node to j through i , so that $I_i \geq I_j$.

Next, we compute the number of walks emanating from two nodes i and j at different levels of the hierarchy. Let $w_k(d)$ denote the number of walks of length d emanating from a node at level ℓ . We show that $w_\ell(d) \geq w_{\ell+1}(d)$. We compute the number of walks recursively:

$$\begin{aligned} w_\ell(d) &= [d(\ell) - 1]w_{\ell+1}(d-1) + w_{\ell-1}(d-1) \text{ for } \ell \geq 1 \\ w_0(d) &= d(0)w_1(d-1) \end{aligned}$$

We also have $w_\ell(0) = 1$ for all ℓ which allows us to start the recursion.

Next we prove that $w_d(\ell) \geq w_{\ell+1}(d)$ for $i = 1, \dots, I-1$ by induction on d . The statement is trivially true for all ℓ at $d = 0$. Now suppose that the statement is true at $d-1$. We first show that the inequality holds for all nodes but the root. For $\ell \geq 1$,

$$\begin{aligned} w_\ell(d) &= [d(\ell) - 1]w_{\ell+1}(d-1) + w_{\ell-1}(d-1) \\ &\geq [d(\ell+1) - 1]w_{\ell+2}(d-1) + w_\ell(d-1) \\ &= w_{\ell+1}(d) \end{aligned}$$

The more difficult step is to show that the statement is also true for the root. To this end, we prove by induction on d that for all $\ell = 1, \dots, h-1$:

$$d(0)w_\ell(d) \geq [d(\ell) - 1]w_{\ell+1}(d) + w_{\ell-1}(d),$$

The statement is true at $d = 0$ because $d(0) \geq d(\ell)$ for all $\ell \geq 1$. Next compute

$$\begin{aligned}
d(0)w_\ell(d) &= d(0)[[d(\ell) - 1]w_{\ell+1}(d-1) + w_{\ell-1}(d-1)], \\
(d(\ell-1) - 1)w_{\ell+1}(d) + w_{\ell-1}(d) &= [d(\ell) - 1][[d(\ell+2) - 1]w_{\ell+2}(d-1) + w_\ell(d-1)] \\
&+ [d(\ell-1) - 1]w_\ell(d-1) + w_{\ell-2}(d-1).
\end{aligned}$$

By the induction hypothesis,

$$d(0)w_{\ell-1}(d-1) \geq [d(\ell-1) - 1]w_\ell(d-1) + w_{\ell-2}(d-1),$$

and

$$d(0)w_{\ell+1}(d-1) \geq [d(\ell+1) - 1]w_{\ell+2}(d-1) + w_\ell(d-1) \geq [d(\ell+2) - 1]w_{\ell+2}(d-1) + w_\ell(d-1).$$

Replacing, we obtain

$$d(0)w_\ell(d) \geq [d(\ell) - 1]w_{\ell+1}(d) + w_{\ell-1}(d),$$

concluding the inductive argument. Applying this formula for $\ell = 1$, we have $w_0(d) = d(0)w_1(d-1) \geq [d(1) - 1]w_2(d-1) + w_0(d-1) = w_1(d)$, completing the proof that $w_\ell(d) \geq w_{\ell+1}(d)$ for all d .

[ONLY IF] Consider a leaf i of the tree. Then $w_i = (0, 0 \dots 0)$. So all leaves have the same centrality based on the intermediary statistic. They must also have the same centrality based on the neighborhood statistic, which implies that the tree is a regular monotone hierarchy. \square

Table 1: Avg degree 2, Erdos Renyi, Decay .15

Cent1	Cent2	Correlation of Rank Vector	Fraction of Sims w same Top Node	Fraction Nodes Switch Rank	Max Change in % Rank
Degree	Decay	0.99	1.00	0.73	0.16
Degree	Closeness	0.83	0.78	0.77	0.34
Degree	Diffusion	0.98	0.98	0.74	0.17
Degree	KatzBon	1.00	1.00	0.73	0.16
Degree	Between	0.83	0.50	0.93	0.30
Decay	Degree	0.99	1.00	0.73	0.16
Decay	Closeness	0.88	0.74	0.61	0.23
Decay	Diffusion	1.00	0.86	0.30	0.08
Decay	KatzBon	0.99	0.86	0.29	0.08
Decay	Between	0.84	0.42	0.90	0.32
Closeness	Degree	0.83	0.78	0.77	0.34
Closeness	Decay	0.88	0.74	0.61	0.23
Closeness	Diffusion	0.86	0.66	0.65	0.24
Closeness	KatzBon	0.83	0.66	0.65	0.25
Closeness	Between	0.67	0.52	0.91	0.35
Diffusion	Degree	0.98	0.98	0.74	0.17
Diffusion	Decay	1.00	0.86	0.30	0.08
Diffusion	Closeness	0.86	0.66	0.65	0.24
Diffusion	KatzBon	0.99	0.98	0.11	0.04
Diffusion	Between	0.83	0.34	0.91	0.36
KatzBon	Degree	1.00	1.00	0.73	0.16
KatzBon	Decay	0.99	0.86	0.29	0.08
KatzBon	Closeness	0.83	0.66	0.65	0.25
KatzBon	Diffusion	0.99	0.98	0.11	0.04
KatzBon	Between	0.83	0.34	0.91	0.35
Between	Degree	0.83	0.50	0.93	0.30
Between	Decay	0.84	0.42	0.90	0.32
Between	Closeness	0.67	0.52	0.91	0.35
Between	Diffusion	0.83	0.34	0.91	0.36
Between	KatzBon	0.83	0.34	0.91	0.35

Table 2: Avg degree 2, Erdos Renyi, Decay .5

Cent1	Cent2	Correlation of Rank Vector	Fraction of Sims w same Top Node	Fraction Nodes Switch Rank	Max Change in % Rank
Degree	Decay	0.89	0.80	0.78	0.34
Degree	Closeness	0.82	0.84	0.78	0.33
Degree	Diffusion	0.91	0.96	0.78	0.34
Degree	KatzBon	1.00	1.00	0.74	0.17
Degree	Between	0.84	0.66	0.92	0.31
Decay	Degree	0.89	0.80	0.78	0.34
Decay	Closeness	0.97	0.96	0.26	0.07
Decay	Diffusion	0.93	0.68	0.57	0.17
Decay	KatzBon	0.90	0.66	0.65	0.25
Decay	Between	0.78	0.78	0.91	0.39
Closeness	Degree	0.82	0.84	0.78	0.33
Closeness	Decay	0.97	0.96	0.26	0.07
Closeness	Diffusion	0.84	0.70	0.61	0.19
Closeness	KatzBon	0.83	0.70	0.66	0.24
Closeness	Between	0.67	0.80	0.91	0.38
Diffusion	Degree	0.91	0.96	0.78	0.34
Diffusion	Decay	0.93	0.68	0.57	0.17
Diffusion	Closeness	0.84	0.70	0.61	0.19
Diffusion	KatzBon	0.92	0.92	0.58	0.23
Diffusion	Between	0.80	0.56	0.93	0.46
KatzBon	Degree	1.00	1.00	0.74	0.17
KatzBon	Decay	0.90	0.66	0.65	0.25
KatzBon	Closeness	0.83	0.70	0.66	0.24
KatzBon	Diffusion	0.92	0.92	0.58	0.23
KatzBon	Between	0.84	0.54	0.91	0.35
Between	Degree	0.84	0.66	0.92	0.31
Between	Decay	0.78	0.78	0.91	0.39
Between	Closeness	0.67	0.80	0.91	0.38
Between	Diffusion	0.80	0.56	0.93	0.46
Between	KatzBon	0.84	0.54	0.91	0.35

Table 3: Avg degree 2, Homophily, Decay .5

Cent1	Cent2	Correlation of Rank Vector	Fraction of Sims w same Top Node	Fraction Nodes Switch Rank	Max Change in % Rank
Degree	Decay	0.89	0.80	0.80	0.35
Degree	Closeness	0.82	0.84	0.79	0.37
Degree	Diffusion	0.90	0.88	0.79	0.34
Degree	KatzBon	1.00	1.00	0.75	0.15
Degree	Between	0.79	0.54	0.93	0.34
Decay	Degree	0.89	0.80	0.80	0.35
Decay	Closeness	0.97	0.96	0.33	0.07
Decay	Diffusion	0.92	0.64	0.62	0.18
Decay	KatzBon	0.90	0.68	0.69	0.26
Decay	Between	0.75	0.64	0.91	0.40
Closeness	Degree	0.82	0.84	0.79	0.37
Closeness	Decay	0.97	0.96	0.33	0.07
Closeness	Diffusion	0.83	0.64	0.64	0.19
Closeness	KatzBon	0.83	0.70	0.68	0.28
Closeness	Between	0.65	0.60	0.92	0.39
Diffusion	Degree	0.90	0.88	0.79	0.34
Diffusion	Decay	0.92	0.64	0.62	0.18
Diffusion	Closeness	0.83	0.64	0.64	0.19
Diffusion	KatzBon	0.91	0.80	0.63	0.24
Diffusion	Between	0.75	0.38	0.93	0.47
KatzBon	Degree	1.00	1.00	0.75	0.15
KatzBon	Decay	0.90	0.68	0.69	0.26
KatzBon	Closeness	0.83	0.70	0.68	0.28
KatzBon	Diffusion	0.91	0.80	0.63	0.24
KatzBon	Between	0.80	0.44	0.91	0.36
Between	Degree	0.79	0.54	0.93	0.34
Between	Decay	0.75	0.64	0.91	0.40
Between	Closeness	0.65	0.60	0.92	0.39
Between	Diffusion	0.75	0.38	0.93	0.47
Between	KatzBon	0.80	0.44	0.91	0.36

Table 4: Avg degree 2, Homophily-Bridge, Decay .5

Cent1	Cent2	Correlation of Rank Vector	Fraction of Sims w same Top Node	Fraction Nodes Switch Rank	Max Change in % Rank
Degree	Decay	0.89	0.82	0.80	0.34
Degree	Closeness	0.82	0.86	0.79	0.36
Degree	Diffusion	0.90	0.88	0.79	0.34
Degree	KatzBon	1.00	1.00	0.75	0.15
Degree	Between	0.79	0.56	0.93	0.34
Decay	Degree	0.89	0.82	0.80	0.34
Decay	Closeness	0.97	0.94	0.32	0.07
Decay	Diffusion	0.93	0.66	0.61	0.18
Decay	KatzBon	0.90	0.70	0.68	0.25
Decay	Between	0.75	0.58	0.92	0.40
Closeness	Degree	0.82	0.86	0.79	0.36
Closeness	Decay	0.97	0.94	0.32	0.07
Closeness	Diffusion	0.84	0.66	0.63	0.19
Closeness	KatzBon	0.83	0.72	0.67	0.27
Closeness	Between	0.65	0.58	0.92	0.39
Diffusion	Degree	0.90	0.88	0.79	0.34
Diffusion	Decay	0.93	0.66	0.61	0.18
Diffusion	Closeness	0.84	0.66	0.63	0.19
Diffusion	KatzBon	0.91	0.80	0.62	0.24
Diffusion	Between	0.75	0.40	0.93	0.46
KatzBon	Degree	1.00	1.00	0.75	0.15
KatzBon	Decay	0.90	0.70	0.68	0.25
KatzBon	Closeness	0.83	0.72	0.67	0.27
KatzBon	Diffusion	0.91	0.80	0.62	0.24
KatzBon	Between	0.79	0.46	0.91	0.36
Between	Degree	0.79	0.56	0.93	0.34
Between	Decay	0.75	0.58	0.92	0.40
Between	Closeness	0.65	0.58	0.92	0.39
Between	Diffusion	0.75	0.40	0.93	0.46
Between	KatzBon	0.79	0.46	0.91	0.36