Linear Geometric Centralities

Paolo Boldi[0000−0002−8297−6255], Flavio Furia[0009−0003−5265−4161], and Chiara Prezioso

Dipartimento di Informatica, Università degli Studi di Milano, Italy
paolo.boldi@unimi.it

Abstract. Centrality indices are ways to measure the importance of nodes in a graph; this need is so obviously relevant that it was discussed many times in sociology, psychology, mathematics and computer science, giving rise to a whole zoo of definitions of centrality. The ideas underlying such definitions are wildly different, but many centrality measures are based on shortest-path distances: such centralities are referred to as *geometric*. Albeit geometric centralities can use the shortest-path–length information in many different ways, most of the existing geometric centralities can be defined as a linear transformation. In this paper we define formally the class of linear geometric centralities in its full generality, and study its main properties and expressivity.

Keywords: network centrality, linear transformation, shortest paths, Farkas' lemma

1 Introduction and Motivation

Network analysis is a fascinating field, as studying human-created networks offers quantitative insights into properties that influence the domain that the network represents. One interesting issue is to determine which nodes in a graph are the most important: centrality indices are used to assess quantitatively how "central" each node is in a graph, under many possible interpretations of the meaning of importance.

In this work, we study indices that can be referred to as *geometric*, as they solely depend on how many nodes exist at every (shortest-path) distance. Notable examples in this category include in-degree, closeness [3, 4], and harmonic [6] centrality. In this paper, in particular, we propose and study a class of geometric measures that can be computed as linear combination of the number of nodes at every distance: these centralities will be referred to as *linear geometric centralities* (or just linear centralities, for short).

A special class of linear centralities are those assigning larger weights to shorter distances, based on the intuition that a node with many nearby connections is more important than a node with many distant ones. These types of decaying centralities first appeared in [12] (albeit with the equivalent but opposite postulation that smaller centrality values denote a higher importance). In [13] the same class was extended to directed graphs, and several of its properties were explored.

In more recent years, decaying centralities appeared in [18], where they are called additive and studied through the so-called "axiomatic approach" (pioneered in [15]). Finally, in [5] the axiomatic approach was used to investigate the impact of edge additions to the final scores and rankings they yield.

In the present paper, we relax the constraint of decaying weights, and study linear geometric centralities in their full generality: our aim is to analyze their expressivity and to understand whether they capture a broader range of behaviors than decaying centralities.

In particular, the main goal of this work is to answer the two following questions: (a) given a pair of linear geometric centralities, is there always a graph on which they rank nodes differently? (b) given a graph, what is the maximum number of rankings of its nodes that linear centralities can induce? how can we characterize precisely those rankings?

To the best of our knowledge, this is the first attempt at studying this wide class of centrality measures. This work, focusing especially on expressiveness, follows the path opened, for instance, by [16] in that it attempts to find general results for large sets of centralities.

2 Notations and Basic Definitions

For every natural number n, we let $[n] = \{0, 1, \ldots, n-1\}$.

Graphs. A *graph* $G = (V_G, E_G)$ is given [7] by a finite set of nodes V_G and[1] a set of arcs $E_G \subseteq V_G \times V_G$. We shall always assume, without loss of generality, that $V = [n]$ where n is the number of nodes.

We write $x \to y$ to mean that $x, y \in V$ and $(x, y) \in E$. A graph is *undirected* iff $x \to y$ implies $y \to x$. When drawing undirected graphs, pairs of opposite arcs are represented as a single undirected edge.

A *path* of length k from $x \in V$ to $y \in V$ is a sequence $(x_0, x_1, \ldots, x_k)$ of nodes such that $x = x_0$, $y = x_k$ and $x_i \to x_{i+1}$ for all $i \in [k]$. The *(shortest path) distance* from x to y in G, denoted by $d_G(x, y)$, is the length of a shortest path from x to y, or ∞ if no path from x to y exists.

Note that in this paper, unless otherwise specified, we will consider *directed* graphs, where $d(x, y)$ and $d(y, x)$ can be different (it can even be the case that one is finite and the other is not!).

Graph homomorphisms. A *graph homomorphism* $\varphi : G \to H$ is a function $\varphi : V_G \to V_H$ such that $x \to_G y$ iff $\varphi(x) \to_H \varphi(y)$, for all $x, y \in V_G$. A bijective homomorphism $\varphi : G \to H$ is called an *isomorphism*; it is an *automorphism* if $G = H$. A graph with no non-trivial automorphism (i.e., whose only automorphism is the identity) is called *rigid*.

[1] For this and similar notations, the subscript G is dropped whenever it is clear from the context.

Matrices and vectors. In this paper, we shall use column vectors; all vectors and matrices are conveniently indexed starting from 0. So, if $\mathbf{x} \in \mathbf{R}^n$ and $A \in \mathbf{R}^{m \times n}$ then

$$\mathbf{x} = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{bmatrix} \qquad \text{and} \qquad A = \begin{bmatrix} a_{0,0} & a_{0,1} & \dots & a_{0,n-1} \\ a_{1,0} & a_{1,1} & \dots & a_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m-1,0} & a_{m-1,1} & \dots & a_{m-1,n-1} \end{bmatrix}$$

Infinite vectors and matrix product. If $\mathbf{a} \in \mathbf{R}^{\mathbf{N}}$, we view $\mathbf{a}$ as an infinite vector of real numbers. If $A \in \mathbf{R}^{m \times n}$, we write $A \cdot \mathbf{a}$ as a shortcut for

$$A \cdot \mathbf{a}[:n]$$

where $\mathbf{a}[:n]$ stands for the vector of the first n entries of $\mathbf{a}$.

Permutations. For every n, we let S_n be the symmetric group of all permutations (i.e., bijections) $\pi : [n] \to [n]$. The permutation matrix R_π is the binary $n \times n$ matrix such that

$$(R_\pi)_{ij} = 1 \text{ iff } \pi(i) = j.$$

3 Centralities, Permutations and Agreements

A *(graph) centrality* $\mathfrak{f}$ [8] is a function associating with each graph G a map $\mathfrak{f}_G : V_G \to \mathbf{R}$ such that for any two graphs G, H, if $\varphi : G \to H$ is an isomorphism then

$$\mathfrak{f}_G(x) = \mathfrak{f}_H(\varphi(x))$$

for all $x \in V_G$.

While usually centralities are restricted to having non-negative values only, here we are going to relax this requirement and allow for negative centrality values. As usual, the value $\mathfrak{f}_G(x)$ is interpreted as a score of how central (i.e., important) node x is in G: larger values imply greater importance.

We say that $\mathfrak{f}$ has *no ties on* G iff $\mathfrak{f}_G(x) \neq \mathfrak{f}_G(y)$ whenever $x, y \in V_G$ and $x \neq y$. Note that a centrality always has ties on non-rigid graphs: if $\phi : G \to G$ is an automorphism then by definition $\mathfrak{f}(\phi(x)) = \mathfrak{f}(x)$ for all x such that $\phi(x) \neq x$.

Here are some notable examples of centralities[2] (see, for instance, [6] for further examples and taxonomy):

- in-degree: $\mathfrak{C}^{\mathrm{in}}_G(x) = |\{y \mid y \to x\}|$;
- closeness [3]: $\mathfrak{C}^{\mathrm{cl}}_G(x) = 1/\sum_{y \in V, d(y,x) < \infty} d(y,x)$;
- Lin [14]: $\mathfrak{C}^{\mathrm{Lin}}_G(x) = |\{y \in V \mid d(y,x) < \infty\}|^2/\sum_{y \in V, d(y,x) < \infty} d(y,x)$;

[2] Some of these definitions were originally given for undirected graphs only. Here we are providing the adaptations for the general (i.e., directed) case, with the usual proviso that incoming paths are more interesting than outgoing paths.

- harmonic [6]: $\mathfrak{C}^{\mathrm{harm}}_G(x) = \sum_{y\in V} 1/d(y,x)$, with the proviso that $1/\infty = 0$;
- Katz [11]: $\mathfrak{C}^{\mathrm{Katz}}_G(x) = \sum_\pi \beta^{|\pi|}$, where the sum ranges over all paths π ending in x, and β is a parameter smaller than the reciprocal of the spectral radius of G;
- betweenness [1]: $\mathfrak{C}^{\mathrm{bet}}_G(x) = \sum_{y,z,\sigma_{yz}\neq 0} \sigma_{yz}(x)/\sigma_{yz}$, where the sum ranges over all pairs of nodes $y \neq z$, σ_{yz} is the number of shortest paths from y to z, and $\sigma_{yz}(x)$ is the number of such paths passing through x.

In most cases, more than the actual value a centrality assigns to each node, we are interested in how nodes are ranked with respect to a given centrality. This concept is captured formally by the following:

Definition 1 (Respecting a permutation). *Given a graph G of n nodes, we say that $\mathfrak{f}$* respects *the permutation $\pi \in S_n$ on G iff $\mathfrak{f}_G(\pi(i)) \geq \mathfrak{f}_G(\pi(i+1))$ for all $i \in [n-1]$ (that is, π orders the nodes of G in non-increasing order of their centrality values with respect to $\mathfrak{f}$).*

Note that if $\mathfrak{f}$ has no ties on G then there is a unique permutation that $\mathfrak{f}$ respects on G, called the *permutation induced by* $\mathfrak{f}$.

Definition 2 (Agreement between centralities). *Two centralities $\mathfrak{f}$ and $\mathfrak{g}$* agree *on a graph G of n nodes iff they respect the same permutations on G. (In particular, if they both have no ties on G, then they must induce the same permutation). Two centralities that agree on all graphs are called* equivalent.

For example, the following two centralities:

$$\mathfrak{C}^{\mathrm{ccl}}_G(x) = \frac{|V|}{\sum_{y\in V, d(y,x)<\infty} d(y,x)} \qquad \text{“classical” closeness}$$

$$\mathfrak{C}^{\mathrm{np}}_G(x) = -\sum_{y\in V, d(y,x)<\infty} d(y,x) \qquad \text{negative peripherality}$$

are both equivalent to $\mathfrak{C}^{\mathrm{cl}}_G$ (closeness). Classical closeness differs from closeness only by a positive multiplicative constant $|V|$, that has no impact on the ranking of nodes. For negative peripherality, instead of taking the reciprocal of the sum of distances to x, we take the *opposite* of the sum of distances: although the actual values are no doubt different, $\mathfrak{C}^{\mathrm{np}}_G$ induces the same orders on the nodes as the one induced by $\mathfrak{C}^{\mathrm{cl}}_G$.

4 Geometric and Linear Centralities

In the wide world of centralities, many depend only on the number of nodes at a(ny) given distance. As explained in Section 1, in this paper we focus on this class.

Definition 3 (Distance-count function). *Given a graph G with n nodes, and a node $i \in V$, let the* distance-count function of G at i *be the function $c_{G,i} : \mathbf{N} \to \mathbf{N}$ defined by*

$$c_{G,i}(k) = |\{j \in V : d(j,i) = k\}|,$$

that is, the number of nodes at distance k to i.

Note that we do not take infinite distances into account (differently from [13]), thus $c_{G,i}(k) = 0$ for all $k \geq n$, because two nodes cannot have *finite* distance larger than $n-1$. Moreover, note also that $c_{G,i}(0) = 1$ always, because there is exactly one node at distance 0 from i, that is, i itself.

A centrality is geometric if it is essentially dependent only on distance counts:

Definition 4 (Geometric centralities). *A centrality $\mathfrak{f}$ is* strictly geometric *iff for all graphs G and G' and all nodes $i \in V_G$ and $i' \in V_{G'}$ we have*[3]

$$c_{G,i} \equiv c_{G',i'} \textit{ implies } \mathfrak{f}_G(i) = \mathfrak{f}_{G'}(i').$$

A centrality $\mathfrak{f}$ is geometric *iff it is equivalent to one that is strictly geometric.*

For instance, in-degree, closeness, Lin, negative peripherality and harmonic are all strictly geometric: their value on a node x only depends on how many nodes are at each distance from x. Classical closeness ($\mathfrak{C}^{\mathrm{ccl}}_G$) is not strictly geometric: to compute its value on x you need to know the number of nodes in the graph, which is not deducible only from the distance-count function unless the graph is strongly connected. Nonetheless, $\mathfrak{C}^{\mathrm{ccl}}_G$ is equivalent to closeness ($\mathfrak{C}^{\mathrm{cl}}_G$), so $\mathfrak{C}^{\mathrm{ccl}}_G$ is geometric.

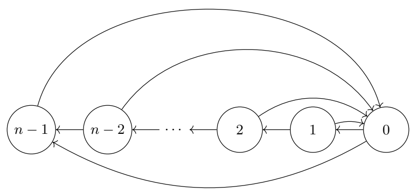
Conversely, Katz centrality is not geometric (because it does not depend on shortest paths, but on *all* paths), neither is betweenness (because it does not depend on the length of shortest paths, but on their number).

Geometric centralities depend only on distance counts, but they may do so in many different ways. We are now going to focus further our attention on the case where the dependence is linear. Let us first define:

Definition 5 (Distance-count matrix). *The* distance-count matrix *of a graph G of n nodes is the matrix $C = C_G \in \mathbf{R}^{n \times n}$ such that the i-th row of C is $\mathbf{c}_{G,i}[: n]$. Said otherwise, $c_{i,k}$ is the number of nodes at distance k to i, for all $i, k \in [n]$.*

In Figure 1, we show an example of distance-count matrix: node 0 (first row) has 1 node at distance 0 (itself), and $n-1$ nodes at distance 1; node 1 (second row) has 1 node at distance 0, 1 node at distance 1 (node 0), and $n-2$ nodes at distance 2; and so on.

[3] We use $\equiv$ to denote that the two functions are equal.



$$C = \begin{bmatrix} 1 & n-1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & 1 & n-2 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & 1 & 1 & n-3 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & 1 & 1 & \dots & 1 & 2 & 0 \\ 1 & 1 & 1 & 1 & 1 & \dots & 1 & 1 & 1 \\ 1 & 2 & n-3 & 0 & 0 & \dots & 0 & 0 & 0 \end{bmatrix}$$

Fig. 1. A graph with n nodes, and its distance-count matrix C.

Definition 6 (Linear centrality). *Given* $\mathbf{a} \in \mathbf{R}^{\mathbf{N}}$ *(the* coefficient vector*), the centrality* $\mathfrak{L}^{\mathbf{a}}_G$ *is defined by*

$$\mathfrak{L}^{\mathbf{a}}_G(i) = \left(C_G \cdot \mathbf{a}\right)_i .$$

A centrality is strictly linear (geometric) *if it is* $\mathfrak{L}^{\mathbf{a}}_G$ *for some* $\mathbf{a} \in \mathbf{R}^{\mathbf{N}}$*. A centrality is* linear (geometric) *if it is equivalent to a strictly linear (geometric) centrality.*

For instance,

- in-degree is strictly linear, with coefficients $(0, 1, 0, 0, \dots)^T$: its value on x is exactly the number of in-neighbors of x (i.e., nodes at distance 1 to x);
- negative peripherality is strictly linear, with coefficients $(0, -1, -2, -3, \dots)^T$;
- harmonic is strictly linear, with coefficients $(0, 1, 1/2, 1/3, 1/4, \dots)^T$;
- closeness and classic closeness are not strictly linear, but they are all linear, since they are all equivalent to negative peripherality.

5 Centralities, Ties and Rigidity

By the very definition of centrality, every centrality (including of course all geometric centralities) has ties on all non-rigid graphs. We can say something more:

Theorem 1. *All centrality measures have ties on graphs that are not rigid. On the other hand, there exists a centrality measure* $\mathfrak{f}$ *such that, for all rigid graphs* G*,* $\mathfrak{f}$ *has no ties on* G*.*

Proof. We exploit the lexicographic canonization process described in [2]: let G be a rigid graph with n nodes, and A_π be the adjacency matrix of G when its nodes are permuted according to $\pi \in S_n$. Suppose that $A_\pi = A_\rho$ for some choice of permutations $\pi, \rho \in S_n$; then for all $x, y \in V_G$:

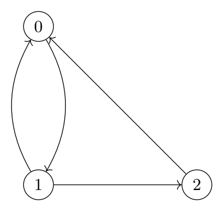
$$A_{\pi(x)\pi(y)} = A_{\rho(x)\rho(y)},$$

that is, $\pi(x) \to_G \pi(y)$ iff $\rho(x) \to_G \rho(y)$. But then $\rho^{-1} \circ \pi$ would be an automorphism of G, hence $\pi = \rho$. So, the matrices A_π are all distinct; hence, there is a unique π such that A_π is minimal in lexicographic order (we can order matrices lexicographically by considering the string obtained concatenating all of their rows). Now you can define $\mathfrak{f}_G$ to be such $\pi : V_G \to [n]$: since π is a permutation, by definition it is injective. $\square$

Geometric centralities, though, are more limited and may give rise to unsolvable ties even on rigid graphs. For instance, the graph in Figure 2 has no non-trivial automorphisms, so according to Theorem 1 there is a centrality that produces no ties on it; nonetheless, all geometric centralities give the same value to nodes 1 and 2, because the corresponding rows in its distance-count matrix C_G are the same. Let us provide a formal definition of this limitation:

Definition 7 (Geometrically rigid). *A graph G is* geometrically rigid *if the rows of its distance-count matrix C_G are all distinct.*

Clearly, every geometrically rigid graph is also rigid, but not the other way round, as the graph in Figure 2 shows. Nonetheless, we can relate the two notions as follows:



$$A_G = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \qquad C_G = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Fig. 2. A rigid graph that is not geometrically rigid, its adjacency matrix A_G and its distance-count matrix C_G.

Theorem 2. *All geometric centrality measures have ties on graphs that are not geometrically rigid. On the other hand, there exists a geometrical centrality measure $\mathfrak{f}$ such that, for all geometrically rigid graphs G, $\mathfrak{f}$ has no ties on G.*

Proof. Define $\mathfrak{f}_G(x)$ to be the rank of x in the lexicographic order of the rows of C_G. Since by definition the rows of C_G are all distinct, the lexicographic order has no ties. On the other hand, if two rows of C_G are the same, the corresponding nodes will have the same value for all geometric centralities (because, by definition, geometric centralities are a function of the distance-count function of the node). $\square$

Whether a similar statement can be proven for *linear* centralities is unclear; what we can prove is a weaker property:

Theorem 3. *For every n, there exists a choice of coefficients $\mathbf{a} \in \mathbf{R}^{\mathbf{N}}$ such that, for all graphs G with at most n nodes, if G is geometrically rigid then $\mathfrak{L}^{\mathbf{a}}_G$ has no ties on G.*

Proof. Define $\mathbf{a} = (n^0, n^1, n^2, \ldots, n^{n-1}, 0, 0, \ldots)^T$. Then $\mathfrak{L}^{\mathbf{a}}_G(i)$ is the base-n encoding of the i-th row of C_G. If G is geometrically rigid, its rows are all distinct, hence $\mathfrak{L}^{\mathbf{a}}_G$ has no ties. □

6 Distinguishable Coefficients

A natural question we may want to ask is the following: how does $\mathfrak{L}^{\mathbf{a}}_G$ change as a function of $\mathbf{a}$? Is it possible that using different coefficients we end up defining equivalent linear centralities? The short answer is no, unless the coefficients differ only by a positive multiplicative constant. Let us provide a specific definition for this case.

Definition 8 (Proportionality). *We say that* $\mathbf{a}, \mathbf{b} \in \mathbf{R}^{\mathbf{N}}$ *are* proportional *iff there is a* $\lambda > 0$ *such that*

$$b_k = \lambda \cdot a_k \text{ for all } k > 0. \tag{1}$$

Observe that in the definition the 0-th coefficient has no role. In the rest of this section, we can safely assume that the 0-th coefficient of all sequences is 0. Some observations are in order:

- if $\mathbf{a}$ and $\mathbf{b}$ are proportional then for all $k > 0$, either a_k and b_k are both zero, or they are both non-zero (and have the same sign);
- no sequence is proportional to $\mathbf{0} = (0, 0, 0, \ldots)$, except $\mathbf{0}$ itself;
- if $\mathbf{a}$ and $\mathbf{b}$ are proportional, and $\mathbf{a} \neq \mathbf{0}$, then the constant $\lambda > 0$ of Definition 8 is unique;
- if $\mathbf{a}$ and $\mathbf{b}$ are not proportional, one of the following statements holds: (a) there is some $k > 0$ such that $a_k b_k = 0$ and $a_k \neq b_k$; (b) for all $i > 0$, $a_i = 0 \iff b_i = 0$ and the minimum index h such that $a_h \neq 0$ satisfies $b_h = \lambda a_h$ for some $\lambda < 0$; (c) for all $i > 0$, $a_i = 0 \iff b_i = 0$ and there exists an index $k > 1$ and $\lambda > 0$ such that $b_i = \lambda \cdot a_i$ for all $0 < i < k$, but $b_k \neq \lambda \cdot a_k$, and moreover $a_h \neq 0$ for some $0 < h < k$.

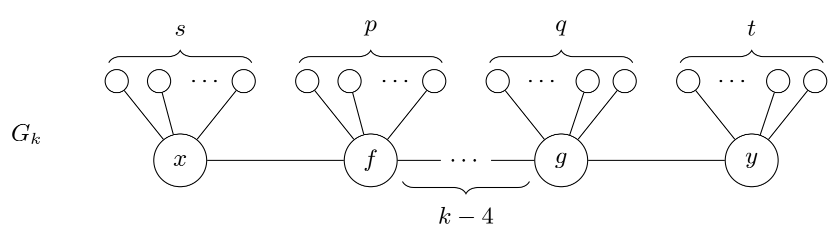


Fig. 3. An undirected graph G_k ($k \geq 4$) used to distinguish two non-proportional sequences. The horizontal dotted line at the bottom represents a path between f and g passing through $k-4$ vertices.

Based on these observations, we are ready to prove the following:

Theorem 4. *For* $\mathbf{a}, \mathbf{b} \in \mathbf{R}^{\mathbf{N}}$ *the following holds:*

1. *if* $\mathbf{a}$ *and* $\mathbf{b}$ *are proportional, then* $\mathfrak{L}^{\mathbf{a}}$ *and* $\mathfrak{L}^{\mathbf{b}}$ *are equivalent centralities;*
2. *if* $\mathbf{a}$ *and* $\mathbf{b}$ *are not proportional, then there exists a graph* G *on which* $\mathfrak{L}^{\mathbf{a}}_G$ *and* $\mathfrak{L}^{\mathbf{b}}_G$ *do not agree; in other words, there is a graph* G *and two nodes* $x, y \in V$ *such that* $\mathfrak{L}^{\mathbf{a}}_G(x) \geq \mathfrak{L}^{\mathbf{a}}_G(y)$ *but* $\mathfrak{L}^{\mathbf{b}}_G(x) < \mathfrak{L}^{\mathbf{b}}_G(y)$.

Proof. Given a graph $G = (V, E)$, if $\mathbf{a}$ and $\mathbf{b}$ are proportional, then $\mathfrak{L}^{\mathbf{b}}_G(x) = \lambda \cdot \mathfrak{L}^{\mathbf{a}}_G(x)$ for all $x \in V$, and it is easy to see that in such a case the two centralities are equivalent.

Let us now assume that $\mathbf{a}$ and $\mathbf{b}$ are not proportional, and look at the last observation before this theorem. If we are in case (a), and k is the first index at which either sequence is zero and the other is not, any bidirectional path of length k proves the statement: just let x and y the two consecutive nodes at one extreme end of the path. The same path can be used for case (b), using a path of length h.

We are left to show the statement for two non-proportional $\mathbf{a}$ and $\mathbf{b}$ both different from $\mathbf{0}$ that satisfy the condition (c) of the last observation. We will first assume that there exists an index $k > 1$ such that

$$(a_1 - a_k)(b_2 - b_{k-1}) \neq (a_2 - a_{k-1})(b_1 - b_k), \tag{2}$$

and the cases where this does not hold will be discussed later; note that necessarily $k \geq 4$.

Consider the smallest k satisfying (2), and the undirected graph G_k in Figure 3, with $n = k + s + p + q + t$ vertices. The linear centralities of x and y with respect to $\mathbf{a}$ and $\mathbf{b}$ in this graph are

$$\begin{aligned}
\mathfrak{L}^{\mathbf{a}}_{G_k}(x) &= e_a + a_1 s + a_2 p + a_{k-1} q + a_k t \\
\mathfrak{L}^{\mathbf{b}}_{G_k}(x) &= e_b + b_1 s + b_2 p + b_{k-1} q + b_k t \\
\mathfrak{L}^{\mathbf{a}}_{G_k}(y) &= e_a + a_1 t + a_2 q + a_{k-1} p + a_k s \\
\mathfrak{L}^{\mathbf{b}}_{G_k}(y) &= e_b + b_1 t + b_2 q + b_{k-1} p + b_k s,
\end{aligned}$$

where $e_a = a_1 + a_2 + \cdots + a_{k-1}$ and $e_b = b_1 + b_2 + \cdots + b_{k-1}$. Then, the linear centralities $\mathfrak{L}^{\mathbf{a}}_G$ and $\mathfrak{L}^{\mathbf{b}}_G$ do not agree on G_k if there is an integer nonnegative solution (s, p, q, t) to the following system of inequalities:

$$\begin{cases}
e_a + a_1 s + a_2 p + a_{k-1} q + a_k t \geq e_a + a_1 t + a_2 q + a_{k-1} p + a_k s \\
e_b + b_1 s + b_2 p + b_{k-1} q + b_k t < e_b + b_1 t + b_2 q + b_{k-1} p + b_k s,
\end{cases}$$

which is equivalent to

$$\begin{cases}
\alpha(s - t) + \alpha'(p - q) \geq 0 \\
\beta(s - t) + \beta'(p - q) < 0,
\end{cases} \tag{3}$$

where we have set $\alpha = a_1 - a_k$, $\beta = b_1 - b_k$, $\alpha' = a_2 - a_{k-1}$ and $\beta' = b_2 - b_{k-1}$. The two inequalities in (3) define two halfplanes (in the coordinate system $X = s - t$

and $Y = p - q$), whose intersection is unbounded and nonempty if $\alpha\beta' \neq \alpha'\beta$, which is exactly the condition (2). As a result, the system (3) has an integral solution (X, Y), whence nonnegative s, t, p, q can be always found.

We are left to consider the case in which there is no k satisfying condition (2), but the condition (c) of the observation holds. Then, we can apply the same line of reasoning as above to the graphs $\widehat{G}_k$ in Figure 4. $\square$

It is a noteworthy fact that all the graphs used to prove the second item of the statement above are undirected, and also connected as long as the condition (2) holds for at least an index k.

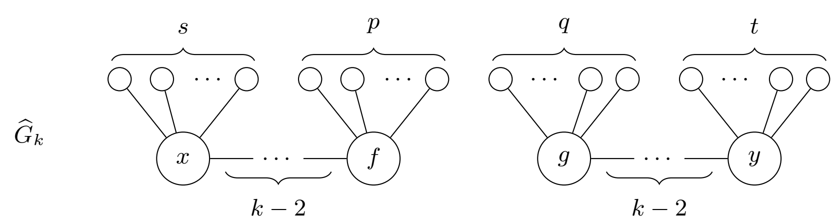


Fig. 4. The graph used to distinguish non-proportional sequences when the condition (2) in the proof of Theorem 4 does not hold. The horizontal dotted lines at the bottom represent paths passing through $k - 2$ vertices.

7 Graphs with Many Representable Permutations

The previous section aimed at discussing when and on which graphs two sequences of coefficients yield different centralities. We now ask a dual question: given a graph G, which permutations of its nodes are induced by some linear centrality? In principle, we may have graphs where all linear centralities agree, regardless of how you choose the coefficients: the fact that there is a graph on which each pair of non-proportional centralities can be distinguished does not rule out the possibility that *on some graph* they may all agree.

On the other hand, it may be possible to have graphs that allow for many different permutations of its nodes, depending on the choice of coefficients. To avoid the complications introduced by nodes having the same centrality value, let us stick to the case with no ties.

Definition 9 (Representable permutation). *We say that a permutation $\pi \in S_n$ is* representable *by a graph G with n nodes iff there is a choice of coefficients $\mathbf{a} \in \mathbf{R}^{\mathbf{N}}$ such that $\mathfrak{L}_G^{\mathbf{a}}$ has no ties and induces the permutation π.*

Of course, we may classify graphs depending on how many permutations they can represent. More precisely, let us define the *representativeness* of a graph G

with n nodes as follows:

$$R(G) = \frac{|\{\pi \in S_n \mid \pi \text{ is representable by } G\}|}{n!} \leq 1.$$

This ratio quantifies how expressive linear centralities are over a rigid graph G. If $R(G) = 1$, linear centralities cannot be less expressible than any other centrality: by choosing coefficients suitably, we can induce any permutation of the nodes, hence achieving what any other centrality can achieve. On the other hand, of course, if $R(G) < 1$ there are permutations that cannot be obtained using a linear centrality, although they may be obtained by some other centrality. In the rest of this section we show that at least in some cases linear centralities are maximally expressive (i.e., they have maximum representativeness).

7.1 Rouché-Capelli Theorem for the Non-Connected Case

Given a graph G of n nodes, and a vector $\mathbf{v} \in \mathbf{R}^n$, determining if there exists an $\mathbf{a} \in \mathbf{R}^{\mathbf{N}}$ such that $\mathfrak{L}^{\mathbf{a}}_G(i) = v_i$ for all $i \in [n]$ can be seen as trying to find a solution $\mathbf{a}$ for the following linear system of equations:

$$C_G \cdot \mathbf{a} = \mathbf{v}. \tag{4}$$

This is a system of n linear equalities with n unknowns, hence by the Rouché-Capelli theorem [17], it has a solution if and only if

$$\mathrm{Rank}(C_G) = \mathrm{Rank}(C_G|\mathbf{v}), \tag{5}$$

where $\mathrm{Rank}(A)$ is the rank of matrix A, and $A|\mathbf{x}$ is the matrix A augmented with the column $\mathbf{x}$.

Now, if $\mathrm{Rank}(C_G) = n$, equation (5) always holds, so if we find a graph G whose distance-count matrix has full rank, then the graph has representativeness 1. Much more than this, in fact: not only can we find linear centralities that rank the nodes in any order we please, but we may even choose the *actual centrality values* for each single node. Such a graph exists:

Theorem 5. *For every $n \geq 3$, the graph G_n in Figure 5 has representativeness $R(G_n) = 1$.*

Proof. The graph in Figure 5 has the following distance-count matrix

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & n-1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & 1 & n-2 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & 1 & 1 & n-3 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & 1 & 1 & \dots & 1 & 2 & 0 \\ 1 & 1 & 1 & 1 & 1 & \dots & 1 & 1 & 1 \end{bmatrix}$$

which is easily seen to have rank n (it is lower-triangular with no zeros on its diagonal). $\square$

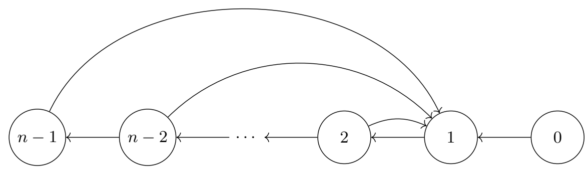


Fig. 5. A graph G_n with representativeness 1.

7.2 Farkas' Lemma for the General Case

The main drawback of the solution in Figure 5 is that the graph G_n is not strongly connected, albeit almost so. So, we still wonder if a *strongly connected* graph can be permuted in all possible ways by linear centralities.

To tackle this problem (or, more in general, to have a tool to decide if a given permutation is representable by a graph), we shall use the following well-known result [9, 10]:

Lemma 1 (Farkas' Lemma). *Given* $A \in \mathbf{R}^{m\times n}$ *and* $\mathbf{b} \in \mathbf{R}^m$*, the following statements are equivalent:*

1. *there exists* $\mathbf{x} \in \mathbf{R}^n$ *such that* $A\mathbf{x} \le \mathbf{b}$*;*
2. *there exists* no $\mathbf{y} \in \mathbf{R}^m$ *such that* $A^T\mathbf{y} = \mathbf{0}$*,* $\mathbf{y} \ge \mathbf{0}$ *and* $\mathbf{b}^T\mathbf{y} < \mathbf{0}$*.*

A consequence of Farkas' Lemma is the following:

Corollary 1. *Given* $A \in \mathbf{R}^{m\times n}$*, the following statements are equivalent:*

1. *there exists* $\mathbf{x} \in \mathbf{R}^n$ *such that* $A\mathbf{x} < \mathbf{0}$*;*
2. *there exists* no $\mathbf{y} \in \mathbf{R}^m$ *such that* $A^T\mathbf{y} = \mathbf{0}$*,* $\mathbf{y} \ge \mathbf{0}$ *and* $\mathbf{y} \ne \mathbf{0}$*.*

We shall use Corollary 1 to determine if a given graph can represent a permutation. Define the matrix $\Delta \in \mathbf{R}^{(n-1)\times n}$ as follows

$$\Delta = \begin{bmatrix} -1 & 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix}$$

For every vector $\mathbf{v} \in \mathbf{R}^n$, the vector is monotonically decreasing (i.e., $v_0 > v_1 > v_2 > \cdots > v_{n-1}$) if and only if $\Delta\mathbf{v} < \mathbf{0}$, because the first inequality is $-v_0 + v_1 < 0$, that is, $v_0 > v_1$, and so on. From this simple observation, we can provide a necessary and sufficient condition for a permutation to be representable.

Theorem 6. *The permutation π is representable by the graph G if and only if there does not exist $\mathbf{y} \in \mathbf{R}^{n-1}$ such that*

$$\left(\Delta R_\pi C_G\right)^T \mathbf{y} = 0 \tag{6}$$

with $\mathbf{y} \geq \mathbf{0}$ and $\mathbf{y} \neq \mathbf{0}$, where R_π is the permutation matrix of π (see Section 2).

Proof. This is obtained applying Corollary 1 with $A = \Delta R_\pi C_G$. □

In other words, to see if π is representable by G we can look at the system (6) of n equations with $n-1$ indeterminates: π is representable if and only if no non-negative non-trivial solution exists for the system. It is useful to rewrite the i-th element of the LHS of the system (6) in a more intelligible form:

$$\begin{aligned}\Big((\Delta R_\pi C)^T \mathbf{y}\Big)_i = \sum_{j<n-1} (\Delta R_\pi C)^T_{i,j}\, y_j = \sum_{j<n-1} (\Delta R_\pi C)_{j,i}\, y_j = \\ = \sum_{j<n-1} \sum_{h_1,h_2} \Delta_{j,h_1} (R_\pi)_{h_1,h_2} c_{h_2,i} y_j\end{aligned}$$

where the variables h_1, h_2 range over $[n]$. Now, recalling the definition of Δ and R_π, we obtain that (6) can be rewritten as:

$$\sum_{j<n-1} \left(c_{\pi(j+1),i} - c_{\pi(j),i}\right) y_j = 0 \qquad \forall i \in [n]. \tag{7}$$

If we let $\rho = \pi^{-1}$, we can re-factor these equations as:

$$\sum_{h:\ 0<\rho(h)} c_{h,i} y_{\rho(h)-1} - \sum_{h:\ \rho(h)<n-1} c_{h,i} y_{\rho(h)} = 0 \qquad \forall i \in [n]. \tag{8}$$

A useful observation is that, if we look at a column i of matrix C having only one single non-zero entry $c_{h,i}$, then the condition (8) becomes

$$y_{\rho(h)} = y_{\rho(h)-1}$$

if $0 < \rho(h) < n-1$; if, instead, $\rho(h) = 0$ ($\rho(h) = n-1$, respectively) then the condition implies $y_0 = 0$ ($y_{n-2} = 0$, resp.).

Note that any linear combination of columns of C yields a valid constraint, too. In other words, we can freely take linear combinations of the columns of C and any $\mathbf{y}$ satisfying the constraints of (8) must also satisfy the new constraints.

7.3 Representativeness of Strongly Connected Graphs

As a first attempt at providing a sequence of strongly connected graphs with large representativeness, we have the following:

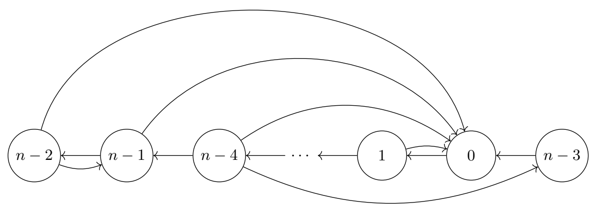
Lemma 2. *For $n > 3$, consider the graph G_n in Figure 1, and let $\pi \in S_n$ and $\rho = \pi^{-1}$. The following two statements are equivalent:*

1. *$\rho(0) < \rho(n-1) < \rho(1)$ or $\rho(1) < \rho(n-1) < \rho(0)$;*
2. *the system* (6) *has no non-negative non-trivial solution.*

Proof. For the sake of space, we omit the proof that is based on an application of (8) and the observations following it. □

It is easy to see (by induction on n) that there are $n!/3$ permutations π satisfying the condition of the above theorem.

Corollary 2. *For every $n > 3$, the graph G_n in Figure 1 is strongly connected and $R(G_n) = 1/3$.*



$$C_{G'_n} = \begin{bmatrix} 1 & n-1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & 1 & n-2 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & 1 & 1 & n-3 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & 1 & 1 & \dots & 1 & 2 & 0 \\ 1 & 1 & 1 & 1 & 1 & \dots & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 & 1 & \dots & 1 & 1 & 0 \end{bmatrix}$$

Fig. 6. The graph G'_n with $R(G'_n) = (n-3)/(n-1)$ and its distance-count matrix $C_{G'_n}$.

The idea of representable permutations, besides being useful to estimate expressiveness of linear centralities on a given graph, has another opposite application that we may call *robustness*. Suppose that you know that an adversary will rank the nodes of your graph using *some unknown* linear centrality (it can be closeness, harmonic, or anything else), and suppose that you want to be sure, say, that a certain node (or set of nodes) is ranked higher than another node (or set of nodes). If you can fiddle a bit with the arcs of the graph, you may want to try to impose a structure which guarantees that only the permutations that you like are representable, thus ruling out the possibility that the event you dislike happens, *whatever linear centrality will be applied by the adversary.* The usage of equation (8) is a very powerful tool to guarantee this kind of properties.

A further quest in the direction of finding graphs with large representativeness led us to consider the graph G'_n of Figure 6. We have experimental results suggesting that $R(G'_n) = (n-3)/(n-1)$, which would imply

$$\lim_{n\to\infty} R(G'_n) = 1,$$

but so far we were unable to prove this result formally. This statement would mean that we have a class of strongly connected graphs that are maximally representative, although only asymptotically so. Whether we can find a family that is ultimately maximally representative remains unknown.

8 Conclusion and Open Problems

This paper presented and studied for the first time the class of linear geometric centralities, providing general properties of expressivity and robustness. Here is a list of problems this paper suggests but that we were unable to solve:

- Does there exist a single linear centrality that has no ties on all geometrically rigid graphs? (In Theorem 3 we solved this problem only for the bounded case)
- Is it possible to find strongly connected graphs to distinguish pairs of non-proportional coefficients, in the cases not covered by the proof of Theorem 4?
- Can we prove formally that $R(G'_n) = (n-3)/(n-1)$ (last paragraph of Section 7.3)?
- Is it possible to find some family G''_n of strongly connected graphs such that $R(G''_n) = 1$, at least for sufficiently large n?

Acknowledgements

This work was supported in part by project SERICS (PE00000014) under the NRRP MUR program funded by the EU - NGEU. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the Italian MUR. Neither the European Union nor the Italian MUR can be held responsible for them.

References

1. Anthonisse, J.M.: The rush in a directed graph. Tech. Rep. BN 9/71, Mathematical Centre, Amsterdam (1971)
2. Babai, L., Luks, E.M.: Canonical labeling of graphs. In: Proceedings of the fifteenth annual ACM symposium on Theory of computing. pp. 171–183 (1983)
3. Bavelas, A.: A mathematical model for group structures. Human Organization **7**, 16–30 (1948)
4. Bavelas, A.: Communication patterns in task-oriented groups. J. Acoust. Soc. Am. **22**(6), 725–730 (1950)
5. Boldi, P., D'Ascenzo, D., Furia, F., Vigna, S.: Score and rank semi-monotonicity for closeness, betweenness, and distance–decay centralities. Social Network Analysis and Mining **14**(1), 183 (Sep 2024)

6. Boldi, P., Vigna, S.: Axioms for centrality. Internet Math. **10**(3-4), 222–262 (2014)
7. Bollobás, B.: Modern graph theory, Graduate Texts in Mathematics, vol. 184. Springer–Verlag (1998)
8. Brandes, U., Erlebach, T.: Network Analysis: Methodological Foundations (Lecture Notes in Computer Science). No. 3418 in Lecture Notes in Computer Science, Springer-Verlag (2005)
9. Farkas, J.: Theorie der einfachen ungleichungen. Journal für die reine und angewandte Mathematik (Crelles Journal) **1902**(124), 1–27 (1902)
10. Gale, D., Kuhn, H.W., Tucker, A.W.: Linear programming and the theory of games. Activity analysis of production and allocation **13**, 317–335 (1951)
11. Katz, L.: A new status index derived from sociometric analysis. Psychometrika **18**(1), 39–43 (1953)
12. Kishi, G.: On centrality functions of a graph. In: Saito, N., Nishizeki, T. (eds.) Graph Theory and Algorithms. pp. 45–52. Springer Berlin Heidelberg, Berlin, Heidelberg (1981)
13. Kishi, G., Takeuchi, M.: A type of centrality functions for a directed graph. Electronics and Communications in Japan (Part I: Communications) **66**(5), 38–47 (1983)
14. Lin, N.: Foundations of Social Research. McGraw-Hill, New York (1976)
15. Sabidussi, G.: The centrality index of a graph. Psychometrika **31**(4), 581–603 (1966)
16. Schoch, D., Brandes, U.: Re-conceptualizing centrality in social networks. European Journal of Applied Mathematics **27**(6), 971–985 (2016)
17. Shafarevich, I.R., Remizov, A.O.: Linear algebra and geometry. Springer Science & Business Media (2012)
18. Skibski, O., Sosnowska, J.: Axioms for distance-based centralities. Proceedings of the AAAI Conference on Artificial Intelligence **32**(1) (Apr 2018)