Axioms for Centrality

Paolo Boldi Sebastiano Vigna*
Dipartimento di informatica, Università degli Studi di Milano, Italy

August 9, 2013

Abstract

Given a social network, which of its nodes are more central? This question has been asked many times in sociology, psychology and computer science, and a whole plethora of *centrality measures* (a.k.a. *centrality indices*, or *rankings*) were proposed to account for the importance of the nodes of a network. In this paper, we try to provide a mathematically sound survey of the most important classic centrality measures known from the literature and propose an *axiomatic* approach to establish whether they are actually doing what they have been designed for. Our axioms suggest some simple, basic properties that a centrality measure should exhibit.

Surprisingly, only a new simple measure based on distances, *harmonic centrality*, turns out to satisfy all axioms; essentially, harmonic centrality is a correction to Bavelas's classic *closeness centrality* [4] designed to take unreachable nodes into account in a natural way.

As a sanity check, we examine in turn each measure under the lens of information retrieval, leveraging state-of-the-art knowledge in the discipline to measure the effectiveness of the various indices in locating web pages that are relevant to a query. While there are some examples of this comparisons in the literature, here for the first time we take into consideration centrality measures based on distances, such as closeness, in an information-retrieval setting. The results match closely the data we gathered using our axiomatic approach.

Our results suggest that centrality measures based on distances, which have been neglected in information retrieval in favour of spectral centrality measures in the last years, are actually of very high quality; moreover, harmonic centrality pops up as an excellent general-purpose centrality index for arbitrary directed graphs.

1 Introduction

In the last years, there has been an ever-increasing research activity in the study of real-world complex networks [51] (the world-wide web, the autonomous-systems graph within the Internet, coauthorship graphs, phone call graphs, email graphs and biological networks, to cite but a few). These networks, typically generated directly or indirectly by human activity and interaction (and therefore hereafter dubbed "social"), appear in a large variety of contexts and often exhibit a surprisingly similar structure. One of the most important notions that researchers have been trying to capture in such networks is "node centrality": ideally, every node (often representing an individual) has some degree of influence or importance within the social domain under consideration, and one expects such importance to be reflected in the structure of the social network; centrality is a quantitative measure that aims at revealing the importance of a node.

Among the types of centrality that have been considered in the literature (see [12] for a good survey), many have to do with distances between nodes.[1] Take, for instance, a node in an undirected

*The authors have been supported by the EU-FET grant NADINE (GA 288956).

[1] Here and in the following, by "distance" we mean the length of a shortest path between two nodes.

connected network: if the sum of distances to all other nodes is large, the node under consideration is *peripheral*; this is the starting point to define Bavelas's *closeness centrality* [4] which is the reciprocal of peripherality (i.e., the reciprocal of the sum of distances to all other nodes).

The role played by shortest paths is justified by one of the most well-known features of complex networks, the so-called *small-world* phenomenon. A small-world network [17] is a graph where the average distance between nodes is logarithmic in the size of the network, whereas the clustering coefficient is larger (that is, neighbourhoods tend to be denser) than in a random Erdős-Rényi graph with the same size and average distance.[2] The fact that social networks (whether electronically mediated or not) exhibit the small-world property is known at least since Milgram's famous experiment [38] and is arguably the most popular of all features of complex networks. For instance, the average distance of the Facebook graph was recently established to be just 4.74 [3].

The purpose of this paper is to pave the way for a formal well-grounded assessment of centrality measures, based on some simple guiding principles; we seek notions of centrality that are at the same time *robust* (they should be applicable to arbitrary directed graphs, possibly non-connected, without modifications) and *understandable* (they should have a clear combinatorial interpretation).

With these principles in mind, we shall present and compare the most popular and well-known centrality measures proposed in the last decades. The comparison will be based on a set of *axioms*, each trying to capture a specific trait.

In the last part of the paper, as a sanity check, we compare the measures we discuss in an information-retrieval settings, using the classic GOV2 collection to extract documents satisfying a query and ranking the resulting induced subgraph of relevant documents based solely on centrality.

The results are somehow surprising, and suggest that simple measures based on distances, and in particular *harmonic centrality* (which we introduce formally in this paper) can give better results than some of the most sophisticated indices used in the literature. These unexpected outcomes are the main contribution of this paper, together with the set of axiom we propose, which provide a conceptual framework for understanding centrality measures in a formal way. We also try to give an orderly account of centrality in social and network sciences, gathering scattered results and folklore knowledge in a systematic way.

2 A Historical Account

In this section we sketch the historical development of centrality, focusing on the ten classical centrality measures that we decided to include in this paper: the overall growth of the field is of course much more complex, and the literature contains a myriad of alternative proposals that will not be discussed here.

Centrality is a fundamental tool in the study of social networks: the first efforts to define formally centrality indices were attempted in the late 1940s by the Group Networks Laboratory at M.I.T. directed by Alex Bavelas [4], in the framework of communication patterns and group collaboration [30, 5]; those pioneering experiments all concluded that centrality was related to group efficiency in problem-solving, and agreed with the subjects' perception of leadership. In the following decades, various measures of centrality were employed in a multitude of contexts (to understand political integration in Indian social life [18], to examine the consequences of centrality in communication paths for urban development [45], to analyse their implications to the efficient design of organizations [7, 34], or even to explain the wealth of the Medici family based on their central position with respect to marriages and financial transactions in the 15th century Florence [42]). We can certainly say that the problem of singling out influential individuals in a social group is a holy grail that sociologists have been trying to capture for at least fifty years.

[2] The reader might find this definition a bit vague, and some variants are often spotted in the literature: this is a general well-known problem, also highlighted recently, for example in [32].

Although all researchers agree that centrality is an important structural attribute of social networks, and that it is directly related to other important group properties and processes, there is no consensus on exactly *what* centrality is or on its conceptual foundations, and there is very little agreement on the proper procedures for its measurement [15, 21]: as Freeman observed, "several measures are often only vaguely related to the intuitive ideas they purport to index, and many are so complex that it is difficult or impossible to discover what, if anything, they are measuring" [21].

Freeman acutely remarks that the implicit starting point of all centrality measures is the same: the central node of a star should be deemed more important than the other vertices; paradoxically, it is precisely the unanimous agreement on this requirement that may have produced quite different approaches to the problem. In fact, the center of a star is at the same time

1. the node with largest degree;
2. the node that is closest to the other nodes (e.g., that has the smallest average distance to other nodes);
3. the node through which most shortest paths pass;
4. the node with the largest number of incoming paths of length k, for every k;
5. the node that maximizes the dominant eigenvector of the graph matrix;
6. the node with highest probability in the stationary distribution of the natural random walk on the graph.

These observations lead to corresponding (competing) views of centrality. Degree is probably the oldest kind of measure of importance ever used, being equivalent to majority voting in elections (where $x \to y$ is interpreted as "x voted for y").

The most classical notion of *closeness*, instead, was introduced by Bavelas [4] for undirected, connected networks as the reciprocal of the sum of distances from a given node. Closeness was originally aimed at establishing how much a vertex can communicate without relying on third parties for his messages to be delivered.[3] In the seventies, Nan Lin proposed to adjust the definition of closeness so to make it usable on directed networks that are not necessarily strongly connected [33].

Centrality indices based on the count of shortest paths were formally developed independently by Anthonisse [2] and Freeman [22], who introduced *betweenness* as a measure of the probability that a random shortest path passes through a given node or edge.

Katz's index [27] is based instead on a weighted count of *all* paths coming into a node: more precisely, the weight of a path of length t is β^t, for some *attenuation factor* β, and the score of x is the sum of the weights of all paths coming into x. Of course, β must be chosen so that all the summations converge.

While the above notions of centrality are combinatorial in nature, and based on the discrete structure of the underlying graph, another line of research studies *spectral* techniques (in the sense of linear algebra) to define a measure of centrality.

The earliest known proposal of this kind is due to Seeley [46], who normalized to sum one the row of an adjacency matrix representing the "I like him" relations among a group of children, and assigned a centrality score using the resulting dominant eigenvector. This is actually equivalent to studying the stationary distribution of the Markov chain defined by the natural random walk on the graph. Few years later, Wei [52] proposed the dominant eigenvector of suitable matrices to rank sport teams.

Curiously enough, the most famous among spectral centrality scores is also one of the most recent, PageRank [43]: PageRank was a centrality measure specifically geared toward web graphs, and it was

[3] The notion can also be generalized to a weighted summation of node contributions multiplied by some *discount* functions applied to their distance to a given node [16].

introduced precisely with the aim of implementing it in a search engine (specifically, Google, that the authors of PageRank founded in 1997).

In the same span of years, Jon Kleinberg defined another centrality measure (actually, a ranking algorithm) called HITS [28] (for "Hyperlink-Induced Topic Search"). The idea[4] is that every node of a graph is associated with two importance indices: one (called "authority score") measures how reliable (important, authoritative…) a node is, and another (called "hub score") measures how good the node is in pointing to authoritative nodes, with the two scores mutually reinforcing each other. The result is again the dominant eigenvector of a suitable matrix. SALSA [31] is a more recent and strictly related score based on the same idea, with the difference that it applies some normalization to the matrix.

3 Definitions and conventions

In this paper we consider directed graphs defined by a set N of n nodes and a set $A \subseteq N \times N$ of arcs; we write $x \to y$ when $\langle x, y \rangle \in A$ and call x and y the source and target of the arc, respectively. An arc with the same source and target is called a *loop*.

The *transpose* of a graph is obtained by reversing all arc directions (i.e., it has an arc $y \to x$ for all arcs $x \to y$ of the original graph). A *symmetric graph* is a graph such that $x \to y$ whenever $y \to x$; such a graph is fixed by transposition, and can be identified with a undirected graph, that is, a graph whose arcs are a subset of unordered pairs of nodes (usually called "edges"). A *successor* of x is a node y such that $x \to y$, and a *precedessor* of x is a node y such that $y \to x$. The *outdegree* $d^+(x)$ of a node x is the number of its successors, and the *indegree* $d^-(x)$ is the number of its predecessors.

A *path* (of length k) is a sequence $x_0, x_1, \ldots, x_{k-1}$, where $x_j \to x_{j+1}$, $0 \leq j < k$. A *walk* (of length k) is a sequence $x_0, x_1, \ldots, x_{k-1}$, where $x_j \to x_{j+1}$ or $x_{j+1} \to x_j$, $0 \leq j < k$. A (strongly) connected *component* of a graph is a maximal subset in which every pair of nodes is connected by a walk (path). Components form a partition of the nodes of a graph. A graph is *(strongly) connected* if there is a single (strongly) connected component, that is, for every choice of x and y there is a walk (path) from x to y. A strongly connected component is *terminal* if its nodes have no arc towards other components.

The *distance* $d(x, y)$ from x to y is the length of a shortest path from x to y, or ∞ if no such path exists. The nodes *reachable* from x are the nodes y such that $d(x, y) < \infty$. The nodes *coreachable* from x are the nodes y such that $d(y, x) < \infty$. A node has *trivial* (co)reachable set if the latter contains only the node itself.

The notation $\bar{A}$, where A is a nonnegative matrix, will be used throughout the paper to denote the matrix obtained by ℓ_1-normalizing the rows of A, that is, dividing each element of a row by the sum of the row (null rows are left unchanged). If there are no null rows, $\bar{A}$ is *stochastic*, that is, it is nonnegative and the row sums are all equal to one.

We use Iverson's notation: if P is a predicate, $[P]$ has value 0 if P is false and 1 if P is true [29]; finally, we denote with H_i the i-th harmonic number $\sum_{1 \leq k \leq i} 1/k$.

3.1 Geometric measures

We call *geometric* those measures assuming that importance is a function of the distances. These are actually some of the oldest measures defined in the literature.

[4]To be true, Kleinberg's algorithm works in two phases; in the first phase, one selects a subgraph of the starting webgraph based on the pages that match the given query; in the second phase, the centrality score is computed on the subgraph. Since in this paper we are looking at HITS simply as a centrality index, will simply apply it to the graph under examination.

3.1.1 Indegree

Indegree, the number of incoming arcs $d^-(x)$, can be considered a geometric measure: it is simply the number of nodes at distance one[5]. It is probably the oldest kind of measure of importance ever used, as it is equivalent to majority voting in elections (where $x \to y$ if x voted for y). Indegree has a number of obvious shortcomings (e.g., it is easy to spam), but it is actually a good baseline, and in some cases turned out to provide better results than more sophisticated methods (see, e.g., [19]).

3.1.2 Closeness

Closeness was introduced by Bavelas in the late forties [6]; the closeness of x is defined by

$$\frac{1}{\sum_y d(y,x)}. \tag{1}$$

The intuition behind closeness is that nodes that are more central have smaller distances, and thus a smaller denominator, resulting in a larger centrality. We remark that for this definition to make sense, the graph must be strongly connected. Lacking that condition, some of the denominators will be ∞, resulting in a rank of zero for all nodes which cannot coreach the whole graph.

It was not probably in Bavelas's intentions to apply the measure to directed graphs, and even less to graph with infinite distances, but nonetheless closeness is sometimes "patched" by simply not including unreachable nodes, that is,

$$\frac{1}{\sum_{d(y,x)<\infty} d(y,x)},$$

and assuming that nodes with an empty coreachable set have centrality 0 by definition: this is actually the definition we shall use in the rest of the paper. These apparently innocuous adjustments, however, introduce a strong bias toward nodes with a small coreachable set.

3.1.3 Lin's index

Nan Lin [33] tried to repair the definition of closeness for graphs with infinite distances by weighting closeness using the square of the number of coreachable nodes; his definition for the centrality of a node x with a nonempty coreachable set is

$$\frac{\left|\{y \mid d(y,x) < \infty\}\right|^2}{\sum_{d(y,x)<\infty} d(y,x)}.$$

The rationale behind this definitions is the following: first, we consider closeness not the inverse of a sum of distances, but rather the inverse of the *average* distance, which entails a first multiplication by the number of coreachable nodes. This change normalizes closeness across the graph. Now, however, we want nodes with a larger coreachable set to be more important, given that the average distance is the same, so we multiply again by the number of coreachable nodes. Nodes with an empty coreachable set have centrality 1 by definition.

Lin's index was (somewhat surprisingly) ignored in the following literature. Nonetheless, it seems to provide a reasonable solution for the problems caused by the definition of closeness.

[5]Most centrality measures proposed in the literature were actually described only for undirected, connected graphs. Since the study of web graphs and online social networks has posed the problem of extending centrality concepts to networks that are directed, and possibly not strongly connected, in the rest of this paper we consider measures depending on the *incoming* arcs of a node (e.g., incoming paths, left dominant eigenvectors, distances from all nodes to a fixed node). If necessary, these measures can be called "negative", as opposed to the "positive" versions obtained by taking the transpose of the graph.

3.1.4 Harmonic centrality

As we noticed, the main problem of closeness lies in the presence of pairs of unreachable nodes. We thus get inspiration from Marchiori and Latora [35], who, faced with the problem of providing a sensible notion of "average shortest path" for a generic directed network, propose to replace the average distance with the *harmonic mean of all distances*. Indeed, in case a large number of pairs of nodes are not reachable, the average distance between reachable pairs can be misleading: a graph might have a very low average distance, while it is almost completely disconnected (e.g., a perfect matching has average distance exactly one). The harmonic mean has the useful property of handling ∞ cleanly (assuming, of course, that $\infty^{-1} = 0$). For example, a perfect matching has harmonic mean of distances $n - 1$.

In general, for each graph-theoretical notion based on arithmetic averaging or maximization there is an equivalent notion based on the harmonic mean. If we consider closeness the reciprocal of a denormalized average of distances, it is natural to consider also the reciprocal of a denormalized harmonic mean of distances. We thus define the *harmonic centrality* of x as[6]

$$\sum_{y \neq x} \frac{1}{d(y,x)} = \sum_{d(y,x)<\infty, y \neq x} \frac{1}{d(y,x)}. \tag{2}$$

The difference with (1) might seem minor, but actually it is a radical change. Harmonic centrality is strongly correlated to closeness centrality in simple networks, but naturally also accounts for nodes y that cannot reach x. Thus, it can be fruitfully applied to graphs that are not strongly connected.

3.2 Spectral measures

Spectral measures compute the left dominant eigenvector of some matrix derived from the graph, and depending on how the matrix is modified before the computation we can obtain a number of different measures. Existence and uniqueness of such measures is usually derivable by the theory of nonnegative matrices started by Perron and Frobenius [8]; we will however avoid to discuss such issues, as there is a large body of established literature about the topic. All observations in this section are true for strongly connected graphs; the modifications for graphs that are not strongly connected can be found in the cited references.

3.2.1 The left dominant eigenvector

The first and most obvious spectral measure is the left dominant eigenvector of the adjacency matrix. Indeed, the dominant eigenvector can be thought as the fixed point of an iterated computation in which every node starts with the same score, and then updates its score with the sum of its predecessors. The vector is then normalized, and the process repeated until convergence.

The usage of dominant eigenvectors to find important nodes in matrices of entities can be traced at least back to Wei's Master thesis [52]. Wei's thesis was then popularized by Kendall, and the technique is actually known in the literature about ranking of sport teams as "Kendall–Wei ranking".[7]

Dominant eigenvectors do not react very well to the lack of strong connectivity. Depending on the dominant eigenvalue of the strongly connected components, the dominant eigenvector might or might not be nonzero on non-terminal components (a detailed characterization can be found in [8]).

[6]We remark that Tore Opsahl already in a March 2010 blog posting observed that in an undirected graph with several disconnected components the inverse of the harmonic mean of distances offers a better notion of centrality than closeness, as it weights less elements that belong to smaller components.

[7]It was rediscovered as a generic way of ranking graphs by Bonacich [11].

3.2.2 Seeley's index

The dominant eigenvector rationale can be slighly amended with the observation that the update rule we described can be thought of as if each node gives away its score to its successors: or even, that each node has a *reputation* and is giving its reputation to its successors so that they can build their own.

Once we take this viewpoint, it is clear that it is not very sensible to give away the same amount of reputation to everybody: it is more reasonable to *divide* equally reputation among our successors. From a linear-algebra viewpoint, this amounts to normalizing each row of the adjacency matrix using the ℓ_1 norm.

This approach was advocated by Seeley [46] for computing the popolarity among groups of children, given a graph representing whether each child liked or not another one. The matrix resulting from the ℓ_1-normalization process is actually stochastic, so the score can be interpreted as the probability distribution of the stationary state of a Markov chain. In particular, if the underlying graph is symmetric Seeley's index collapses to the degree (modulo normalization) because of the very well-known characterization of the stationary distribution of the natural random walk on a symmetric graph.

Also Seeley's index does not react very well to the lack of strong connectivity, but in a more predictable way: the only nodes with a nonzero rank are those belonging to terminal components.

3.2.3 Katz's index

Katz introduced his celebrated index [27] using a summation over all paths coming into a node, but weighting each path so that the summation would actually be finite. Due to the interplay between the powers of the adjacency matrix and the number of paths connecting two nodes, Katz's index can be expressed as

$$\boldsymbol{k} = \mathbf{1}\sum_{i=0}^{\infty}\beta^i A^i.$$

For the summation above to be finite, the *attenuation factor* β must be smaller than $1/\lambda$, where λ is the dominant eigenvalue of A.

Katz immediately noted that the index was actually expressible using linear algebra operations:

$$\boldsymbol{k} = \mathbf{1}(1-\beta A)^{-1}.$$

It took some more time to realize that, due to Brauer's theorem on the displacement of eigenvalues [14], Katz's index is actually the left dominant eigenvector of a *perturbed matrix*

$$\beta\lambda A + (1-\beta\lambda)\boldsymbol{e}^T\mathbf{1}, \tag{3}$$

where $\boldsymbol{e}$ is a right dominant eigenvector such that $\mathbf{1}\boldsymbol{e}^T = \lambda$ [50]. An easy generalization (actually suggested by Hubbell [25]) replaces the vector $\mathbf{1}$ with some preference vector $\boldsymbol{v}$, so that paths are also weighted differently depending on their starting node.[8]

If the underlying graph is strongly connected, the limit of Katz's index when $\beta \to 1/\lambda$ is exactly the dominant eigenvector [50]. This is also true under the much more general condition that the dominant eigenvalue of A is *semisimple* [37], but in that case the limit is a specific dominant eigenvector that depends on the preference vector $\boldsymbol{v}$.

[8] We must note that the original definition of Katz's index is $\mathbf{1}A\sum_{i=0}^{\infty}\beta^i A^i = \mathbf{1}/\beta\sum_{i=0}^{\infty}\beta^{i+1}A^{i+1} = (\mathbf{1}/\beta)\sum_{i=0}^{\infty}\beta^i A^i - \mathbf{1}/\beta$. This additional multiplication by A is somewhat common in the literature, even for PageRank; clearly, it alters the order induced by the ranking only when there is a nonuniform preference vector. Our discussion can be easily adapted for this version.

3.2.4 PageRank

PageRank [43] is one of the most discussed and quoted spectral indices in use today, mainly because of its alleged use in Google's ranking algorithm.[9]

By definition, PageRank is the left dominant eigenvector (i.e., the stationary distribution) $\boldsymbol{p}$ of the Markov chain

$$\alpha \bar{A} + (1-\alpha)\mathbf{1}^T \boldsymbol{v},$$

where again $\bar{A}$ is the ℓ_1-normalized adjacency matrix of the graph, and $\boldsymbol{v}$ is a *preference vector* (which must be a distribution). The reader will immediately notice the similarity with (3): indeed, we can work backwards and rewrite PageRank as

$$\boldsymbol{p} = \boldsymbol{v}\bigl(1 - \alpha \bar{A}\bigr)^{-1},$$

leading to

$$\boldsymbol{p} = \boldsymbol{v} \sum_{i=0}^{\infty} \alpha^i \bar{A}^i,$$

which shows immediately that Katz's index and PageRank differ only by the ℓ_1 normalization applied to A, similarly to the difference between the dominant eigenvector and Seeley's index.

Analogously to what happens with Katz's index, the limit of PageRank when α goes to 1 is exactly the dominant eigenvector of $\bar{A}$, that is, Seeley's index [10, 24]. The statement is always true, because in stochastic matrices the dominant eigenvalue is always semisimple [8], but if the graph is not strongly connected the limit is a specific dominant eigenvector that depends on the preference vector $\boldsymbol{v}$ [10].

3.2.5 HITS

Kleinberg introduced his celebrated HITS algorithm [28] using the web metaphore of "mutual reinforcement": a page is authoritative if it is pointed by many good *hubs*—pages which contain good list of authoritative pages—, and a hub is good if it points to authoritative pages. This suggests an iterative process that computes at the same time an authoritativeness score $\boldsymbol{a}_i$ and a "hubbiness" score $\boldsymbol{h}_i$ starting with $\boldsymbol{a}_0 = \mathbf{1}$, and then applying the update rule

$$\boldsymbol{h}_{i+1} = \boldsymbol{a}_i A^T \qquad \boldsymbol{a}_{i+1} = \boldsymbol{h}_{i+1} A.$$

This process converges to the left dominant eigenvector of the matrix $A^T A$, which gives the final authoritativeness score, which is the score we label with "HITS" throughout the paper.[10]

Inverting the process, and considering the left dominant eigenvector of the matrix AA^T, gives the final hubbiness score. The two vectors are actually the left and right *singular vectors* associated with the largest *singular value* in the singular-value decomposition of A. Note also that hubbiness is the positive version of authoritativeness.

3.2.6 SALSA

Finally, we consider SALSA, a measure introduced by Lempel and Moran [31] always using the metaphore of mutual reinforcement between authoritativeness and hubbiness, but ℓ_1-normalizing the matrices A and A^T. We start with $\boldsymbol{a}_0 = \mathbf{1}$ and proceed with

$$\boldsymbol{h}_{i+1} = \boldsymbol{a}_i \overline{A^T} \qquad \boldsymbol{a}_{i+1} = \boldsymbol{h}_{i+1} \bar{A}.$$

[9] The reader should be aware, however, that the literature about the actual effectiveness of PageRank in information retrieval is rather scarce, and comprises mainly negative results such as [40] and [19].

[10] As discussed in [20], the dominant eigenvector may not be unique; equivalently, the limit of the recursive definition given above may depend on the way the authority and hub scores are initialized. Here we consider the result of the iterative process starting with $\boldsymbol{a}_0 = \mathbf{1}$.

We remark that this normalization process is analogous to the one that moves us from the dominant eigenvector to Seeley's index, or from Katz's index to PageRank.

Similarly to what happens with Seeley's index on symmetric graphs, SALSA does not need such an iterative process to be computed.[11] First, one computes the connected components of the symmetric graph induced by the matrix $A^T A$; in this graph, x and y are adjacent if x and y have some common predecessor in the original graph. Then, the score of a node is the ratio between its indegree and the sum of the indegrees of nodes in the same component, multiplied by the ratio between the component size and n. Thus, contrarily to HITS, a single linear scan of the graph is sufficient to compute SALSA, albeit the computation of the intersection graph requires time proportional to $\sum_x d^+(x)^2$.

3.3 Path-based measures

Path-based measures exploit not only the existence of a shortest paths, but actually take into examination all shortest paths (or all paths) coming into a node. We remark that indegree can be considered a path-based measure, as it is the equivalent to the number of incoming paths of length one.

3.3.1 Betweeness

Betweenness centrality was introduced by Anthonisse [2] for edges, and then rephrased by Freeman for nodes [22]. The idea is to measure the probability that a random shortest path passes through a given node: if σ_{yz} is the number of shortest paths going from y to z, and $\sigma_{yz}(x)$ is the number of such paths that pass through x, we define the *betweeness* of x as

$$\sum_{y,z\neq x,\sigma_{yz}\neq 0} \frac{\sigma_{yz}(x)}{\sigma_{yz}}.$$

The intuition behind betweeness is that if a large fraction of shortest paths passes through x, than x is an important point of passage for the network. Indeed, removing nodes in betweeneess order causes a very quick disruption of the network [9].

3.3.2 Spectral measures as path-based measures

It is a general observation that all spectral measures can actually be interpreted as path-based measures, as they depend on taking the limit of some summations of powers of A, or on the limit of powers of A, and in both cases we can express these algebraic operations in terms of suitable paths.

For instance, the left dominant eigenvector of a nonnegative matrix can be computed with the power method by taking the limit of $\mathbf{1}A^k/\|\mathbf{1}A^k\|$ for $k \to \infty$. Since, however, $\mathbf{1}A^k$ is a vector associating with each node the number of paths of length k coming into the node, we can see that dominant eigenvector expresses the relative growth of the number of paths coming into each node as their length increases.

Analogously, Seeley's index can be computed (modulo a normalization factor) by taking the limit of $\mathbf{1}\bar{A}^k$ (in this case, the ℓ_1 norm cannot grow, so we do not need to renormalize at each iteration). The vector $\mathbf{1}\bar{A}^k$ has the following combinatorial interpretation: it assigns to each x the sums of the *weights* of the paths coming into x, where the weight of a path $x_0, x_1, \ldots, x_t$ is

$$\prod_{i=0}^{t-1} \frac{1}{d^+(x_i)}. \tag{4}$$

[11] This property, which appears to be little known, is proved in Proposition 2 of the original paper [31].

When we switch to the attenuated versions of the previous indices (that is, Katz's index and PageRank), we switch from limits to infinite summations and at the same time modify the weight of a path of length t with β^t or α^t. Actually, the Katz index of x was originally defined as the summation over all t of the number of paths of length t coming into x multiplied by β^t, and PageRank is the summation over all paths coming into x of the weight (4) multiplied by α^t.

The reader can easily work out similar definitions for HITS and SALSA, which depend on a suitable definition of alternate "back-and-forth path" (see, e.g., [13])

4 Axioms for Centrality

The comparative evaluation of centrality measures is a challenging, difficult, arduous task, for many different reasons. The datasets that are classically used in social sciences are very small (typically, some tens of nodes) and it is hard to draw conclusions out of them. Nonetheless, some attempts were put forward, like [48]; sometimes, the attitude was actually to provide evidence that different measures highlight different kinds of centralities and are therefore equally incomparably interesting [23]. Whether the latter attitude is the only sensible conclusion or not is debatable. While it is clear that the notion of centrality, in its vagueness, can be interpreted differently giving rise to many good but incompatible measures, we will provide evidence that some measures tend to reward nodes that are in no way central.

If results obtained on small corpora may be misleading, a comparison on larger corpora is much more difficult to deal with, due to the lack of ground truth and to the unavailability of implementations of efficient algorithms to compute the measures under consideration (at least in the cases where efficient, possibly approximate, algorithms do exist). Among the few attempts that try a comparison on large networks we cite [49] and [41], that nevertheless focus only on web graphs and on a very limited number of centrality indices.

In this paper, we propose to understand (part of) the behaviour of a centrality measure using a set of axioms. While, of course, it is not sensible to prescribe a set of axioms that *define* what centrality should be (in the vein of Shannon's definiton of entropy [47] or Altman and Tennenholtz axiomatic definition of Seeley's index [1][12]), as different indices serve different purposes, it is reasonable to set up some *necessary* axioms that an index should satisfy to behave predictably and follow our intuition.

The other interesting aspect of defining axioms is that, even if one does not believe they are really so discriminative or necessary, they provide a very specific, formal, provable piece of information about a centrality measure that is much more precise than folklore intuitions like "this centrality is really correlated to indegree" or "this centrality is really fooled by cliques". We believe that a theory of centrality should exactly provide this kind of compact, meaningful, reusable information (in the sense that it can be used to prove other properties). This is indeed what happens, for example, in topology, where the information that a space is T_0, rather than T_1, is a compact way to provide a lot of information about the structure of the space.

Definining such axioms is a delicate matter. First of all, the semantics of the axioms must be very clear. Second, the axioms must be evaluable in an exact way on the most common centrality measures. Third, they should be formulated avoiding the trap of small, finite (counter)examples, on which many centrality measures collapse (e.g., using an asymptotic definition). We assume from the beginning that the centrality measures under examination are invariant by isomorphism, that is, that they depend just on the structure of the graph, and not on particular labelling chosen for each node.

To meet these constraints, we propose to study the reaction of centrality measures to *change of size*, to *(local) change of density* and their *monotonicity with respect to arc additions*. We expect that nodes belonging to larger groups, when all other parameters are fixed, should be more important,

[12] The authors claim to formalize PageRank [44], but they do not consider the damping factor (equivalently, they are setting $\alpha = 1$), so they are actually formalizing Seeley's venerable index [46].

and that nodes with a denser neighbourhood (i.e., having more friends), when all other parameter are fixed, should also be more important. We also expect that adding an arc should increase the importance of the target.

How can we actually determine if this happens in an exact way, and possibly in an asymptotic setting? To do so, we need to do something entirely new—evaluating *exacly* (i.e., in algebraic closed form) all measures of interest on all nodes of some representative classes of networks.

4.1 The size axiom

An obvious approach to reduce to a minimum the amount of computation is using strongly connected *vertex-transitive*[13] graphs as basic building blocks: these graphs have as much symmetry as possible, which entails a simplification of the computations. Finally, since we want to compare density, the obvious choice is to pick the *densest* strongly connected vertex-transitive graph, the clique, and the *sparsest* strongly connected, the directed cycle. Choosing two graphs at the extreme of the density spectrum guarantees that best possible highlight of the reaction of centrality measures to densities. Moreover, k-cliques and directed p-cycles obviously exist for every k and p (this might not happen for more complicated structures, e.g., a cubic graph).

Let us consider a graph made by a k-clique and a p-cycle (see the figure in Table 1).[14] Because of invariance by isomorphism, all nodes of the clique has the same score, and all nodes of the cycle have the same score. But which nodes are more important? Probably everybody would answer that if $p = k$ the elements on the clique are more important, and indeed this axiom is so trivial that is satisfied by almost any measure we are aware of. But we are interested in assessing the sensitivity to *size*, and thus we state our first axiom:

Definition 1 (Size axiom) *Consider the graph $S_{k,p}$ made by a k-clique and a directed p-cycle. A centrality measure satisfies the* size axiom *if for every k there is a P_k such that for all $p \geq P_k$ in $S_{k,p}$ the centrality of a node of the p-cycle is strictly larger than the centrality of a node of the k-clique, and if for every p there is a K_p such that for all $k \geq K_p$ in $S_{k,p}$ the centrality of a node of the k-clique is strictly larger than the centrality of a node of the p-cycle.*

Intuitively, when $p = k$ we do expect nodes of the cycle to be less important than nodes of the clique. (Note that because of vertex transitivity and invariance by isomorphism we can speak of the "centrality of a node of the p-cycle", without specifying which node.) The rationale behind the case $k \to \infty$ is rather obvious: the denser community is also getting larger, and thus its members are expected to become even more important.

On the other hand, if the cycle becomes very large (more precisely, when its size goes to infinity), its nodes are still part of a very large (albeit badly connected) community, and we expect them to achieve at some size greater importance than the node of a fixed-size community, no matter how dense it can be.

Since one might devise some centrality measures that satisfy the size axiom for p and not for k, which we would not certaintly want to pass our screening, stating both properties in Definition 1 gives us a finer granularity and avoids pathological cases.

4.2 The density axiom

Designing an axiom for density is a more delicate issue, since we must be able to define an increase of density "with all other parameters fixed", including size. Let us start ideally from a graph made

[13] A graph is vertex-transitive if for every nodes x and y there is an automorphism exchanging x and y.

[14] The graph is of course disconnected. It is a common theme of this work that centrality measures should work also on graphs that are not strongly connected, for the very simple reason that we meet this kind of graphs in the real world, the web being an obvious example.

by a directed k-cycle and a directed p-cycle, and connect a vertex x of the k-cycle with a vertex y of the p-cycle through a bidirectional arc, the *bridge*. If $k = p$, the vertices x and y are symmetric, and thus must have necessarily the same ranking. Now, we increase the density of the k-cycle as much as possible, turning it into a k-clique (see the figure in Table 2). Note that this change of density is local to x, as the degree of y has not changed. We are thus *strictly increasing the local density around x, leaving all other parameters fixed*, and in these circumstances we expect that the ranking x increases.

Definition 2 (Density axiom) *Consider the graph $D_{k,p}$ made by a k-clique and a p-cycle ($p, k \geq 3$) connected by a bidirectional bridge $x \leftrightarrow y$, where x is a node of the clique and y is a vertex of the cycle. A centrality measure satisfies the* density axiom *if for $k = p$ the centrality of x is strictly larger than the centrality of y.*

Note that our axiom does not specify any constraint when $k \neq p$. While studying the behaviour of the graph $D_{k,p}$ of the previous definition when $k \neq p$ shades some lights of the inner behaviour of centrality measures, it is essential, in an axiom asserting the sensitivity to density, that size is not involved.

In our proofs for the density axiom, we actually let k and p be independent parameters (even if the axiom requires $k = p$) because in this way we can compute the *watershed*, that is, the value of k (expressed as a function of p) at which the axiom becomes true (if any). The watershed can give some insight as to how badly a measure can miss to satisfy the density axiom.

4.3 The monotonicity axiom

Finally, we propose a seemingly trivial axiom that specifies strictly monotonic behaviour upon the addition of an arc:

Definition 3 (Monotonicity axiom) *Consider an arbitrary graph G and a pair of nodes x, y such that $x \not\to y$. A centrality measure satisfies the* monotonicity axiom *if when we add $x \to y$ the centrality of y increases.*

Actually, in some sense this axiom *is* trivial: it is satisfied by essentially all centrality measures we consider *on strongly connected graphs*. Thus, it is an excellent test to verify that a measure is able to handle correctly partially disconnected graphs.

We remark that the reader might be tempted to define a *weak* monotonicity axiom which just require the rank of y to be nondecreasing. However, the constant ranking associating one to every node of every network would satisfy such an axiom, which makes it not very interesting for our goals.

5 Proofs and Counterexamples

We have finally reached the core of this paper: given that we are considering eleven centralities and three axioms, we have to verify 33 statements. For the size and density axioms, we compute in closed form the values of all measures, from which we can derived the desired results, whereas for the monotonicity axiom we provide directly proofs or counterexamples.

We remark that in all our tables we use the proportionality symbol $\propto$ to mark values that have been rescaled by a common factor to make them more readable.

5.1 Size

Table 1 provides scores for the graph $S_{p,k}$, from which we can check whether the size axiom is satisfied. The scores are all immediately computable from the basic definitions, because as we noticed $S_{k,p}$ is highly symmetrical and so there are actually only two scores—the score of a node of the clique and the score of a node of the cycle. Note that in the case of some spectral centrality measure there are actually several possible solutions, in which case we use the one returned by the power method starting from the uniform vector.

5.2 Density

Table 2 provides scores for the graph $D_{p,k}$. Being the graph strongly connected, there is no uniqueness issue. While the computation of geometric and path-based centrality measures, being just a matter of finite summations, is tedious but rather straightforward, spectral indices require some more care. In the case of $D_{k,p}$, we have to write down parametric equations expressing the matrix computation that defines the centrality, and solve them. As noted before, we prefer to perform the computation with two independent parameters k and p (even if the axiom requires $k = p$) because in this way we can compute the watershed.

In all cases, we can always use the bounds imposed by symmetry to write down just a small number of variables: c for the centrality of an element of the clique, ℓ for the clique bridge ("left"), r for the cycle bridge ("right"), and some function $t(d)$ of the distance from the cycle bridge for the nodes of the cycle (with $0 < d < p$), with the condition $t(0) = r$.

5.2.1 The left dominant eigenvector

In this case, the equations are given by the standard eigenvalue problem of the adjacency matrix:

$$\begin{aligned}\lambda\ell &= r + (k-1)c\\ \lambda c &= \ell + (k-2)c\\ \lambda r &= \ell + \frac{r}{\lambda^{p-1}},\end{aligned}$$

subject to the condition that we choose the λ with maximum absolute value. Note that in the case of the last equation we "unrolled" the equations about the elements of the cycle, $\lambda t(d+1) = t(d)$. Solving the system and choosing $c = 1/(\lambda - k + 1)$ gives the solutions found in Table 2.

Since for nonnegative matrices the dominant eigenvalue is monotone in the matrix entries, $\lambda \geq k-1$, because the k-clique has dominant eigenvalue equal to $k-1$. On the other hand, $\lambda \leq k$ by row-sum bounds, and the eigenvalue equations have no solution for $\lambda = k-1$, so we conclude that $k - 1 < \lambda \leq k$.

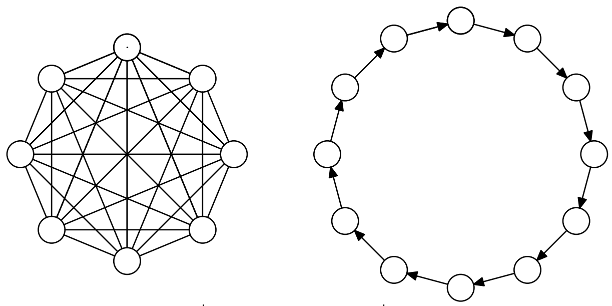
5.2.2 Katz's index

In this case, the equations can be obtained by the standard technique of "taking one summand out", that is, writing

$$\boldsymbol{k} = \mathbf{1}\sum_{i=0}^{\infty}\beta^i A^i = \mathbf{1} + \mathbf{1}\sum_{i=1}^{\infty}\beta^i A^i = \mathbf{1} + \left(\mathbf{1}\sum_{i=0}^{\infty}\beta^i A^i\right)\beta A = \mathbf{1} + \boldsymbol{k}\beta A.$$

The equations are then

$$\begin{aligned}\ell &= 1 + \beta r + \beta(k-1)c\\ c &= 1 + \beta\ell + \beta(k-2)c\\ r &= 1 + \beta\ell + \beta\left(\frac{1-\beta^{p-1}}{1-\beta} + \beta^{p-1}r\right),\end{aligned}$$



Centrality	k-clique	p-cycle
Degree	$k-1$	1
Harmonic	$k-1$	H_{p-1}
Closeness	$\frac{1}{k-1}$	$\frac{2}{p(p-1)}$
Lin	$\frac{k^2}{k-1}$	$\frac{2p}{p-1}$
Betweenness	0	$\frac{(p-1)(p-2)}{2}$
Dominant $\propto$	1	0
Seeley $\propto$	1	1
Katz	$\frac{1}{1-(k-1)\beta}$	$\frac{1}{1-\beta}$
PageRank $\propto$	1	1
HITS $\propto$	1	0
SALSA $\propto$	1	1

Table 1: Centrality scores for the graph $S_{k,p}$. H_i denotes the i-th harmonic number. The parameter β is Katz's attenuation factor.

where again we "unrolled" the equations about the elements of the cycle, as we would just have $t(d+1) = 1 + \beta t(d)$, so

$$t(d) = \frac{1-\beta^d}{1-\beta} + \beta^d r.$$

The explicit values of the solutions are quite ugly, so we present them in Table 2 as a function of the centrality of the clique bridge ℓ.

5.2.3 PageRank

To simplify the computation, we use **1**, rather than $\mathbf{1}/(k+p)$, as preference vector (the result obtained is obviously the same up to proportionality). We use the same technique employed in the computation of Katz's index, leading to

$$\begin{aligned}
\ell &= 1-\alpha + \frac{1}{2}\alpha r + \alpha c \\
c &= 1-\alpha + \frac{\alpha}{k}\ell + \alpha\frac{k-2}{k-1}c \\
r &= 1-\alpha + \frac{\alpha}{k}\ell + \alpha\left(1-\alpha^{p-1} + \frac{1}{2}\alpha^{p-1}r\right),
\end{aligned}$$

noting once again that unrolling the equation of the cycle $t(1) = 1-\alpha+\alpha r/2$ and $t(d+1) = 1-\alpha+\alpha t(d)$ for $d > 1$ we get

$$t(d) = 1-\alpha^d + \frac{1}{2}\alpha^d r.$$

The explicit values for PageRank are even uglier than those of Katz's index, so again we present them in Table 2 as a function of the centrality of the clique bridge ℓ.

5.2.4 Seeley's index

This is a freebie, as we can just compute PageRank's limit when $\alpha \to 1$.

5.2.5 HITS

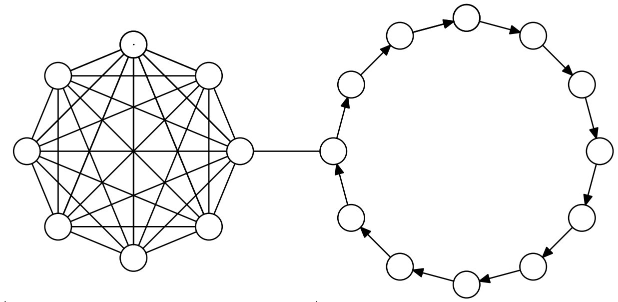
In this case we write down an eigenvalue problem for $A^T A$. We have

$$\begin{aligned}
\mu c &= (k-1)c + (k-2)^2 c + (k-2)\ell + r \\
\mu \ell &= k\ell + (k-1)(k-2)c + t \\
\mu r &= 2r + (k-1)c \\
\mu t &= t + \ell.
\end{aligned}$$

By normalizing the result so that $c = \mu^2 - \mu(k+1) + k - 1$, and noting that the dominant eigenvalue of $A^T A$ is the square of the dominant eigenvalue of A, we obtain the complex but somewhat readable values shown in Table 2. Note that p has no role in the solution, because $A^T A$ can be decomposed into two independent blocks, one of which is an identity matrix corresponding to all elements of the cycle except for the first two.

5.2.6 SALSA

It is easy to check that the components of the intersection graph of precedessors are given by the clique together with the cycle bridge and its successor, and then by one component for each node of the cycle. The computation of the scores is then trivial using the non-iterative rules.



Centrality	Clique	Clique bridge	Cycle bridge	Cycle ($d > 0$ from the bridge)	Watershed
Degree	$k-1$	k	2	1	—
Harmonic	$k-2+H_{p+1}$	$k-1+H_p$	$1+\frac{k-1}{2}+H_{p-1}$	$\frac{1}{d+1}+\frac{k-1}{d+2}+H_{p-1}$	—
Closeness	$\frac{1}{k-1+2p+p(p-1)/2}$	$\frac{1}{k-1+p+p(p-1)/2}$	$\frac{1}{2k-1+p(p-1)/2}$	$\frac{1}{k(d+2)-1+p(p-1)/2}$	$k \le p$
Betweenness	0	$2p(k-1)$	$2k(p-1)+\frac{(p-1)(p-2)}{2}$	$2k(p-2)+\frac{(p-1)(p-2)}{2}$	$k \le \frac{p^2+p+2}{4}$
Dominant	$\frac{1}{\lambda-k+1}$	$1+\frac{1}{\lambda-k+1}$	$1+\lambda$	$\frac{1+\lambda}{\lambda^d}$	—
Seeley $\propto$	$k-1$	k	2	1	—
Katz $\propto$	$\frac{1+\beta\ell}{1-\beta(k-2)}$	ℓ	$\frac{1}{1-\beta}+\frac{\beta}{1-\beta^p}\ell$	$\frac{1}{1-\beta}+\frac{\beta^{d+1}}{1-\beta^p}\ell$	—
PageRank $\propto$	$\frac{(k-1)(k-\alpha k+\alpha\ell)}{k(k-1-\alpha(k-2))}$	ℓ	$2+2\frac{\alpha\ell-k}{k(2-\alpha^p)}$	$1+\alpha^d\frac{\alpha\ell-k}{k(2-\alpha^p)}$	—
HITS $\propto$	$\lambda^4-\lambda^2(k+1)+k-1$	$(k-1)(k-2)(\lambda^2-1)$	$\lambda^6-(k^2-2k+4)\lambda^4$ $+(3k^2-7k+6)\lambda^2-(k-1)^2$	$[d=1](k-1)(k-2)$	—
SALSA $\propto$	$(k-1)(k+2)$	$k(k+2)$	$2(k+2)$	$k+2+[d\neq 1](k^2-2k+2)$	—

Table 2: Centrality scores for the graph $D_{k,p}$. The parameter β is Katz's attenuation factor, whereas $\alpha \in [0\,..\,1]$ is PageRank's damping factor. The value $k-1 < \lambda \le k$ is the dominant eigenvalue of the adjacency matrix A. Lin's centrality is omitted because it is proportional to closeness (the graph being strongly connected).

Armed with our closed-form description of the scores, we have now to prove whether the density axiom actually holds, that is, whether $\ell > r$ when $k = p$. In Table 2 we report the *watershed*, that is, the point at which the axiom becomes true. When there is no watershed, the axiom is true for every $k, p \geq 3$. Note that the determination of the watershed is trivial in almost all cases. We will now discuss the remaining cases.

Theorem 1 *HITS satisfies the density axiom.*

Proof. As we have seen, we can normalize the solution to the HITS equation so that

$$\begin{aligned}\ell &= (k-1)(k-2)(\mu-1)\\ r &= \mu^3 - (k^2-2k+4)\mu^2 + (3k^2-7k+6)\mu - (k-1)^2\end{aligned}$$

Moreover, the characteristic polynomial can be computed explicitly from the set of equations and some simple observations on the eigenvectors for the eigenvalue 1:

$$p(\mu) = \big(\mu^4 - (k^2-2k+6)\mu^3 + (5k^2-12k+15)\mu^2 - (6k^2-16k+14)\mu + k^2 - 2k + 1\big)(1-\mu)^{k+p-4}.$$

The largest eigenvalue μ_0 satisfies the inequation $(k-1)^2 \leq \mu_0 \leq k^2 - 2k + 5/4$ for every $k \geq 9$, as shown below (the statement of the theorem can be verified in the remaining cases by explicit computation, as it does not depend on p). Using the stated upper and lower bounds on μ_0, we can say that

$$\begin{aligned}\ell - r &= (k-1)(k-2)(\mu-1) - (\mu^3 - (k^2-2k+4)\mu^2 + (3k^2-7k+6)\mu - (k-1)^2)\\ &= -\mu^3 + (k^2-2k+4)\mu^2 - (2k^2-4k+4)\mu + k - 1\\ &\geq -\left(k^2-2k+\frac{5}{4}\right)^3 + (k^2-2k+4)(k-1)^4 - (2k^2-4k+4)\left(k^2-2k+\frac{5}{4}\right) + k - 1\\ &= \frac{1}{4}k^4 - k^3 - \frac{19}{16}k^2 + \frac{43}{8}k - \frac{253}{64},\end{aligned}$$

which is positive for $k > 4$.

We are left to prove the bounds on μ_0. The lower bound can be easily obtained by monotonicity of the dominant eigenvalue in the matrix entries, because the dominant eigenvalue of a k-clique is $k-1$. For the upper bound, first we observe that μ_0 can be computed explicitly (as it is the solution of a quartic equation) and using its expression in closed form it is possible to show that $\lim_{k\to\infty} \mu_0 = (k-1)^2$. This guarantees that the bound $\mu_0 \leq k^2 - 2k + 5/4$ is true ultimately. To obtain an explicit value of k after which the bound holds true, observe that $k^2 - 2k + 5/4 = \mu_0$ implies $q(k) = p(\mu_0) = 0$, where $q(k) = p(k^2 - 2k + 5/4)$. Computing the Sturm sequence associated to $q(k)$ one can prove that $q(k)$ has no zeroes for $k \geq 9$, hence our lower bound on k. ∎

Theorem 2 *Katz's index satisfies the density axiom.*

Proof. Recall that the equations for Katz's index are

$$\begin{aligned}\ell &= 1 + \beta r + \beta(k-1)c\\ c &= 1 + \beta\ell + \beta(k-2)c\\ r &= 1 + \beta\ell + \beta\left(\frac{1-\beta^{p-1}}{1-\beta} + \beta^{p-1} r\right).\end{aligned}$$

First, we remark that as $\beta \to 1/\lambda$ Katz's index tends to the dominant eigenvector, so ultimately $\ell > r$. Thus, by continuity, we just need to show that $\ell = r$ never happens in the range of our parameters. If we solve the equations above for c, ℓ and r and impose $\ell = r$, we obtain

$$p = \frac{\ln \dfrac{\beta^2 + k - 2}{k - 1}}{\ln \beta}.$$

Now observe that

$$\beta \le \frac{\beta^2 + k - 2}{k - 1}$$

is always true for $\beta \le 1$ and $k \ge 3$. This implies that under the same conditions $p \le 1$, which concludes the proof. ▮

Theorem 3 *PageRank with constant preference vector satisfies the density axiom.*

Proof. The proof is similar to that of Theorem 2. Recall that the equations for PageRank are

$$\begin{aligned}
\ell &= 1 - \alpha + \frac{1}{2}\alpha r + \alpha c \\
c &= 1 - \alpha + \frac{\alpha}{k}\ell + \frac{\alpha(k-2)}{k-1}c \\
r &= 1 - \alpha + \frac{\alpha}{k}\ell + \alpha\left(1 - \alpha^{p-1} + \frac{1}{2}\alpha^{p-1} r\right).
\end{aligned}$$

First, we remark that as $\alpha \to 1$ PageRank tends to Seeley's index, so ultimately $\ell > r$. By continuity, we thus just need to show that $\ell = r$ never happens in our range of parameters. If we solve the equations above for c, ℓ and r and impose $\ell = r$, we obtain

$$p = 1 + \frac{\ln\left(-\dfrac{2\alpha^2 - (k^2 - 4k + 6)\alpha + k^2 - 3k + 2}{(k^2 - 3k + 2)\alpha^2 - (2k^2 - 3k)\alpha + k^2 - k}\right)}{\ln \alpha}.$$

Now observe that $2\alpha^2 - (k^2 - 4k + 6)\alpha + k^2 - 3k + 2 \ge 0$ for $k \ge 3$. Thus, a solution for p exists only when the denominator is negative. However, in that region

$$-\frac{2\alpha^2 - (k^2 - 4k + 6)\alpha + k^2 - 3k + 2}{(k^2 - 3k + 2)\alpha^2 - (2k^2 - 3k)\alpha + k^2 - k} \ge 1.$$

This implies that under the same conditions $p \le 1$, which concludes the proof. ▮

5.3 Monotonicity

For the monotonicity axiom, we discuss briefly the nontrivial cases.

5.3.1 Harmonic

If you add an arc $x \to y$ the harmonic centrality of y can only increase, because this addition can only reduce the distances (possibly even turning some of them from infinite to finite), so it will increase their reciprocals (strictly increasing the one from x).

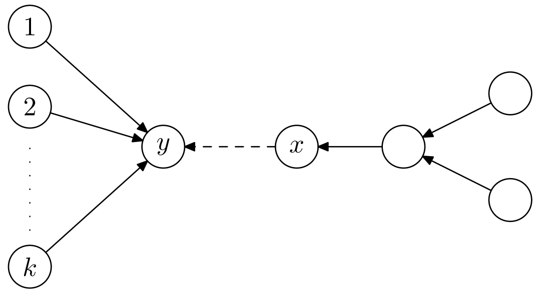


Figure 1: A counterexample showing that Lin's index fails to satisfy the monotonicity axiom.

5.3.2 Closeness

If you consider a one-arc graph $z \to y$ and add an arc $x \to y$, the closeness of y decreases from 1 to $1/2$.

5.3.3 Lin

Consider the graph in Figure 1: the Lin centrality of y is $(k+1)^2/k$. After adding an arc $x \to y$, the centrality becomes $(k+5)^2/(k+9)$, which is smaller than the previous value when $k > 3$.

5.3.4 Betweenness

If you consider a graph made of two isolated nodes x and y, the addition of the arc $x \to y$ leaves the betweenness of x and y unchanged.

5.3.5 Katz

The score of y after adding $x \to y$ can only increase, because the set of paths coming into y now contains new elements[15].

5.3.6 Dominant eigenvector, Seeley's index, HITS

If you consider a clique and two isolated nodes x, y, the rank given by the dominant eigenvector, Seeley's index and HITS to x and y is zero, and it remains unchanged when the arc $x \to y$ is added.

5.3.7 SALSA

Consider the graph in Figure 2: the indegree of y is 1, and its component in the intersection graph of predecessors is trivial, so its SALSA centrality is $(1/1) \cdot (1/6) = 1/6$. After adding an arc $x \to y$, the indegree of y becomes 2, but now its component is $\{ y, z \}$; so the sum of indegrees within the component is $2 + 3 = 5$, hence the centrality of y becomes $(2/5) \cdot (2/6) = 2/15 < 1/6$.

5.3.8 PageRank

The case of PageRank turns out to be definitely nontrivial:

Theorem 4 *PageRank satisfies the monotonicity axiom if* $\alpha \in (0 \mathinner{.\,.} 1)$.

[15]It should be noted, however, that this is true only for the values of the parameter β that still make sense after the addition.

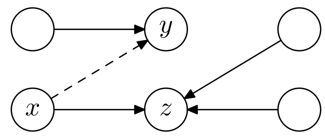


Figure 2: A counterexample showing that SALSA fails to satisfy the monotonicity axiom.

Proof. For this proof, we define PageRank as $\boldsymbol{v}\big(1-\alpha\bar{A}\big)^{-1}$ (i.e., without the normalizing factor $1-\alpha$), so to simplify our calculations. By linearity, the result for the standard definition follows immediately.

Consider two nodes x and y of a graph G such that there is no arc from x to y, and let d be the outdegree of x. Given the normalized matrix $\bar{A}$ of G, and the normalized matrix $\bar{A}'$ of the graph G' obtained by adding to G the arc $x \to y$, we have

$$\bar{A} - \bar{A}' = \boldsymbol{\chi}_x \boldsymbol{\delta},$$

where $\boldsymbol{\chi}_x$ is the characteristic vector of x, and $\boldsymbol{\delta}$ is the difference between the rows corresponding to x in $\bar{A}$ and $\bar{A}'$, which contains $1/d(d+1)$ in the positions corresponding to the successors of x in G, and $-1/(d+1)$ in the position corresponding to y (note that if $d=0$, we have just the latter entry).

We now use the Sherman–Morrison formula to write down the inverse of $1-\alpha\bar{A}'$ as a function of $1-\alpha\bar{A}$. More precisely,

$$\begin{aligned}\big(1-\alpha\bar{A}'\big)^{-1} = \Big(1-\alpha\big(\bar{A}-\boldsymbol{\chi}_x^T\boldsymbol{\delta}\big)\Big)^{-1} &= \big(1-\alpha\bar{A}+\alpha\boldsymbol{\chi}_x^T\boldsymbol{\delta}\big)^{-1}\\ &= \big(1-\alpha\bar{A}\big)^{-1} - \frac{\big(1-\alpha\bar{A}\big)^{-1}\alpha\boldsymbol{\chi}_x^T\boldsymbol{\delta}\big(1-\alpha\bar{A}\big)^{-1}}{1+\alpha\boldsymbol{\delta}\big(1-\alpha\bar{A}\big)^{-1}\boldsymbol{\chi}_x^T}.\end{aligned}$$

We now multiply by the preference vector $\boldsymbol{v}$, obtaining the explicit PageRank correction:

$$\begin{aligned}\boldsymbol{v}\big(1-\alpha\bar{A}'\big)^{-1} &= \boldsymbol{v}\big(1-\alpha\bar{A}\big)^{-1} - \boldsymbol{v}\frac{\big(1-\alpha\bar{A}\big)^{-1}\alpha\boldsymbol{\chi}_x^T\boldsymbol{\delta}\big(1-\alpha\bar{A}\big)^{-1}}{1+\alpha\boldsymbol{\delta}\big(1-\alpha\bar{A}\big)^{-1}\boldsymbol{\chi}_x^T}\\ &= \boldsymbol{r} - \frac{\alpha\boldsymbol{r}\boldsymbol{\chi}_x^T\boldsymbol{\delta}\big(1-\alpha\bar{A}\big)^{-1}}{1+\alpha\boldsymbol{\delta}\big(1-\alpha\bar{A}\big)^{-1}\boldsymbol{\chi}_x^T}\boldsymbol{r} - \frac{\alpha r_x\boldsymbol{\delta}\big(1-\alpha\bar{A}\big)^{-1}}{1+\alpha\boldsymbol{\delta}\big(1-\alpha\bar{A}\big)^{-1}\boldsymbol{\chi}_x^T}.\end{aligned}$$

We now note that $\big(1-\alpha\bar{A}\big)^{-1}\boldsymbol{\chi}_x^T$ is the vector of positive contributions to the PageRank of x, modulo the normalization factor $1-\alpha$. As such, it is made of positive values adding up to at most $1/(1-\alpha)$. When the vector is multiplied by $\boldsymbol{\delta}$, in the worst case ($d=0$) we obtain $1/(1-\alpha)$, so given the conditions on α it is easy to see that the denominator is positive. This implies that we can gather all constants in a single positive constant c and just write

$$\boldsymbol{v}\big(1-\alpha\bar{A}'\big)^{-1} = \big(\boldsymbol{v}-c\boldsymbol{\delta}\big)\big(1-\alpha\bar{A}\big)^{-1}.$$

The above equation rewrites the rank-one correction due to the addition of the arc $x \to y$ as a formal correction of the preference vector. We are interested in the difference

$$\big(\boldsymbol{v}-c\boldsymbol{\delta}\big)\big(1-\alpha\bar{A}\big)^{-1} - \boldsymbol{v}\big(1-\alpha\bar{A}\big)^{-1} = -c\boldsymbol{\delta}\big(1-\alpha\bar{A}\big)^{-1},$$

as we can conclude our proof by just showing that its y-th coordinate is strictly positive.

We now note that being $(1-\alpha\bar{A})$ strictly diagonally dominant, the (nonnegative) inverse $B = (1-\alpha\bar{A})^{-1}$ has the property that the entries b_{ii} on the diagonal are strictly larger than off-diagonal entries b_{ki} on the same column [36, Remark 3.3], and in particular they are nonzero. Thus, if $d = 0$

$$\left[-c\boldsymbol{\delta}\left(1-\alpha\bar{A}\right)^{-1}\right]_y = \frac{c}{d+1}b_{yy} > 0,$$

and if $d \neq 0$

$$\left[-c\boldsymbol{\delta}\left(1-\alpha\bar{A}\right)^{-1}\right]_y = \frac{c}{d+1}b_{yy} - \sum_{x\to z}\frac{c}{d(d+1)}b_{zy} > \frac{c}{d+1}b_{yy} - \sum_{x\to z}\frac{c}{d(d+1)}b_{yy} = 0. \blacksquare$$

6 Roundup

All our proofs are summarized in Table 3, where we distilled our results into simple yes/no answers to the question: does a given centrality measure satisfy the axioms?

It was surprising for us to discover that *only harmonic centrality satisfies all axioms*.[16] All spectral centrality measures are sensitive to density. Row-normalized spectral centrality measures (Seeley's index, PageRank and SALSA) are insensitive to size, whereas the remaining ones are only sensitive to the increase of k (or p in the case of betweenness). All non-attenuated spectral measures are also non-monotone. Both Lin's and closeness centrality fail density tests[17]. Closeness has indeed the worst possible behaviour, failing to satisfy all our axioms. While this result might seem counterintuitive, it is actually a consequence of the known tendence of very far nodes to dominate the score, hiding the contribution of closer nodes, whose presence is more correlated to local density.

All centralities satisfying the density axiom have no watershed: the axiom is satisfied for all $p, k \geq 3$. The watershed for closeness (and Lin's index) is $k \leq p$, meaning that they just miss it, whereas the watershed for betweeness is a quite pathological condition ($k \leq (p^2+p+2)/4$): you need a clique whose size is *quadratic* in the size of the cycle before the node of the clique on the bridge becomes more important than the one on the cycle (compare this with closeness, where $k = p + 1$ is sufficient).

We remark that our results on geometric indices do not change if we replace the directed cycle with a symmetric (i.e., undirected) cycle. It is possible that the same is true also of spectral rankings, but the geometry of the paths of the undirected cycle makes it extremely difficult to carry on the analogous computations in that case.

7 Sanity check via information retrieval

Information retrieval has developed in the last fifty years a large body of research about extracting knowledge from data. In this section we want to leverage the work done in that field to check that our axioms actually describe interesting features centrality measures. We are in this sense following the same line of thought as in [40]: in that paper, the authors tried to establish in a methodologically sound way which of degree, HITS and PageRank works better as a feature in web retrieval. Here we ask the same question, but we include for the first time also geometric indices, which had never been considered before in the literature about information retrieval, most likely because it was not possible to compute them efficiently on large networks.

[16]It is interesting to note that it is actually the only centrality satisfying the size axiom—in fact, you need a cycle of $\approx e^k$ nodes to beat a k-clique.

[17]We note that since $D_{k,p}$ is strongly connected, closeness and Lin's centrality differ just by a multiplicative constant.

Centrality	Size	Density	Monotonicity
Degree	only k	yes	yes
Harmonic	yes	yes	yes
Closeness	no	no	no
Lin	only k	no	no
Betweenness	only p	no	no
Dominant	only k	yes	no
Seeley	no	yes	no
Katz	only k	yes	yes
PageRank	no	yes	yes
HITS	only k	yes	no
SALSA	no	yes	no

Table 3: For each centrality and each axiom, we report whether it is satisfied.

The community working on information retrieval developed a number of standard datasets with associated queries and ground truth about which documents are relevant for every query; those collections are typically used to compare the (de)merits of new retrieval methods; since many of those collections are made of hyperlinked documents, it is possible to use them to assess centrality measures, too.

In this paper we consider the somewhat classical TREC GOV2 collection (aboud 25 million web documents) and the 149 associated queries. For each query (*topic*, in TREC parlance), we have solved the corresponding Boolean conjunction of the terms, obtaining a subset of matching web pages. Each subset induces a graph (whose nodes are the pages satisfying the conjunctive query), which can then be ranked using any centrality measure. Finally, the pages in the graph are listed in rank order as results of the query, and standard relevance measures can be applied to see how much they correspond to the available ground truth about the assessed relevance of pages to queries.

There are a few methodological remarks that are necessary before discussing the results:

- The results we present are for GOV2; there are other publicly available collections with queries and relevant documents that can be used to this purpose.
- As observed in earlier works [40], centrality scores in isolation have a very poor performance when compared with text-based ranking functions, but can improve the results of the latter. We purposely avoid measuring performance in conjunction with text-based ranking because this would introduce further parameters. Moreover, our idea is using information-retrieval techniques to judge centrality measures, not improving retrieval performance *per se* (albeit, of course, a better centrality measure could be used to improve the quality of retrieved documents).
- Because of the poor performance, even for the best documents about half of the queries have null score. Thus, the data we report must be taken with a grain of salt—confidence intervals would be largely overlapping (i.e., our experiments have limited statistical significance).
- Some methods are claimed to work better if *nepotistic links* (that is, links between pages of the same host) are excluded from the graph. We therefore report also results on the procedure applied to GOV2 with all intra-host links removed.
- There are several ways to build a graph associated with a query. Here we choose the simplest possible way—we solve the query in conjunctive form and build the induced subgraph. Variants may include enlarging the resulting graph with successors/predecessors, possibly by sampling [39].

- There are many measures of effectiveness that are used in information retrieval; among those, we focus here on the Precision at 10 (P@10, i.e., fraction of relevant documents retrieved among the first ten) and on the NDCG@10 [26].

The results obtained are presented in Table 4: even if obtained in a completely different way, they confirm the information we have been gathering with our axioms. Harmonic centrality has the best overall scores. When we eliminate nepotistic links, the landscape changes drastically—SALSA and PageRank lead now the results—but the best performances are *worse* than those obtained using the whole structure of the web. Note that, again consistently with the information gathered up to now, closeness performs very badly and betweenness performs essentially like using no ranking at all (i.e., showing the documents in some arbitrary order).

There are two new centrality measures appearing in Table 4 which deserve an explanation. When we first computed these tables, we were very puzzled: HITS is supposed to work very badly on disconnected graphs (it fails monotonicity), whereas it was the second best ranking after harmonic centrality. Also, when you eliminate nepotistic links the graphs become highly disconnected and all rankings tend to correlate with one another simply because most nodes obtain a null score. How is it possible that PageRank and SALSA work so well (albeit less than harmonic centrality on the whole graph) with so little information?

Our suspect was that *these ranking were actually picking up some much more elementary signal than their definition could make you think*. In a highly disconnected graph, the values assigned by such algorithms depends mainly on the indegree and on some additional ranking provided by coreachable (or weakly reachable) nodes.

We thus devised two somewhat paradoxically simple centrality measures, Windegree and Salsina. Windegree is simply the indegree weighted (i.e., multiplied) by the number of coreachable nodes. Salsina is the indegree multiplied by the number of weakly reachable nodes (which is somewhat similar to the way you compute SALSA). Both rankings have been designed to satisfy *all* our axioms. As it is evident from Table 4, such simple rankings outperform in this test most of the very sophisticated rankings proposed in the literature: this shows on one hand that it is possible to extract information from the graph underlying a query in very simple ways that do not involve any spectral technique, and on the other hand that designing centralities around our axioms actually pays off. We consider this fact a further confirmation that the traits of centrality represented by our axioms are important.

8 Conclusions and future work

We have presented a set of axioms that try to capture part of the intended behaviour of centrality measures. We have proved or disproved all our axioms for twelve classical centrality measures and for *harmonic centrality*, a small variant to Bavelas's closeness that we define formally in this paper for the first time. The results are surprising and confirmed by some information-retrieval experiments: harmonic centrality is a very simple measure providing a good notion of centrality. It is almost identical to closeness centrality on undirected, connected networks, but provides a centrality notion for arbitrary directed graphs.

There is of course a large measure of arbitrariness in what we have done: other researchers could come up with other axioms. We believe that this is actually a *feature*—building an ecosystem of interesting axioms is just a healthy way of understanding centrality better and less anecdotally. Promoting the growth of such an ecosystem is one of the goals of this work.

As a final note, the experiments on information retrieval that we have reported are just a start. Testing with different collections (and possibly with different ways of generating the graph associated to a query) may lead to different results. Nonetheless, we believe we have made the important point that *geometric measures are relevant not also to social networks, but also to information retrieval*. In the literature comparing exogenous (i.e., link-based) rankings one can find different instances of

All links

	NDCG@10	P@10
BM25	0.5842	0.5644
Harmonic	0.1438	0.1416
Windegree	0.1373	0.1356
HITS	0.1364	0.1349
Salsina	0.1357	0.1349
Lin	0.1307	0.1289
Katz $3/4\lambda$	0.1228	0.1242
Katz $1/2\lambda$	0.1222	0.1228
Indegree	0.1222	0.1208
Katz $1/4\lambda$	0.1204	0.1181
SALSA	0.1194	0.1221
Closeness	0.1093	0.1114
PageRank $1/2$	0.1091	0.1094
PageRank $1/4$	0.1085	0.1107
Dominant	0.1061	0.1027
PageRank $3/4$	0.1060	0.1094
Betweenness	0.0595	0.0584
—	0.0588	0.0577

Inter-host links only

	NDCG@10	P@10
BM25	0.5842	0.5644
SALSA	0.1384	0.1282
PageRank $1/4$	0.1347	0.1295
Salsina	0.1318	0.1255
PageRank $1/2$	0.1315	0.1268
PageRank $3/4$	0.1313	0.1255
Katz $1/2\lambda$	0.1297	0.1262
Windegree	0.1295	0.1262
Harmonic	0.1293	0.1262
Katz $1/4\lambda$	0.1289	0.1255
Lin	0.1286	0.1248
Indegree	0.1283	0.1248
Katz $3/4\lambda$	0.1278	0.1242
HITS	0.1179	0.1107
Closeness	0.1168	0.1121
Dominant	0.1131	0.1067
Betweenness	0.0588	0.0577
—	0.0588	0.0577

Table 4: Normalized discounted cumulative gain (NDCG) and precision at 10 retrieved documents (P@10) for the GOV2 collection using all links and using only inter-host links. The tables include, for reference, the results obtained using a state-of-the-art text ranking function, BM25, and a final line obtained by applying no ranking function at all (documents are sorted by the document identifier).

spectral rankings and indegree, but up to know that venerable measures based on distances have been neglected. We suggest that it is time to change this attitude.

9 Acknowledgements

We thank David Gleich for useful pointers leading to the proof of the monitonicity axiom for PageRank, and Edith Cohen for useful discussions on the behaviour of centrality indices. Marco Rosa participated to the first phases of development of this paper.

References

[1] Alon Altman and Moshe Tennenholtz. Ranking systems: the PageRank axioms. In *Proceedings of the 6th ACM conference on Electronic commerce*, pages 1–8. ACM, 2005.

[2] Jac M. Anthonisse. The rush in a graph. Technical report, Amsterdam: University of Amsterdam Mathematical Centre, 1971.

[3] Lars Backstrom, Paolo Boldi, Marco Rosa, Johan Ugander, and Sebastiano Vigna. Four degrees of separation. In *ACM Web Science 2012: Conference Proceedings*, pages 45–54. ACM Press, 2012. Best paper award.

[4] A. Bavelas. A mathematical model for group structures. *Human Organization*, 7:16–30, 1948.

[5] A. Bavelas, D. Barrett, and American Management Association. *An experimental approach to organizational communication*. Publications (Massachusetts Institute of Technology. Dept. of Economics and Social Science).: Industrial Relations. American Management Association, 1951.

[6] Alex Bavelas. Communication patterns in task-oriented groups. *Journal of the Acoustical Society of America*, 1950.

[7] Murray A. Beauchamp. An improved index of centrality. *Behavioral Science*, 10(2):161–163, 1965.

[8] Abraham Berman and Robert J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Classics in Applied Mathematics. SIAM, 1994.

[9] Paolo Boldi, Marco Rosa, and Sebastiano Vigna. Robustness of social and web graphs to node removal. *Social Network Analysis and Mining*, 2013.

[10] Paolo Boldi, Massimo Santini, and Sebastiano Vigna. PageRank: Functional dependencies. *ACM Trans. Inf. Sys.*, 27(4):1–23, 2009.

[11] Phillip Bonacich. Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*, 2(1):113–120, 1972.

[12] Stephen P. Borgatti. Centrality and network flow. *Social Networks*, 27(1):55–71, 2005.

[13] Allan Borodin, Gareth O. Roberts, Jeffrey S. Rosenthal, and Panayiotis Tsaparas. Link analysis ranking: algorithms, theory, and experiments. *ACM Transactions on Internet Technology (TOIT)*, 5(1):231–297, 2005.

[14] Alfred Brauer. Limits for the characteristic roots of a matrix. IV: Applications to stochastic matrices. *Duke Math. J.*, 19:75–91, 1952.

[15] Robert L. Burgess. Communication networks and behavioral consequences. *Human Relations*, 22(2):137–159, 1969.

[16] Edith Cohen and Haim Kaplan. Spatially-decaying aggregation over a network. *Journal of Computer and System Sciences*, 73(3):265–288, 2007.

[17] Reuven Cohen and Shlomo Havlin. *Complex Networks: Structure, Robustness and Function*. Cambridge University Press, 2010.

[18] B.S. Cohn and M. Marriott. Networks and centres of integration in Indian civilization. *Journal of Social Research*, 1:1–9, 1958.

[19] Nick Craswell, David Hawking, and Trystan Upstill. Predicting fame and fortune: Pagerank or indegree. In *In Proceedings of the Australasian Document Computing Symposium, ADCS2003*, pages 31–40, 2003.

[20] Ayman Farahat, Thomas Lofaro, Joel C. Miller, Gregory Rae, and Lesley A. Ward. Authority rankings from HITS, PageRank, and SALSA: Existence, uniqueness, and effect of initialization. *SIAM Journal on Scientific Computing*, 27:1181–1201, 2006.

[21] L. Freeman. Centrality in social networks: Conceptual clarification. *Social Networks*, 1(3):215–239, 1979.

[22] Linton C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, 40(1):35–41, 1977.

[23] N.E. Friedkin. Theoretical foundations for centrality measures. *The American Journal of Sociology*, 96(6):1478–1504, 1991.

[24] Roger A. Horn and Stefano Serra-Capizzano. A general setting for the parametric Google matrix. *Internet Math.*, 3(4):385–411, 2006.

[25] Charles H. Hubbell. An input-output approach to clique identification. *Sociometry*, 28(4):377–399, 1965.

[26] Kalervo Järvelin and Jaana Kekäläinen. Cumulated gain-based evaluation of ir techniques. *ACM Trans. Inf. Syst.*, 20(4):422–446, 2002.

[27] Leo Katz. A new status index derived from sociometric analysis. *Psychometrika*, 18(1):39–43, 1953.

[28] Jon M. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5):604–632, September 1999.

[29] Donald E. Knuth. Two notes on notation. *American Mathematical Monthly*, 99(5):403–422, May 1992.

[30] H. J. Leavitt. Some effects of certain communication patterns on group performance. *J Abnorm Psychol*, 46(1):38–50, January 1951.

[31] Ronny Lempel and Shlomo Moran. SALSA: the stochastic approach for link-structure analysis. *ACM Trans. Inf. Syst.*, 19(2):131–160, 2001.

[32] Lun Li, David L. Alderson, John Doyle, and Walter Willinger. Towards a theory of scale-free graphs: Definition, properties, and implications. *Internet Math.*, 2(4), 2005.

[33] Nan Lin. *Foundations of Social Research*. McGraw-Hill, New York, 1976.

[34] Kenneth Mackenzie. Structural centrality in communications networks. *Psychometrika*, 31(1):17–25, 1966.

[35] Massimo Marchiori and Vito Latora. Harmony in the small-world. *Physica A: Statistical Mechanics and its Applications*, 285(3-4):539 – 546, 2000.

[36] J.J. McDonald, M. Neumann, H. Schneider, and M.J. Tsatsomeros. Inverse m-matrix inequalities and generalized ultrametric matrices. *Linear Algebra and its Applications*, 220:321–341, 1995.

[37] Carl D. Meyer. *Matrix analysis and applied linear algebra*. Society for Industrial and Applied Mathematics, pub-SIAM:adr, 2000.

[38] Stanley Milgram. The small world problem. *Psychology Today*, 2(1):60–67, 1967.

[39] Marc Najork, Sreenivas Gollapudi, and Rina Panigrahy. Less is more: sampling the neighborhood graph makes salsa better and faster. In *Proceedings of the Second ACM International Conference on Web Search and Data Mining*, pages 242–251. ACM, 2009.

[40] Marc Najork, Hugo Zaragoza, and Michael J. Taylor. HITS on the web: how does it compare? In Wessel Kraaij, Arjen P. de Vries, Charles L. A. Clarke, Norbert Fuhr, and Noriko Kando, editors, *SIGIR 2007: Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Amsterdam, The Netherlands, July 23-27, 2007*, pages 471–478. ACM, 2007.

[41] Marc A. Najork, Hugo Zaragoza, and Michael J. Taylor. HITS on the web: how does it compare? In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '07, pages 471–478. ACM, 2007.

[42] John F. Padgett and Christopher K. Ansell. Robust Action and the Rise of the Medici, 1400-1434. *The American Journal of Sociology*, 98(6):1259–1319, 1993.

[43] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The PageRank citation ranking: Bringing order to the web. Technical report, Stanford Digital Library Technologies Project, Stanford University, Stanford, CA, USA, 1998.

[44] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The PageRank citation ranking: Bringing order to the web. Technical report, Stanford Digital Library Technologies Project, Stanford University, Stanford, CA, USA, 1998.

[45] Forrest R. Pitts. A graph theoretic approach to historical geography. *The Professional Geographer*, 17(5):15–20, 1965.

[46] John R. Seeley. The net of reciprocal influence: A problem in treating sociometric data. *Canadian Journal of Psychology*, 3:234–240, 1949.

[47] Claude E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27:379–423, 623–656, 1948.

[48] Karen Stephenson and Marvin Zelen. Rethinking centrality: Methods and examples. *Social Networks*, 11(1):1 – 37, 1989.

[49] Trystan Upstill, Nick Craswell, and David Hawking. Query-independent evidence in home page finding. *ACM Trans. Inf. Syst.*, 21(3):286–313, 2003.

[50] Sebastiano Vigna. Spectral ranking, 2009.

[51] Stanley Wasserman and Katherine Faust. *Social network analysis: Methods and applications*. Cambridge Univ Press, 1994.

[52] T.H. Wei. The algebraic foundations of ranking theory, 1952.