# **Axioms for Distance-Based Centralities**

### Oskar Skibski, Jadwiga Sosnowska University of Warsaw, Poland

#### Abstract

We study the class of *distance-based centralities* that consists of centrality measures that depend solely on distances to other nodes in the graph. This class encompasses a number of centrality measures, including the classical Degree and Closeness Centralities, as well as their extensions: the Harmonic, Reach and Decay Centralities. We axiomatize the class of distance-based centralities and study what conditions are imposed by the axioms proposed in the literature. Building upon our analysis, we propose the class of *additive distance-based centralities* and pin-point properties which combined with the axiomatic characterization of the whole class uniquely characterize a number of centralities from the literature.

#### Introduction

Identifying the nodes that play the most important role in the network is a fundamental challenge in network analysis (Brandes and Erlebach 2005). This area of research, named *centrality analysis*, is rooted in social network literature, but attracted attention in many fields, including AI. It is, for instance, essential both in determining key infrastructure nodes in the Internet (Page et al. 1999) and central hubs in a transportation networks (Guimera et al. 2005), but also key proteins in protein-protein networks (Jeong et al. 2001).

Arguably, the most well-known centrality measures are the Degree, Closeness and Betweenness Centralities (Freeman 1979). The Degree Centrality assesses the importance of a node simply by looking at the number of its links. The Closeness Centrality, defined as the inverse of the sum of distances to others in the network, promotes nodes which are close to others in the network. In contrast, the Betweenness Centrality measures how many shortest paths in the network traverse a given node. To date, many extensions of these standard centrality measures have been proposed in the literature (Koschützki et al. 2005a).

Unfortunately, the multitude of centrality measures with unclear distinctions between them makes it difficult to determine which one should be used in a specific application. To help understand the differences between various centralities, various classifications have been proposed. Notably, Borgatti (2005) argued that all standard centrality measures can be classified according to two factors: the type of flow and the type of a path considered. Another well-known characterization is due to Borgatti and Everett (2006) who characterized centrality measures in term of cohesiveness. Their approach basically boils down to two categories, both with two options: the type of path and the type of unit.

While these characteristics provide additional insights, they do not help much in answering the question which centrality should be applied to a specific, perhaps non-standard, setting. In particular, criteria chosen are not intuitive and are mostly based on the mathematical formulations and not on the properties stemming from them. For instance, it is hard to decide whether a measure for key nodes in a social networks should be based on paths or walks.

To address the problems with characterization, various authors proposed axiomatic foundations for some centrality measures. This approach was first used by Sabidussi (1966), who proposed several simple axioms that should be satisfied by all centrality measures based on two graph operations – adding an edge, and moving an edge. Since then, a number of authors used axiomatic approach to characterize specific centralities (see the Related Work section for details). Unfortunately, from the three standard centralities only the Degree Centrality has been extensively studied. In other words, there are still no axiomatic characterizations of the other two standard measures – the Closeness and Betweenness Centralities.

Against this background, we study a large class of distance-based centralities which rely only on distances from a node in question to other nodes in the network. This class encompasses both the Degree and Closeness Centralities, as well as their many extensions, such as the Decay, Reach, and Harmonic Centralities. In our analysis we attempt to answer two questions - in which settings distancebased centralities should be used? and, given those settings, which centrality should be selected? To this end, we axiomatize the class of distance-based centralities using the Sabidussi's operations of adding and moving edges, but limited to versions invariant in this class. Building upon this axiomatic characterization, we propose the first axiom system for the Closeness Centrality. Furthermore, we propose and axiomatically characterize the class of additive distancebased centralities and characterize several other centralities from this class.

Copyright © 2018, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

## **Preliminaries**

A graph is a pair, G = (V, E), where V is a set of nodes and E is a set of undirected edges, i.e., subsets of V of size 2.

A path,  $p = (v_1, \ldots, v_k)$ , is a sequence of nodes in which every two consecutive nodes are connected by an edge, i.e.,  $\{v_i, v_{i+1}\} \in E, \forall i \in \{1, \ldots, k-1\}$ . If  $v_1 = v$  and  $v_k = u$ , we say that path is between v and u. The length of a path is the number of edges in it (i.e., the number of nodes in it minus 1). We write  $v \in p$  if v is one of the nodes in p. A maximal subset of nodes such that between every two nodes there is a path is called a *connected component*. The set of connected components of a graph G is denoted K(G). Note that K(G) is a partition of V. A bridge is an edge whose deletion increases the number of connected components. A *cut-vertex* is a node whose deletion increases the number of connected components.

The distance between two nodes  $v, u \in V$  in graph G = (V, E) is denoted by  $d_G(v, u)$ , and is defined as the length of the shortest path between them. If there exists no path between v and u, we assume that  $d_G(v, u) = \infty$ . For a node v, the nodes at distance 1 from v, are called *neighbours* and their set is denoted  $N_G(v)$ . Formally,  $N_G(v) = \{u \in V : \{v, u\} \in E\}$ . For  $k \in \mathbb{N}$ , the set of nodes at distance k from v in graph G is denoted by  $N_G^k(v)$ :

$$N_G^k(v) = \{ u \in V : d_G(v, u) = k \}.$$

In particular,  $N_G^0 = \{v\}$  and  $N_G^1(v) = N_G(v)$ . We will use two shorthand notations:  $N_G^{\leq k}(v)$  for the set of nodes at distance at most k, called k-neighbourhood of node v, and  $N_G^{\leq k}(v)$  for the set of nodes at distance smaller than k. Formally:  $N_G^{\leq k}(v) = \bigcup_{0 \leq i \leq k} N_G^k(v)$ , and  $N_G^{\leq k}(v) =$  $N_G^{\leq k-1}(v)$ . Especially,  $N_G^{\leq \infty}(v)$  denotes the set of all nodes in the same component as v. See Figure 1 for an illustration.

A special subclass of graphs are *paths and stars*. For  $k \in \mathbb{N}$ ,  $P_k$  is a graph which is a path of k nodes, denoted  $u_1, \ldots, u_k$ :

$$P_k = (\{u_1, \dots, u_k\}, \{\{u_i, u_{i+1}\} : i \in \{1, \dots, k-1\}\}).$$

By  $P_k^v$  we denote  $P_k$  such that  $u_1 = v$ . For  $k \in \mathbb{N}$ ,  $S_k^v$  is a graph which is a star, center of which is v:

$$S_k^v = (\{v, v_1, \dots, v_{k-1}\}, \{\{v, v_i\} : i \in \{1, \dots, k-1\}\}).$$

We use shorthand notation to denote graphs obtained by adding/removing edge e or node v to/from graph  $G: G+e = (V, E \cup \{e\}), G-e = (V, E \setminus \{e\}), G+v = (V \cup \{v\}, E)$  and  $G-v = (V \setminus \{v\}, E \setminus \{\{v, u\} : u \in V\})$ .

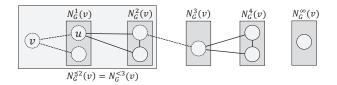


Figure 1: The neighbourhood notation. The dotted nodes are cut vertices. The dotted lines are bridges.

**Centrality measures:** A function that assigns to every node a number reflecting its importance is called a *centrality measure* and defined as  $F : \mathcal{G}^V \to \mathbb{R}^V$ , where  $\mathcal{G}^V$  denotes the set of all possible graphs with nodes V. There is a plethora of centrality measures proposed in the literature. The first and the most well-know centrality indices are the following:

• *Degree Centrality* (*D<sub>v</sub>*) is the number of neighbours of a node:

$$D_v(G) = |N_G(v)|;$$

• Closeness Centrality  $(C_v)$ , defined only for connected graphs, is the inverse of the sum of distances to other nodes (Sabidussi 1966):

$$C_v(G) = \frac{1}{\sum_{u \in V \setminus \{v\}} d_G(v, u)};$$

• Betweenness Centrality  $(B_v)$  is the sum of percentages of shortest paths between any two other nodes that goes through the node under consideration (Freeman 1977). Formally, if we denote by  $\Pi_s(s, t)$  the set of shortest paths between s and t, then:

$$B_v(G) = \sum_{s,t \in V \setminus \{v\}} \frac{|\{p \in \Pi_s(s,t) : v \in p\}|}{|\Pi_s(s,t)|}.$$

All these three centralities are based on the concept of a *distance*. However, in this paper we focus on centralities that assess the importance of node v based only on distances between v and other nodes in the graph. The Betweenness Centrality does not belong to this category, as the Betweenness Centrality of node v depends on the shortest paths, but between nodes other than v.

Other particular centralities studied in this paper are:

 k-Step Reach Centrality (R<sup>k</sup><sub>v</sub>) (or k-Degree Centrality) is the number of distinct nodes within k links of a given node (Borgatti, Everett, and Johnson 2013):

$$R_v^k(G) = |N_G^{\leq k}(v)| - 1.$$

Decay Centrality (Y<sub>v</sub>) is the number of nodes at distance 1 plus the number of nodes at distance 2 multiplied by the decay parameter δ ∈ (0, 1), plus the number of nodes at distance 3 multiplied by δ<sup>2</sup>, and so on (Jackson 2008):<sup>1</sup>

$$Y_v(G) = \sum_{k \ge 1} |N_G^k(v)| \cdot \delta^{k-1}$$

• Harmonic Centrality  $(H_v)$  is an alternative version of the Closeness Centrality – and sometimes called by this name – for graphs that do not have to be connected; it is the sum of inverses of distances to other nodes (assuming  $\frac{1}{\infty} = 0$ ) (Rochat 2009):

$$H_v(G) = \sum_{u \in V \setminus \{v\}} \frac{1}{d_G(v, u)};$$

• Component-Size Centrality  $(S_v)$ , a toy centrality proposed in this paper, is a borderline case of the k-Step Reach Centrality when  $k \to \infty$  and of the Decay Centrality when  $\delta \to 1$ :

$$S_v(G) = |N_G^{<\infty}(v)| - 1.$$

<sup>&</sup>lt;sup>1</sup>Note that sometimes in the definition of the Decay Centrality  $\delta^k$  instead of  $\delta^{k-1}$  is used.

#### **Distance-based centralities**

In this section, we define a class of *distance-based centralities*. This class encompasses all measures such that centrality of a node depends solely on its distances from other nodes in a graph.

**Definition 1.** A centrality measure F is a distance-based centrality if for every two graphs  $G_1 = (V, E_1)$  and  $G_2 = (V, E_2)$  and every  $v \in V$ , condition  $|N_{G_1}^k(v)| = |N_{G_2}^k(v)|$  for every  $k \in \mathbb{N} \cup \{\infty\}$  implies  $F_v(G_1) = F_v(G_2)$ .

We characterize the class of distance-based centralities with three axioms – *Anonymity*, *Add Edge Distance*, and *Move Edge Distance*. Our characterization is based on a seminal work by Sabidussi (1966). The first axiom, *Anonymity*, comes directly from this work. It states that centrality measures should not depend on nodes names. In particular, if there exists an automorphism that transforms one node into another, then they should have the same centrality.

**Anonymity:** For every graph G = (V, E), node  $v \in V$ and bijection  $f: V \to V$ 

$$F_v(V, E) = F_{f(v)}(V, \{\{f(u), f(w)\} : \{u, w\} \in E\}).$$

Two other axioms – Add Edge Distance and Move Edge Distance – states that a centrality measure is invariant under two operations studied by Sabidussi – adding and moving an edge. However, while Sabidussi considered adding and moving an arbitrary edge, to define distance-based centralities, we restrict our attention to adding and moving an edge between equidistant nodes, i.e., nodes at a same distance from the node in question. Specifically, Add Edge Distance states that adding an edge between equidistant nodes does not affect the centrality of v; Move Edge Distance states that moving such an edge to a neighbour of one of incident nodes does not affect the centrality of v either.

Add Edge Distance: For every graph G = (V, E) and every triple of nodes  $v, u, w \in V$  such that  $d_G(v, u) = d_G(v, w)$  $F_v(G) = F_v(G + \{u, w\}).$ 

*Move Edge Distance:* For every graph G = (V, E)and every quadruple of nodes  $v, u, w, t \in V$  such that  $d_G(v, u) = d_G(v, w)$  and  $\{u, w\} \in E, \{w, t\} \in E$ 

$$F_v(G) = F_v(G - \{u, w\} + \{u, t\}).$$

Theorem 1 characterizes distance-based centralities.

**Theorem 1.** A centrality is a distance-based centrality iff it satisfies Anonymity, Add Edge Distance and Move Edge Distance.

*Proof.* It is easy to check that distance-based centralities satisfy Anonymity, Add Edge Distance and Move Edge Distance. Assume F satisfies Anonymity, Add Edge Distance and Move Edge Distance. Let  $G_1 = (V, E_1)$  and  $G_2 = (V, E_2)$  be two graphs and  $v \in V$  be a node such that  $|N_{G_1}^k(v)| = |N_{G_2}^k(v)|$  for every  $k \in \mathbb{N}$ . We will prove that  $F_v(G_1) = F_v(G_2)$ .



Figure 2: The illustration for the proof of Theorem 1. Graph  $G^*$  can be obtained from G by adding edges between the nodes equidistant from v and moving one end of such edges to a neighbour. Thus, based on Add Edge Distance and Move Edge Distance, the centralities of v in G and  $G^*$  are equal.

First, let us assume that there exists a node,  $u \in V$ , such that  $d_{G_1}(v, u) \neq d_{G_2}(v, u)$ . Then, there exists a bijection  $f: V \to V$  such that  $d_{G_1}(v, u) = d_{f(G_2)}(v, f(u))$  and from Anonymity  $F_v(G_1) = F_v(f(G_2))$ . Therefore, in what follows, we assume that  $d_{G_1}(v, u) = d_{G_2}(v, u)$  for every  $u \in V$ .

Now, let us consider graph  $G^* = (V, E^*)$  defined as follows:

$$\begin{split} E^* = \{\{u,w\}: k \geq 0, u \in N_{G_1}^k(v), w \in N_{G_1}^k(v) \cup N_{G_1}^{k+1}(v)\}. \\ \text{Graph } G^* \text{ is the maximal graph with the same set of distances as } G_1 \text{ and } G_2 \text{ and } E_1 \subseteq E^*, E_2 \subseteq E^* \text{ (see Figure 2)}. \\ \text{We will show that } F_v(G_1) = F_v(G^*) = F_v(G_2). \\ \text{Without loss of generality, let us consider graph } G_1. \\ \text{To this end, we will show that } G^* \text{ can be obtained from } G_1 \text{ by adding edges that does not affect the centrality of } v. \end{split}$$

Note that  $E^*$  consists of two types of edges – (1) edges between two nodes from a set  $N_{G_1}^k(v)$  for some k (i.e.,  $\{u, w\}$  such that  $d_{G_1}(v, u) = d_{G_1}(v, w) = k$ ), and (2) edges between a node from a set  $N_{G_1}^k(v)$  and a node from set  $N_{G_1}^{k+1}(v)$  for some k (i.e.,  $\{u, w\}$  such that  $d_{G_1}(v, u) + 1 = d_{G_1}(v, w) = k + 1$ ).

Regarding (1), we know that adding an edge between any two nodes at the same distance from node v does not affect the distances between node v and other nodes. Let G' = (V, E') be a graph obtained from  $G_1$  by adding all such edges from  $G^*$ :

$$E' = E_1 \cup \{\{u, w\} : k \ge 0, u, w \in N_{G_1}^k(v)\}.$$

Thus, for every node  $u \in V$  the distance between v and u is the same in both graphs  $G_1$  and G' and from Add Edge Distance we get  $F_v(G_1) = F_v(G')$ .

Regarding (2), consider  $k \ge 0$  and two nodes  $u \in N_{G_1}^k(v)$ and  $w \in N_{G_1}^{k+1}(v)$  such that  $\{u, w\} \notin E'$ . We will show that adding  $\{u, w\}$  to the graph G' does not change the centrality of v. Since  $d_{G'}(v, w) = d_{G_1}(v, w) = k + 1$ , then there must exist a node  $t \in N_{G_1}^k(v)$ , which is a neighbour of node w in G', i.e.,  $\{t, w\} \in E'$ . Node u is connected with node t in graph G', so from Move Edge Distance, we know that replacing edge  $\{u, t\}$  with  $\{u, w\}$  does not change the centrality of v:

$$F_v(G') = F_v(G' - \{u, t\} + \{u, w\}).$$

After this replacement, from Add Edge Distance, we can add edge  $\{u, t\}$  again, without changing the centrality of node v. Thus, we showed that:

$$F_v(G') = F_v(G' + \{u, w\})$$
 for every  $\{u, w\} \in G^*$ .

Repeating this argument for every edge we obtain graph  $G^*$ and get  $F_v(G_1) = F_v(G^*)$ . This concludes the proof of Theorem 1.

From now on, we will consider only distance-based centralities. We begin with the Closeness Centrality, which is one of the three standard centrality measures (Freeman 1979). It assesses how "central" a given node is in the network by looking at the sum of distances to other nodes in the network. Since the bigger this sum is, the farther the node is, the Closeness Centrality is defined as an inverse of the sum.

To axiomatize this measure, we propose three axioms. First axiom, called *Bridge*, states that out of two endpoints of a bridge the bigger centrality has a node from a bigger part of the graph, i.e., the bigger component in a graph with this bridge removed. Roughly speaking, in such a case, the "center" of a graph is expected to be located in a bigger component; that is why a node from the bigger component is more "central" in the graph.

**Bridge:** For every graph, 
$$G = (V, E)$$
, and edge  $\{v, u\}$   
such that  $K(G - \{v, u\}) = \{C_v, C_u\}, v \in C_v, u \in C_u, |C_v| \le |C_u| \Leftrightarrow F_v(G) \le F_u(G).$ 

The next axiom – *Cut-vertex Average* – relates the centrality of a cut-vertex to the centrality of this node in parts it connects. Specifically, if  $(V_1, E_1)$  and  $(V_2, E_2)$  are two graphs with only one joint node,  $v = V_1 \cap V_2$ , then the centrality of v in the sum of these graphs is half of the harmonic average of centralities in both graphs:  $H(x, y) = 2/(\frac{1}{x} + \frac{1}{y})$ ; i.e., it is an inverse of the sum of inverses of centralities in both graphs.

**Cut-vertex Average:** For every two graphs  $G_1 = (V_1, E_1)$  and  $G_2 = (V_2, E_2)$  such that  $V_1 \cap V_2 = \{v\}$  and  $F_v(G_1), F_v(G_2) > 0$ :

$$F_v(V_1 \cup V_2, E_1 \cup E_2) = \frac{1}{2} \cdot H(F_v(V_1, E_1), F_v(V_2, E_2)).$$

In the next section, we propose an additive version of this axiom, called *Cut-vertex Additivity*, which will be a basis of the class of Additive Distance-Based Centralities.

Finally, we propose a simple axiom saying that in a connected graph centrality should be positive.

**Positivity:** For every connected graph 
$$G = (V, E)$$
,  $|V| > 1$ , and node  $v \in V$ :  $F_v(G) > 0$ .

The following theorem shows that, in the class of distance-based centralities, Bridge, Cut-vertex Average, and Positivity characterize the Closeness Centrality up to an affine transformation of distances.

**Theorem 2.** If a distance-based centrality F satisfies Bridge, Cut-vertex Average and Positivity, then there exists  $\alpha, \beta \in \mathbb{R}, \alpha \ge 0, \alpha + \beta > 0$  such that for every connected graph G = (V, E) and node  $v \in V$ 

$$F_v(G) = \frac{1}{\sum_{u \in V \setminus \{v\}} (\alpha \cdot d(v, u) + \beta)}.$$
 (1)

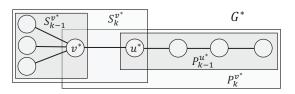


Figure 3: The illustration for the proof of Theorem 2. Bridge implies centralities of  $v^*$  and  $u^*$  are equal. From Cut-vertex Average, we know that the centrality of  $v^*$  is a combination of centralities in  $S_{k-1}^{v^*}$  and  $P_k^{v^*}$ . Analogously, the centrality of  $u^*$  is a combination of centralities in  $P_{k-1}^{u^*}$  and  $S_k^{v^*}$ . For known centralities in a star this yields a recursive formula for centralities in paths.

*Proof.* Assume F is a distance-based centrality that satisfies Bridge, Cut-vertex Average and Positivity. Let us denote by  $a, b \in \mathbb{R}$  two centralities:  $a = F_v(\{v, u\}, \{\{v, u\}\})$  and b = $F_v(\{v, u, w\}, \{\{v, u\}, \{u, w\}\})$ . We will prove that  $F_v(G)$ in every connected graph is uniquely characterized based on a and b. More formally, if F, F' satisfy both these equations for some  $a, b \in \mathbb{R}$ , then F(G) = F'(G) for every connected graph G. To this end, first consider a star,  $S_k^{v^*}$ , with k nodes, the center of which is  $v^*$ . Based on Cut-vertex Average we get a recursive formula for  $F_{v^*}(S_k^{v^*})$ :

$$F_{v^*}(S_k^{v^*}) = \frac{1}{2}H(F_{v^*}(S_{k-1}^{v^*}), F_{v^*}(\{v^*, v_{k-1}\}, \{\{v^*, v_{k-1}\}\}))$$

with the borderline case:  $F_{v^*}(S_2^{v^*}) = a$ . From this formula, we get that:  $F_{v^*}(S_k^{v^*}) = a/(k-1)$ .

Now, consider the centrality of a leaf  $v_1$  in star  $S_k^{v^*}$ . Let w be an additional node, not from the graph. From Move Edge Distance and Cut-vertex Average we get the formula:

$$H(F_{v_1}(S_k^{v^*}), F_{v_1}(\{v_1, w\}, \{\{v_1, w\}\})) =$$

 $H(F_{v_1}(S_{k-1}^{v^*}), F_{v_1}(\{v_1, w, v_{k-1}\}, \{\{v_1, w\}, \{w, v_{k-1}\}\}))$ 

with the borderline cases  $F_{v_1}(S_2^{v^*}) = a$  and  $F_{v_1}(S_3^{v^*}) = b$ . By solving this formula we get  $F_{v_1}(S_k^{v^*}) = ab/(a(k-2) - b(k-3))$ .

Next, let us analyze paths. To this end, consider a graph  $G^*$  obtained by connecting star  $S_{k-1}^{v^*}$  with  $P_{k-1}^{u^*}$  by an edge  $\{v^*, u^*\}$  (see Figure 3). From Bridge we know that  $F_{v^*}(G^*) = F_{u^*}(G^*)$ . Observe that  $v^*$  added to  $P_{k-1}^{u^*}$  forms a path with k nodes:  $P_k^{v^*}$ , and  $u^*$  added to  $S_{k-1}^{v}$  forms a star with k nodes:  $S_k^{v^*}$  with  $u_{k-1} = u^*$ . Using Cut-vertex Average for both nodes we get:

$$H(F_{v^*}(S_{k-1}^{v^*}), F_{v^*}(P_k^{v^*})) = H(F_{u^*}(S_k^{v^*}), F_{u^*}(P_{k-1}^{u^*})).$$

which based on calculated centralities in star gives us recursive formula for  $P_k^{v^*}$ , with the borderline cases  $F_{v^*}(P_2^{v^*}) = a$  and  $F_{v^*}(P_3^{v^*}) = b$ . These conditions imply  $F_{v^*}(P_k^{v^*}) = 2ab/((k-2)(k-1)a - 2(k-3)(k-1)b)$ .

Finally, consider arbitrary connected graph G = (V, E), node  $v^* \in V$  and let k be the distance to the farthest node from  $v^*$ , denoted by  $u^*$ . Consider a graph obtained by adding path  $P_k^{v^*}$  to G. Observe that when we remove all edges of  $u^*$  and add edge  $\{u_k, u^*\}$  no distances of node  $v^*$  will change. Thus, from the definition of distance-based class and Cut-vertex Average we get:

$$H(F_{v^*}(G), F_{v^*}(P_k^{v^*})) = H(F_{v^*}(G - u^*), F_{v^*}(P_{k+1}^{v^*})),$$

with already defined borderline case for  $F_{v^*}(S_j^{v^*})$  for every  $j \in \mathbb{N}$  (all graphs with k = 1). Solving this, we get that for every connected graph G = (V, E), |V| > 1, and  $v \in V$ :

$$F_v(G) = \frac{1}{\sum_{u \in V \setminus \{v\}} \left(\frac{a-2b}{ab} \cdot d(v, u) + \frac{3b-a}{ab}\right)}$$

Consider  $\alpha = \frac{a-2b}{ab}$  and  $\beta = \frac{3b-a}{ab}$ . Since a > 0 from Positivity, we see that  $\alpha + \beta = 1/a > 0$ . Moreover, we proved that  $F_u(S_3^u) = a/2$ , so from the assumption and Bridge, we get that  $b = F_{v_1}(S_3^u) \leq F_u(S_3^u) = a/2$  and  $\alpha = \frac{a-2b}{ab} \geq 0$ . Thus, we proved that if a distance-based centrality F satisfies Bridge, Cut-vertex Average and Positivity, then F satisfies (1) with  $\alpha, \beta$  defined above. This concludes the proof of Theorem 2.

For  $\alpha = 1$  and  $\beta = 0$  (1) simplifies to the Closeness Centrality. A corollary from Theorem 2 is the fact that to characterize the Closeness Centrality we need to additionally specify the centralities of end-points of paths of length 2 and 3.

**Corollary 3.** If a distance-based centrality F satisfies Bridge, Cut-vertex Average and Positivity and  $F_v(P_2^v) = 1$ and  $F_v(P_3^v) = 1/3$ , then for every connected graph it is equal to the Closeness Centrality.

### Additive distance-based centralities

In this section, we propose a new class of centralities – *additive distance-based centralities*. This class consists of distance-based centralities in which the profit from each node at a given distance is fixed.

**Definition 2.** For a sequence of real values:  $a = (a_1, a_2, ..., a_{\infty}), a_i \in \mathbb{R}$  the centrality  $F_v^a$  is defined as

$$F_v^a(G) = \sum_{u \in V \setminus \{v\}} a_{d(v,u)},$$

for every graph G = (V, E). A distance-based centrality is additive if it is equal to  $F_v^a$  for some sequence a.

Table 1 lists centralities that belong to the class of additive distance-based centralities.

If a distance-based centrality is additive, then the centrality of a cut-vertex equals the sum of centralities in each parts it connects. We call this property *Cut-vertex Additivity*.

*Cut-vertex Additivity:* For every two graphs 
$$G = (V_1, E_1), G_2 = (V_2, E_2)$$
 such that  $V_1 \cap V_2 = \{v\}$ :  
 $F_v(V_1 \cup V_2, E_1 \cup E_2) = F_v(V_1, E_1) + F_v(V_2, E_2).$ 

Interestingly, in the following theorem, we show that a distance-based centrality is additive if and only if it satisfies Cut-vertex Additivity.

Centrality	$a_1$	$a_2$	.	$a_k$	$a_{k+1}$	•	$a_{\infty}$
Degree	1	0	•	0	0	•	0
k-Step Reach	1	1	•	1	0	.	0
CompSize	1	1	•	1	1	.	0
Decay	1	$\delta^1$	•	$\delta^{k-1}$	$\delta^k$	.	0
Harmonic	1	$\frac{1}{2}$		$\frac{1}{k}$	$\frac{1}{k+1}$	.	0

Table 1: Additive distance-based centralities.

**Theorem 4.** A distance-based centrality is additive iff it satisfies Cut-vertex Additivity.

*Proof.* Any additive distance-based centrality satisfies Cutvertex Additivity. It remains to prove that if distance-based centrality satisfies Cut-vertex Additivity, then it is additive.

Let F be a distance-based centrality that satisfies Cutvertex Additivity, G = (V, E) be a graph, and  $v \in V$  be an arbitrary node. We will prove that:

$$F_{v}(G) = F_{v}^{a}(G) \text{ for } a_{k} = \begin{cases} F_{v}(P_{k+1}^{v}) - F_{v}(P_{k}^{v}) & \text{if } k < \infty \\ F_{v}(\{v, u\}, \emptyset) & \text{if } k = \infty. \end{cases}$$

$$(2)$$

We will prove this by induction over the number of nodes in the graph. If |V| = 1, then from Cut-vertex Additivity we get  $F_v(G) + F_v(G) = F_v(G) = 0$ , and (2) is satisfied:  $F_v(G) = \sum_{u \in V \setminus \{v\}} a_{d(v,u)} = 0$ . Assume (2) holds if |V| < n. Now, consider graph G = (V, E) with |V| = n. Let u be the farthest node from v: d(v, u) = kand  $k \ge d(v, w)$  for every  $w \in V$ . If  $k = \infty$ , then from Cut-vertex Additivity we immediately get  $F_v(G) =$  $F_v(G-u) + F_v(\{v, u\}, \emptyset)$ . Since  $F_v(\{v, u\}, \emptyset) = a_\infty$ , from the inductive assumption (2) is satisfied. Assume  $k < \infty$  and consider graph  $G^*$  obtained by adding a path graph  $P_v^v$  to G(see Figure 4). From Cut-vertex Additivity we get that:

$$F_v(G) + F_v(P_k^v) = F_v(G^*).$$
 (3)

Let us move u from its position in the graph  $G^*$  to the end of the path  $P_k^v$  (formally, we remove edges of u and add the edge  $\{u_k, u\}$ ). In so doing, we obtain a new graph such that the distance from v to any node in the graph is the same. Therefore, since F is a distance-based centrality, the centrality of node v in this graph is the same. Using Cut-vertex Additivity again we get:

$$F_v(G^*) = F_v(G-u) + F_v(P_{k+1}^v).$$
(4)

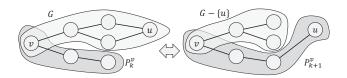


Figure 4: The illustration for the proof of Theorem 4. For a distance-based centrality F the centralities of v in both graphs are equal. For an arbitrary graph G with the farthest node (u) at distance k, Cut-vertex Additivity implies that  $F_v(G) + F_v(P_k^v) = F_v(G-u) + F_v(P_{k+1}^v)$ . Thus, the profit from node at distance k is fixed and centrality is additive.

From (3) and (4) we know that:

$$F_{v}(G) - F_{v}(G - u) = F_{v}(P_{k+1}^{v}) - F_{v}(P_{k}^{v})$$

Thus, removing u from G decreases the centrality of v by  $a_{d(v,u)}$  defined in (2). Using inductive assumption we get (2). This concludes the proof of Theorem 4.

Combining Theorem 4 with Theorem 1 we get that the class of additive distance-based centralities is characterized by Anonymity, Add Edge Distance, Move Edge Distance, and Cut-vertex Additivity.

In theory, any sequence  $(a_1, a_2, \ldots, a_{\infty})$  will characterize an additive distance-based centrality. However, centralities proposed and used in practice (see Table 1) are based on sequences that share several common properties: (1) all values are between 0 and 1; (2)  $a_1$  equals 1; (3)  $a_{\infty}$  equals 0; (4) all sequences are (weakly) decreasing. As we will show, these properties correspond to two known axioms from the literature – *Normalization* and *Monotonicity*.

We begin with Normalization proposed by Skibski et al. (2016). Normalization specifies boundaries for the centrality -0 and (|V| - 1) – where the minimum value should be obtained by isolated nodes and maximal – by the center of a star.

*Normalization:* For every graph, G = (V, E), and every node  $v \in V$ , we have:

(a)  $F_v(G) \in [0, |V| - 1];$ 

- (b)  $F_v(G) = 0$  where v is isolated in G;
- (c)  $F_v(G) = |V| 1$  where G is a star with v in the center.

*Monotonicity* axiom is a part of Sabidussi's axiom (A4) (Sabidussi 1966). Here, we state it in the form proposed by Skibski et al. (2016). Monotonicity requires that adding an edge does not decrease the centrality of any node.

Monotonicity: For every graph, 
$$G = (V, E), v, u, w \in V$$
  
 $F_w(G + \{v, u\}) \ge F_w(G).$ 

The following proposition presents conditions imposed by Normalization and Monotonicity on sequence a.

**Proposition 5.** An additive distance-based centrality  $F^a$  satisfies:

- Normalization iff  $a_1 = 1$ ,  $a_{\infty} = 0$ , and  $a_i \in [0, 1]$  for every  $i \in \{2, 3, ...\}$ ;
- Monotonicity iff  $a_1 \ge a_2 \ge \ldots \ge a_{\infty}$ .

Unlike Normalization, most axioms impose only relation between elements of *a*, but do not indicate specific values. The following proposition states that for every centrality there exists at most one centrality equal to it up to an affine transformation that satisfies Normalization. That is why, in what follows, we will axiomatize centrality measures up to an affine transformation and use Normalization only to pinpoint a specific centrality.

**Definition 3.** We say that an additive distance-based centrality  $F^a$  is equal to  $F^{a'}$  up to an affine transformation if there exists  $\alpha, \beta \in \mathbb{R}$  such that  $a_k = \alpha \cdot a'_k + \beta$  for every  $k \in \{1, 2, ...\} \cup \{\infty\}$ .

**Proposition 6.** For every additive distance-based centrality there exists at most one centrality equal to it up to an affine transformation that satisfies Normalization.

#### The Degree and Component-Size Centralities

In this section, we propose an axiomatization of the Degree Centrality and the Component-Size Centrality. To this end, we consider two axioms from the literature – *Fairness* and *Gain-Loss*. The former one, proposed by Myerson (1977), states that the profit of an edge is the same for both its adjacent nodes. The later one, proposed as an alternative to Monotonicity, states that the sum of centralities does not change if an edge is added to a connected component in the graph (Sosnowska and Skibski 2017).

**Fairness**: For every graph G = (V, E) and every  $v, u \in V$ 

$$F_v(G + \{v, u\}) - F_v(G) = F_u(G + \{v, u\}) - F_u(G).$$

**Gain-Loss:** For every graph, G = (V, E), and every  $v, u \in C \in K(G)$ 

$$\sum_{w \in V} F_w(G + \{v, u\}) = \sum_{w \in V} F_w(G).$$

There exists a number of centralities that satisfy both properties. In fact, Fairness is satisfied by a large class of separable game-theoretic centralities (Skibski, Michalak, and Rahwan 2017). On the other hand, Gain-Loss is satisfied by many game-theoretic centralities based on the Myerson value. In particular, the Attachment Centrality satisfies both axioms (Skibski et al. 2016). Interestingly, in Theorems 7 and 8 we prove that every additive distancebased centrality that satisfies Fairness is equal to the Degree Centrality (up to affine transformation) and every additive distance-based centrality that satisfies Gain-Loss is equal to the Component-Size Centrality (up to affine transformation).

**Theorem 7.** An additive distance-based centrality satisfies Fairness iff it is equal to the Degree Centrality up to an affine transformation. If it also satisfies Normalization, then it is equal to the Degree Centrality.

**Theorem 8.** An additive distance-based centrality satisfies Gain-Loss iff it is equal to the Component-Size Centrality up to an affine transformation. If it also satisfies Normalization, then it is equal to the Component-Size Centrality.

#### The k-Step Reach Centrality

The k-Step Reach Centrality is a middle-point between the Degree Centrality and the Component-Size Centrality. As we will show in this section, it can be characterized by using the same axioms – Fairness and Gain-Loss – but extended/limited to the neighbourhood of a node. Specifically, k-Fairness states that if (k - 1)-neighbourhoods of two nodes do not overlap, then adding an edge does not affect the sum of centralities of these neighbourhoods. On the other hand, k-Gain-Loss states that if adding an edge does no affect (k - 1)-neighbourhood of any node, then the sum of centralities in the graph does not change.

*k*-Fairness: For every graph G = (V, E) and every  $e = \{v, u\} \subseteq V$  such that  $N_G^{\leq k}(v) \cap N_G^{\leq k}(u) = \emptyset$ :

$$\sum_{w \in N_G^{\leq k}(v)} F_w(G+e) - F_w(G) = \sum_{w \in N_G^{\leq k}(u)} F_w(G+e) - F_w(G).$$

**k-Gain-Loss:** For every graph, G = (V, E), and every  $v, u \in V$  such that  $N_G^{\leq k}(w) = N_{G+\{v,u\}}^{\leq k}(w)$  for every  $w \in V$ :

$$\sum_{w \in V} F_w(G + \{v, u\}) = \sum_{w \in V} F_w(G).$$

We define k-Fairness and k-Gain-Loss based on (k-1)neighbourhood, as in the (k-1)-neighbourhood nodes from k different distances are considered: distance  $0, 1, \ldots, (k-1)$ . In result, 1-Fairness is equivalent to Fairness and  $\infty$ -Gain-Loss is equivalent to Gain-Loss. The following theorems characterize the conditions imposed by k-Fairness and k-Gain-Loss.

**Proposition 9.** An additive distance-based centrality  $F^a$  satisfies:

- k-Fairness iff  $a_i = a_j$  for every  $i, j \in \{k+1, k+2, \ldots\} \cup \{\infty\}$ ;
- *k*-Gain-Loss iff  $a_i = a_j$  for every  $i, j \in \{1, 2, ..., k-1\}$ .

*Proof (sketch). (k-Fairness,*  $\Rightarrow$ ): Let  $F^a$  be an additive distance-based centrality that satisfies k-Fairness. We will prove that  $a_i = a_j$  for every  $i, j \in \{k+1, k+2, \ldots\} \cup \{\infty\}$ . Consider graph  $G = (\{u_1, \ldots, u_{2k+m}, w\}, \{\{u_i, u_{i+1}\}\}:$  $i \in \{1, \dots, 2k + m - 1\} \cup \{\{u_k, w\}\}), v = u_1, u = u_{2k+m}$ (see Figure 5 for an illustration). We will also consider graph  $G + \{v, u\}$ . Observe that only the distances to node w are significant. Adding edge  $\{v, u\}$  does not affect distances of nodes from  $N_G^{\leq k}(v)$ . Thus, by analyzing distances from  $N_G^{\leq k}(u)$  to w we get:  $\sum_{i=2k+\lfloor m/2+1 \rfloor}^{2k+m} a_{3k+m+1-i} = 2k+\lfloor m/2+1 \rfloor a_{3k+m+1-i}$  $\sum_{i=2k+\lfloor m/2+1\rfloor}^{2k+m} a_{i-k+1}$ . Using induction over *m* it boils down to  $a_{k+1} = a_{k+m+1}$  for every  $m \in \mathbb{N}_+$ . To prove that  $a_{\infty} = a_{k+1}$  we consider a path  $P_{k+1}^{v}$  with an isolated node. (k-Fairness,  $\Leftarrow$ ): Assume  $a_i = a_j$  for every  $i, j \in$  $\{k+1, k+2, \ldots\} \cup \{\infty\}$ . We will prove that  $F^a$  satisfies k-Fairness. Let G = (V, E) be a graph,  $v, u \in V$  be two nodes such that  $N_G^{\leq k}(v) \cap N_G^{\leq k}(u) = \emptyset$  and denote  $G' = G + \{v, u\}$ . Consider node w from  $N_G^{\leq k}(v)$  and node t such that the distance between w and t is shorter in G' than in G. If  $d_{G'}(w,t) > k$ , then from the assumption  $a_{d_{G'}(w,t)} = a_{d_G(w,t)}$  and the difference of centralities of node w in k-Fairness condition equals 0. Otherwise, if  $d_{G'}(w,t) \leq k$ , then t must be in  $N_G^{\leq k}(u)$  and the expression  $a_{d_{G'}(w,t)} - a_{d_G(w,t)}$  appears both as the profit of (k-1)neighbourhood of v and of u.

(k-Gain-loss,  $\Rightarrow$ ): omitted due to space constraints.

(k-Gain-loss,  $\Leftarrow$ ): Assume  $a_i = a_j$  for every  $1 \le i, j < k$ . We will prove that  $F^a$  satisfies k-Gain-loss. Fix graph  $G = (V, E), G' = G + \{v, u\}$  and assume adding the edge  $\{v, u\}$  to E does not affect (k - 1)-neighbourhood of

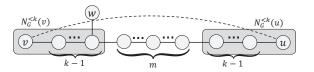


Figure 5: The illustration for the proof of Proposition 9. Adding edge  $\{v, u\}$  does not affect distances from  $N_G^{\leq k}(v)$  to node w, but do affect distances from  $N_G^{\leq k}(u)$  to node w. Considering different m shows  $a_{k+1} = a_{k+m+1}$  for m > 0.

any node. By contradiction, assume there exist two nodes,  $w, t \in V$ , for which  $d_G(w,t) > d_{G'}(w,t)$  and  $a_{d_G(w,t)} \neq a_{d_{G'}(w,t)}$ . Since  $a_i = a_j$  for every  $1 \leq i, j < k$  and from the fact that (k - 1)-neighbourhood of t does not change we know that  $d_G(w,t) > d_{G'}(w,t) \geq k$ . Consider a path  $p = (u_1, u_2, \ldots, u_s)$  between w and t in graph G', i.e.,  $u_1 = w$  and  $u_s = t$ . Since adding the edge  $\{v, u\}$  has decreased the distance between nodes w and t, nodes v, u must appear on the path. Without loss of generality, assume that  $u_m = v$  and  $u_{m+1} = u$ . If  $m \geq k$ , then it can be shown that the (k - 1)-neighbourhood of  $u_{m-k+2}$  has changed. If m < k, then the (k - 1)-neighbourhood of node  $u_k$  has changed – node  $u_1$  is at distance k - 1 in G' and was farther in G. In both cases we get the contradiction.  $\Box$ 

In result, we get the following characterization of the k-Step Reach Centrality.

**Theorem 10.** An additive distance-based centrality satisfies k-Fairness and (k + 1)-Gain-Loss iff it is equal to the k-Step Reach Centrality up to an affine transformation. If it also satisfies Normalization, then it is equal to k-Step Reach Centrality.

### The Decay Centrality

The last centrality measure considered by us is the Decay Centrality. To this end, we propose an axiom named *Leaf Proportionality*. Let v be an isolated node and consider a graph obtained by adding an edge between v and some node  $u \in V \setminus \{v\}$ . Leaf Proportionality states the centrality of v in the new graph minus the profit from a single edge is proportional to the centrality of u in the original graph.

Leaf Proportionality: There exists  $\alpha \in (0,1)$  such that for every graph, G = (V, E), an isolated node  $v \in V$ and node  $u \in V \setminus \{v\}$ 

$$F_v(G + \{v, u\}) - F_v(V, \{\{v, u\}\}) = \alpha \cdot F_u(G).$$

**Theorem 11.** An additive distance-based centrality satisfies Leaf Proportionality iff it is equal to the Decay Centrality multiplied by a scalar. If it also satisfies Normalization, then it is the Decay Centrality.

#### **Related Work**

Since the seminal work by Sabidussi (1966), there have been an extensive literature on axiomatizing centrality measures. See (Koschützki et al. 2005b) and (Boldi and Vigna 2014, Section 4.4) for an overview.

In a related work, Bloch, Jackson, and Tebaldi (2016) propose several *nodal statistics* which are vectors of data that describes the position of a node in the network. One of such statistic is a vector of distances to other nodes:  $(|N_G^1(v)|, |N_G^2(v)|, ...)$ . The authors' main claim is that the classic centrality measures differ solely in terms of vectors of statistics and not in the manner in which they process that information. Our claim can be considered orthogonal, as we studied a large class of centralities based on the same nodal statistic (called *neighbourhood statistic*) and highlighted the differences in a way that they process this information.

As far as we know, in this paper we provided the first axiomatization of the Closeness Centrality and Reach Centrality. There is a couple of papers that axiomatize other distance-based centralities. van den Brink et al. (2008), Dequiedt and Zenou (2014), and Skibski et al. (2016) proposed different axiomatizations of the Degree Centrality. Boldi and Vigna (2014) proposed three axioms and checked that they are satisfied only by the Harmonic Centrality but did not provide characterization results. Finally, in an unfinished manuscript, Garg (2009) characterized the Degree, Decay and Harmonic Centralities (under the name the Closeness Centrality). Two axioms used by Garg are Breadth-First Search, which states the centrality in a graph is equal to the centrality in a breadth-first search tree, and C-Additivity that explicitly states that the profit of a node at a given distance is constant. These axioms combined are equivalent to our definition (but not the axiomatization) of additive distancebased centralities.

#### Conclusions

In this paper, we axiomatized distance-based centralities and its most prominent representative – the Closeness Centrality. Furthermore, we axiomatized the class of additive distancebased centralities and with additional axioms provided axiomatizations of the Degree, Reach, Component-Size, and Decay Centralities. In our future work, we plan to axiomatize the Harmonic Centrality and extend our approach to edge-weighted graphs, which will result in a continuous, not discrete function. We also plan to study the connection between additive distance-based centralities and methods of evaluating candidates scores based on similar score vectors in social choice theory.

## Acknowledgments

Oskar Skibski and Jadwiga Sosnowska were supported by the Foundation for Polish Science within the Homing programme (Project title: "Centrality Measures: from Theory to Applications").

### References

Bloch, F.; Jackson, M. O.; and Tebaldi, P. 2016. Centrality measures in networks. *Available at SSRN 2749124*.

Boldi, P., and Vigna, S. 2014. Axioms for centrality. *Internet Mathematics* 10(3-4):222–262.

Borgatti, S. P., and Everett, M. G. 2006. A graph-theoretic perspective on centrality. *Social Networks* 28(4):466–484.

Borgatti, S. P.; Everett, M. G.; and Johnson, J. C. 2013. *Analyzing social networks*. SAGE Publications Limited.

Borgatti, S. P. 2005. Centrality and network flow. *Social networks* 27(1):55–71.

Brandes, U., and Erlebach, T. 2005. *Network analysis: Methodological foundations*. Springer-Verlag.

Dequiedt, V., and Zenou, Y. 2014. Local and consistent centrality measures in networks. CEPR Discussion Paper No. DP10031.

Freeman, L. C. 1977. A set of measures of centrality based on betweenness. *Sociometry* 35–41.

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

Garg, M. 2009. Axiomatic foundations of centrality in networks. Available at SSRN 1372441.

Guimera, R.; Mossa, S.; Turtschi, A.; and Amaral, L. N. 2005. The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. *Proceedings of the National Academy of Sciences* 102(22):7794–7799.

Jackson, M. O. 2008. *Social and economic networks*, volume 3. Princeton university press.

Jeong, H.; Mason, S.; Barabási, A.-L.; and Oltvai, Z. 2001. Lethality and centrality in protein networks. *Nature* 411(6833):41.

Koschützki, D.; Lehmann, K. A.; Peeters, L.; Richter, S.; Tenfelde-Podehl, D.; and Zlotowski, O. 2005a. Centrality indices. In *Network Analysis*, volume 3418 of *Lecture Notes in Computer Science*. Springer. 16–61.

Koschützki, D.; Lehmann, K. A.; Tenfelde-Podehl, D.; and Zlotowski, O. 2005b. Advanced centrality concepts. In *Network Analysis*, volume 3418 of *Lecture Notes in Computer Science*. Springer. 83–111.

Myerson, R. B. 1977. Graphs and cooperation in games. *Mathematical Methods of Operations Research* 2(3):225–229.

Page, L.; Brin, S.; Motwani, R.; and Winograd, T. 1999. The pagerank citation ranking: bringing order to the web. *Stanford InfoLab*.

Rochat, Y. 2009. Closeness centrality extended to unconnected graphs: The harmonic centrality index. In *ASNA*, number EPFL-CONF-200525.

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

Skibski, O.; Rahwan, T.; Michalak, T. P.; and Yokoo, M. 2016. Attachment centrality: An axiomatic approach to connectivity in networks. In *Proceedings of 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 168–176.

Skibski, O.; Michalak, T. P.; and Rahwan, T. 2017. Axiomatic characterization of game-theoretic network centralities. In *Proceedings* of 31st AAAI Conference on Artificial Intelligence (AAAI), 698– 705.

Sosnowska, J., and Skibski, O. 2017. Attachment centrality for weighted graphs. In *Proceedings of 26th International Joint Conference on Artificial Intelligence (IJCAI)*, 416–422. AAAI Press.

van den Brink, R.; Borm, P.; Hendrickx, R.; and Owen, G. 2008. Characterizations of the  $\beta$ -and the degree network power measure. *Theory and Decision* 64(4):519–536.