Artificial Intelligence 274 (2019) 151–179

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

Attachment centrality: Measure for connectivity in networks

Oskar Skibski [a,*], Talal Rahwan [b], Tomasz P. Michalak [a], Makoto Yokoo [c]

[a] Institute of Informatics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland
[b] Computer Science, New York University, Abu Dhabi, United Arab Emirates
[c] Faculty of Information Science and Electrical Engineering, Kyushu University, 744 Motooka, Nishiku, Fukuoka 819-0395, Japan

ARTICLE INFO

Article history:
Received 30 July 2018
Accepted 24 March 2019
Available online 30 March 2019

Keywords:
Networks
Centrality measures
Axiomatization
Myerson value

ABSTRACT

Centrality indices aim to quantify the importance of nodes or edges in a network. Much interest has been recently raised by the body of work in which a node's connectivity is understood less as its contribution to the *quality or speed of communication* in the network and more as its role in *enabling communication altogether*. Consequently, a node is assessed based on whether or not the network (or part of it) becomes disconnected if this node is removed.

While these new indices deliver promising insights, to date very little is known about their theoretical properties. To address this issue, we propose an axiomatic approach. Specifically, we prove that there exists a unique centrality index satisfying a number of desirable properties. This new index, which we call the *Attachment centrality*, is equivalent to the Myerson value of a certain graph-restricted game. Building upon our theoretical analysis we show that, while computing the Attachment centrality is #P-complete, it has certain computational properties that are more attractive than the Myerson value for an arbitrary game. In particular, it can be computed in chordal graphs in polynomial time.

1. Introduction

Typically, in a network, certain nodes play more important roles than others. For example, popular individuals with frequent social contacts are more likely to spread a disease in the event of an epidemic [18]; popular websites concentrate users' traffic on the Internet [45]; and certain parts of the brain's neural network may be indispensable for breathing or to perform other vital activities [31]. Consequently, the concept of *centrality index* has been extensively studied in the literature [15,44]; such indices aim at identifying the nodes that play a key role in terms of connecting the network and facilitating communication between the nodes therein.

Arguably, the most well-known such indices are: *Degree*, *Closeness*, *Betweenness*, and *Eigenvector* centralities [22,13]. Each of these centralities looks at the connectivity of a node from a different perspectives. In particular, Degree centrality quantifies the power of a node by the number of its connections, i.e., how many other nodes it can directly communicate with. Closeness centrality promotes nodes that are on average close to other nodes in the network and can thus communicate with others more efficiently. Betweenness centrality counts the shortest paths between any two nodes in the network, and ranks nodes according to the number of the shortest paths they belong to; thus nodes with high Betweenness centrality can be thought of as those exert more control over communication in the network. Finally, Eigenvector centrality assumes that

* Corresponding author.
E-mail addresses: oskar.skibski@mimuw.edu.pl (O. Skibski), talal.rahwan@nyu.edu (T. Rahwan), tomasz.michalak@mimuw.edu.pl (T.P. Michalak), yokoo@inf.kyushu-u.ac.jp (M. Yokoo).

https://doi.org/10.1016/j.artint.2019.03.002
0004-3702/

the importance of a node is conditioned on the importance of its neighbors, and the importance of every such neighbor is in turn conditioned on the importance of its neighbors and so on and so forth. Apart from these fundamental centrality indices, many others were proposed in the literature, each looking at the role of a node in the network in a different way [34].

Much interest has been recently raised by the body of work in which a node's connectivity is understood less as its contribution to the *quality or speed of communication* in the network and more as its role in *enabling communication altogether* [27,3,39,4]. Consequently, a node is assessed based on *whether or not the network (or part of it) becomes disconnected if this node is removed*. Such an approach is appealing in numerous settings, especially when the goal(s) of a particular system can be achieved by various groups of agents, the performance of which depends on whether or not the members are connected. The centrality of an agent in such a networked system can be evaluated based on the agent's contribution towards enabling communication in various groups. Such centrality indices are often built upon the well-known coalitional-game model of Myerson [43], where cooperation is restricted to groups, or "coalitions", that are connected in a given network. Within this general framework, different centrality indices are obtained by specifying (i) a characteristic function that assigns a payoff to each connected coalition; and (ii) a method with which the contribution of each node is evaluated, with the Shapley value and the Myerson value being two of main such methods in coalitional game theory.

Unfortunately, the centrality indices based on this approach to connectivity have been so far evaluated only empirically. The lack of theoretical underpinnings not only makes it difficult to choose among those indices, but also hinders the development of efficient algorithms to compute them, especially given the inherent computational difficulty of Myerson's model [55].

To address these limitations, we adopt an *axiomatic approach*, which typically involves two main steps: (i) identifying a set of requirements, or "*axioms*" that are desirable for a centrality index to have; (ii) identifying a centrality index that satisfies those requirements and proving that no other index could possibly satisfy them. In so doing, the axioms provide a strong theoretical foundation for the identified index. Following this approach, we identify five basic axioms (those will be formally defined in subsequent sections):

(1) *Locality*: the centrality of a node depends only on the connected component to which this node belongs;
(2) *Normalization*: the centrality of a node is minimized and equals 0 when the node has no neighbors, and it is maximized and equals $n-1$ when the node is the center of a star;
(3) *Fairness*: the addition of an edge equally affects the centralities of both nodes connected by it;
(4) *Gain-loss*: the addition of an edge in a connected graph does not change the sum of the centralities of all the nodes in this graph;
(5) *Monotonicity*: the addition of an edge does not decrease the centrality of any node.

The first two, namely *Locality* and *Normalization*, seem to be reasonable requirements for any centrality index (indeed, all four standard centrality indices satisfy them). The remaining requirements are concerned with the impact of the addition of an edge. The third axiom, *Fairness*, proposed by Myerson [43], seems reasonable when an edge is considered to be equally owned by both nodes. That is to say, both nodes have an equal say in whether the edge is formed or disbanded (e.g., both sides of a relationship must consent to, and can break, the relation). From this point of view, it seems "fair" that both nodes equally benefit from their edge, since they both equally own it.[1] The final two axioms that we focus on are *Gain-loss* and *Monotonicity*—two somewhat-opposing views of how adding an edge affects other nodes (i.e., those that are not adjacent to the edge).

Our first main result is to prove that there exists a unique index satisfying Normalization, Locality, Fairness and Gain-loss, which we call: *Attachment centrality*; this is the first axiomatized centrality index focusing on the new approach to connectivity in the literature. Interestingly, we prove that by replacing just a single axiom—Gain-loss with Monotonicity—we obtain an axiomatization of Degree centrality.

The Attachment centrality has a simple and elegant interpretation: If we were to remove nodes from the network one by one in a random order, then the Attachment centrality of a node is the expected number of components created immediately after the removal of this node from the network, multiplied by 2 for the normalization purposes.

To get a better insight into the inner workings of the Attachment centrality, we analyze in detail the effects of adding an edge to the network. We capture these effects by introducing what we call *Attachment Delta*—a centrality index parametrized by two nodes, $s, t : s \neq t$. Intuitively, this index represents the role that v plays in connecting nodes s and t which is lost when the direct connection between s and t is added to the network. Importantly, we will show that the Attachment Delta of v equals the change in the Attachment centrality of v that results from the addition of the edge between s and t.

Moreover, our analysis shows that the Attachment centrality is based on the notion of *induced paths*—paths in which (i) there are no repeated nodes, and (ii) any two non-consecutive nodes are not connected by an edge in the network. These paths are important from the perspective of enabling communication, because each node on an induced path is *required* for the existence of that path, or "communication channel". In particular, we show that adding an edge between nodes

[1] Of course one can think of many alternative ways to interpret "fairness". We do not argue that this one is necessarily *the best way*, but rather that it seems to be a *reasonable* way.

s and t affects only the Attachment centrality of nodes that are on induced paths between s and t. Consequently, the Attachment centrality can be considered as an alternative to other centrality measures that are based on shortest paths (e.g., Betweenness centrality), flow (e.g., Flow Betweenness centrality), or walks (e.g., Eigenvector centrality).

Unfortunately, focusing on induced paths entails a high computational complexity. In particular, even determining whether a given node is on an induced path between two other nodes is **NP**-complete. In fact, we show that computing the Attachment centrality is #**P**-complete. This comes from the fact that the Attachment centrality—like many previous attempts to measure a node's ability to enable communication in the network—is based on Myerson's graph-restricted games and on the Myerson value [43].

Nevertheless, our analysis shows that the Attachment centrality has more desirable computational properties than an arbitrary centrality index that is based on the Myerson value. Specifically, we show that our index can be computed in polynomial time in chordal graphs. Furthermore, we propose a general-purpose algorithm that uses cut clique decomposition and perform the standard Myerson-value calculation only for a part of the graph. Finally, using our algorithm, we perform the first analysis focused on identifying communication in the terrorist network responsible for the 2004 attacks on Madrid trains. Our analysis shows that the Attachment centrality gives significantly different results than standard centrality measures.

The remainder of the paper is structured as follows. Section 2 discusses related bodies of literature. Section 3 presents the necessary notation. Section 4 introduces the Attachment centrality along with its axiomatization. Section 5 discusses various properties of the Attachment centrality and how it is related to the notion of induced paths. Finally, Section 6 studies the computational properties of the new centrality index. Conclusions follows.

2. Related work

The literature on the interface of artificial intelligence, graph theory, and social network analysis has grown considerably over the past decades. On one hand, networks are a natural abstraction of many settings of interest to artificial intelligence, such as the domain of multi-agent systems [50] and, more generally, the emerging 4.0 industry [68]. In this context, potential applications include sensor networks and communication networks [49], logistics [30], smart energy and ride-sharing [10], cloud computing [29], just to name a few. Furthermore, social network analysis and graph theory are directly employed to develop new artificial intelligence tools, e.g., to support topic modeling in natural language processing [26], or to construct new classes of deep learning models [41,40]. On the other hand, various artificial intelligence tools can be applied to the analysis of networks [16]. Examples include feature selection methods that are used to support the identification of susceptibility hubs in networks with scale-free structure [36] and deep learning methods that are used to infer the characteristics on different nodes [46].

The remainder of this section is divided into three parts. First, we discuss alternative centrality indices relevant to our approach. Next, we discuss the literature on axiomatizing centrality measures. Finally, we briefly discuss the literature on computing the Myerson value and the computational aspects of graph-restricted games.

2.1. Relevant centrality measures

We begin with the game-theoretic centrality measure studied by Amer and Giménez [3], whereby the centrality of a node is the Shapley value of the following characteristic function: $f(S) = 1$ if the subgraph induced by S is connected and $f(S) = 0$ otherwise. This was later expanded by Lindelauf et al. [37] to an arbitrary $f(S)$, where the subgraph induced by S is connected. As far as the axiomatic underpinnings are concerned, the key difference between both those centrality measure and the Attachment centrality proposed in this article is that the former ones satisfy Fairness and Gain-loss, but not Normalization and Locality.

While the aforementioned centrality indices were based on the Shapley value, other authors also considered the Myerson value. In particular, Gómez et al. [27] proposed a general approach in which the centrality of a node is defined as the difference between the Myerson value and the Shapley value of this node in a given coalitional game. Skibski et al. [55] considered several characteristic functions (e.g., $f(S) = |S|^2$, or $f(S) =$ the number of edges in the subgraph induced by S) combined with the graph-restrictions from Myerson's model. The resulting centrality measures satisfy Locality and Gain-loss. Nonetheless, depending on the characteristic function f that is being used, they may not satisfy Normalization nor Fairness (note that when f is based on the graph, the Myerson value may violate Fairness).

There exists also game-theoretic centrality measures tailored for specific applications. Kötter et al. [35] considered the graph of projections between neuronal structures in a brain, and defined the value of a coalition as the number of strongly connected components of the subgraph induced by that coalition. Moretti et al. [42] studied protein-protein interactions using the general approach proposed by Gómez [27] described above. Here, for a fixed set of *key genes*, the value of any group of genes is equal to the number of key genes that are directly connected (i.e., by an edge) only to members of that group. Finally, Bianzino et al. [7] focused on reducing energy consumption in communication networks, and defined the value of a group as the amount of traffic that this group can effectively transport.

We mention also another group of game-theoretic centralities, which does not make a distinction between connected and disconnected coalitions [38,62]. All these measures do not satisfy Fairness and Normalization. Also, if we consider their normalized version then they do not satisfy Locality.

Finally, note that all of the above centrality measures were only tested empirically, without any axiomatic analysis.

2.2. Axiomatic characterizations of centrality measures

The first attempt to creating axiomatic characterization of centrality measures was proposed by Sabidussi [51]. In his paper, Sabidussi introduced several axioms and defined centrality measures as functions that satisfy these axioms. Sabidussi's axioms are based on two basic operations in a graph—adding an edge and moving an edge (i.e., changing one endpoint of that edge). In fact, the Monotonicity axiom that we consider in Section 4 is one the weakest axioms proposed by Sabidussi. Several other axioms proposed by Sabidussi are violated even by standard centrality measures such as the Closeness and Betweenness centralities.

Since then, many centrality measures have been axiomatically characterized. Because of their (often) complicated nature, a lot of attention in the literature has been devoted to the class of *feedback centralities*, which includes the Eigenvector, Katz, PageRank centralities, among others. For such centrality measures, the importance of a node is directly correlated with the importance of its neighbors. One of the first such axiomatic characterizations was created by van den Brink and Gilles [65] for a centrality measure called the β-measure in the context of directed graphs. Later on, van den Brink et al. [64] proposed an axiomatization of this measure in the context of undirected graphs. Dequiedt and Zenou [17] considered graphs in which some nodes have fixed centralities. Based on this, the authors axiomatized Eigenvector and Katz centralities in undirected graphs. Kitti [32] proposed another axiomatization of the Eigenvector centrality, which relies on algebraic properties of the adjacency matrix. More recently, Wąs and Skibski [66] proposed an axiomatization of the Eigenvector and Katz centralities in directed graphs. One of the axiom used by the authors is a modification of the Locality axiom that we consider in Section 4.

Furthermore, a couple of papers focused on PageRank [45]. Palacios-Huerta and Volij [47] provided an axiomatization of a simplified version of PageRank in the context of journal citation networks, called the *Invariant method*. Altman and Tennenholtz [2] also focused on the simplified version of PageRank, but instead of axiomatizing the resulting numerical values, the authors axiomatized the resulting ranking. Finally, Wąs and Skibski [67] proposed the first axiomatization of PageRank in its general form.

Several papers focused on classical centralities based on distance. Garg [25] characterized the Degree, Decay and Harmonic centralities by considering axioms that refer to the *breadth-first-search* structure of the graph. Boldi and Vigna [12] proposed three axioms and showed that, out of many centrality measures that they considered from the literature, the Closeness centrality for disconnected graphs (called *Harmonic centrality*) was the only one that satisfies all of them. However, they did not propose an axiomatic characterization of this centrality, i.e., they did not show that Harmonic centrality is *the only possible* measures that satisfies all those axioms. Skibski and Sosnowska [57] formally characterized the class of distance-based centralities and its subclass—additive distance based centralities—using axioms similar to the ones proposed by Sabidussi [51]. The authors also proposed axiomatizations of Closeness, Decay, and k-Step Reach centralities. Later on, Sosnowska and Skibski [59] used an axiomatic approach to extend classical distance-based centralities to edge-weighted graphs.

Bloch et al. [11] considered a more general approach. The authors proposed the notion of *nodal statistics*, which describes the position of nodes in the network. The authors' main claim is that the classical centrality measures use different nodal statistics, but process them in the same manner. Thus, for arbitrary nodal statistics, the authors propose a method to axiomatize a centrality measure based on these statistics.

2.3. Computational aspects of the Myerson value and graph-restricted games

The first computational study of the Myerson value was proposed by Bilbao [9], who developed explicit formulas that traverse only connected induced subgraphs in order to compute this value for an arbitrary graph-restricted game. Later on, a different closed-form formulas were developed by Elkind [19] and Skibski et al. [55]. In addition to traversing connected induced subgraphs, the formulas developed by Bilbao [9] and Elkind [19] require finding all the cut vertices in every such subgraph. This, however, is not the case for the formula developed by Skibski et al. [55], which can compute the Myerson value while avoiding the aforementioned extra requirement. Furthermore, we mentioned the work by Algaba et al. [1], who proposed a method for computing the Myerson value based on Harsanyi's dividends in time $O(3^n)$ for n players.

The algorithm for an arbitrary graph proposed in Section 6 builds upon the theoretical analysis from Section 5, which shows that the computation of the Attachment centrality (the Myerson value of a particular game) can be accelerated via the cut clique decomposition of the network. By doing so, the standard Myerson-value calculation only needs to be carried out only for parts of the graph; which can potentially speed up the computation by orders of magnitude, depending on the topology of the network. This allowed us to perform the first analysis focused on identifying communication in the terrorist network responsible for the 2004 attacks on Madrid trains; see Section 6.4. Thus far, this network has never been analyzed with a centrality index that focuses on identifying nodes that enable communication within the network. This is due to the computational challenge entailed by the size of this network, and by the number of connected induced subgraphs in particular.

While most of the literature on the computational aspects of the Myerson value assumes that the game is arbitrary, little is known about the complexity of computing the Myerson value for specific classes of games. One such result was proposed

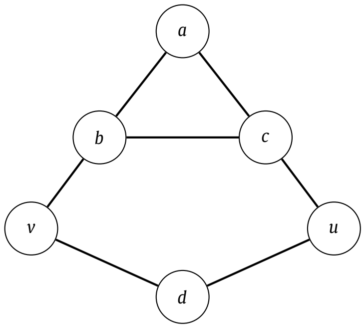

Fig. 1. A sample graph of 6 nodes. In this graph, there are three simple paths between v and u: $p_1 = (v, b, a, c, u)$, $p_2 = (v, b, c, u)$ and $p_3 = (v, d, u)$. Out of these, only p_2 and p_3 are *induced* paths. Furthermore, out of those two induced paths, only p_3 is the *shortest* path between v and u.

by Benati et al. [5], who proved that if a game is a weighted voting game, then computing the Myerson value is #P-complete. In Section 6, we show that even in a unanimity game of two players—one of the simplest games possible—computing the Myerson value is also #P-complete.

Finally, we mention a number of works that considered computing the Myerson value for certain graph types [1,28,20,54], but not for chordal graphs, as is the case in our work (Section 6.2).

3. Preliminaries

Since our work falls at the interface of graph theory and coalitional game theory, we introduce in this section the necessary background and notation pertaining to both of them.

3.1. Graph theory

A graph, or a network, is a pair, $G = (V, E)$, where V is the set of $n = |V|$ nodes, and E is the set of undirected edges. Edge $\{v, u\} \in E$ is said to be *incident* to nodes v and u. The set of all the incident edges of node v is denoted by $\Gamma_G(v)$, i.e., $\Gamma_G(v) = \{\{v, u\} \in E : u \in V\}$. The *degree* of a node v is the number of edges incident to v, i.e., $|\Gamma_G(v)|$. Two nodes, $v, u \in V$, are said to be *neighbors* if they are connected by an edge. The set of all the neighbors of a node v is denoted $N_G(v)$. Formally, $N_G(v) = \{u \in V : \{v, u\} \in E\}$. If a node, v, has no neighbors, i.e., if $N_G(v) = \Gamma_G(v) = \emptyset$, we say that v is *isolated*. If a node has exactly one neighbor, we call it a *leaf*.

A path, $p = (v_1, \dots, v_k)$, is a sequence of nodes in which every two consecutive nodes are connected by an edge, i.e., $\{v_i, v_{i+1}\} \in E,\ \forall i \in \{1, \dots, k-1\}$. The length of a path is the number of edges in it (which equals to the number of nodes in the path minus 1). A path is said to be *simple* if it contains no repeated nodes. We write $v \in p$ if v is one of the nodes in p. The set of all paths between v and u is denoted by $\Pi(v, u)$. The *distance* between any two nodes, $v, u \in V$, is denoted by $dist(v, u)$, and is defined as the length of a shortest path between them. If there exists no path between v and u, we assume that $dist(v, u) = \infty$. There may be multiple shortest paths between certain two nodes. The set of all *shortest* paths between v and u is denoted by $\Pi_s(v, u)$.

Nodes $v, u \in V$ in graph G are said to be *connected* if there exists a path between them, in which case we write $v \sim_G u$. A graph G is said to be connected if every two nodes in it are connected. A node is a *cut vertex* if its removal makes a connected graph disconnected. An edge with the same property, i.e., whose removal makes a graph disconnected, is called a *bridge*.

For any subset of nodes, $S \subseteq V$, the *subgraph induced by* S is denoted by $G[S]$ and is defined as a graph whose nodes are S and whose edges are those in G that connect some nodes in S. Formally:

$$G[S] = (S, E[S]), \text{ where } E[S] = \{\{v, u\} \in E : v, u \in S\}.$$

Any subset of nodes, $S \subseteq V$, is said to be connected if the subgraph induced by S is connected. We denote by $K(G)$ the partition of V into disjoint sets of nodes, called *connected components*, that each induce a *maximal connected subgraph* in G, i.e., a connected subgraph that is not connected to any other node in G. Finally, we denote by $K_v(G)$ the connected component containing v in G.

Next, we introduce the concept of an *induced path*. In particular, a path is said to be *induced* if it is simple and the graph does not contain an edge between any two non-consecutive nodes in that path. Formally, $p = (v_1, \dots, v_k)$ is an induced path if $E[\{v_1, \dots, v_k\}] = \{\{v_i, v_{i+1}\} : i \in \{1, \dots, k-1\}\}$. Fig. 1 illustrates the difference between a simple path, an induced path, and a shortest path. The set of all *induced* paths between v and u is denoted by $\Pi_i(v, u)$. Note that every *shortest* path between v and u is an *induced* path between u and v, i.e., $\Pi_s(v, u) \subseteq \Pi_i(v, u)$.

A *cycle*, $c = (v_1, \dots, v_k)$, is a path such that $v_1 = v_k$. Analogously to paths, the cycle c is *induced* if $E[\{v_1, \dots, v_{k-1}\}] = \{\{v_i, v_{i+1}\} : i \in \{1, \dots, k-1\}\}$. A *tree* is a graph in which there exists *exactly one* path between any two nodes. A *forest* is a graph in which there exists *at most one* path between any two nodes. A *star* is a tree in which there exists a node, v, called the *center* of the star, which is connected by an edge to every other node in the graph, i.e., we have: $E = \{\{v, u\} : u \in V \setminus \{v\}\}$. A graph or a subgraph is said to be a *clique* if every two nodes in it are connected by an edge. A *cut clique* is a clique whose removal makes a connected graph disconnected.

For a graph $G = (V, E)$ and a set of edges E', we use the shorthand notation $G + E'$ to denote the graph obtained by adding E' to E, i.e., the graph $(V, E \cup E')$. Similarly, we use $G - E'$ to denote the graph obtained by removing E' from E, i.e., the graph $(V, E \setminus E')$. When $E' = \{e\}$, we simply write $G + e$ and $G - e$ instead of $G + \{e\}$ and $G - \{e\}$, respectively.

A function that assigns to every node a number reflecting its importance is called a *centrality index* and is defined as $F : \mathcal{G}^V \to \mathbb{R}^V$, where $\mathcal{G}^V$ denotes the set of all possible graphs whose set of nodes is V. Typically, the higher the centrality index is, the more *important* or *central* the node is. The following are three widely-studied centrality indices described by Freeman [22]; we will refer to them as *the standard centrality indices*:

- *Degree centrality*, D_v, of a node, v, is simply the degree of v, i.e.,

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

- *Closeness centrality*, C_v, of a node, v, is the sum of the inverses of distances from v to other nodes (under the assumption that $\frac{1}{\infty} = 0$). Formally, it is defined as follows:

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

  Note that the original definition places the summation sign in the denominator [22], which is only applicable to connected graphs. To avoid this limitation, we consider the above widely-used alternative, which is applicable to any graph [12].
- *Betweenness centrality*, B_v, of a node, v, is the percentage of shortest paths on which v lies, averaged over all pairs of nodes (other than v) that are in the same connected component as v. More formally:

$$B_v(G) = \frac{1}{|K_v(G)| - 2} \sum_{\substack{s,t \in K_v(G) \setminus \{v\} \\ s \neq t}} \frac{|\{p \in \Pi_s(s, t) : v \in p\}|}{|\Pi_s(s, t)|}.$$

  Note that the above formula is normalized to ensure that it yields the same range of values as Degree and Closeness centralities, i.e., $[0, n - 1]$.

3.2. Coalitional game theory

A *game* is a pair (V, f), where V is the set of *players* and $f : 2^V \to \mathbb{R}$ is the *characteristic function*, which assigns a real number to each subset of players (with the only assumption being that $f(\emptyset) = 0$). Any subset of players, $S \subseteq V$, is called a *coalition*, and $f(S)$ is called *the value of coalition* S. We will often refer to a game simply by its characteristic function f. For a set of players, $S \subseteq V$, a *unanimity game*, denoted by u_S, is defined as follows:

$$u_S(T) = \begin{cases} 1 & \text{if } S \subseteq T, \\ 0 & \text{otherwise.} \end{cases}$$

A *value of a game* is a function that assigns a *payoff* to each player $v \in V$, i.e., $\varphi : (2^V \to \mathbb{R}) \to \mathbb{R}^V$. This payoff traditionally represents v's share out of the value of the *grand coalition*, i.e., the coalition of all players. Alternatively, the payoff of v can be interpreted as an assessment of the importance of v in the game. Thus, a *value* of a game plays the same role as a *centrality index* of a network; the former ranks players whereas the latter ranks nodes.

Shapley [52] was the first to propose an axiomatic approach to the problem of payoff division. In particular, Shapley proved that there exists a unique value satisfying some intuitive and desirable properties (also known as *axioms*):

- *Efficiency*: $\sum_{v \in V} \varphi_v(f) = f(V)$, for every game f;
- *Symmetry*: $\varphi_v(f) = \varphi_{\pi(v)}(\pi(f))$, for every game f and every bijection $\pi : V \to V$, where $\pi(f)(S) = f(\{\pi(v) : v \in S\})$;
- *Additivity*: $\varphi(f) + \varphi(f') = \varphi(f + f')$ for every pair of games, f, f', where $(f + f')(S) = f(S) + f'(S)$;
- *Null-player Axiom*: if $f(S \cup \{v\}) = f(S)$ for every $S \subseteq V$, then $\varphi_v(f) = 0$, for every game f.

This value—now widely known as *the Shapley value*—is denoted for player v by $SV_v(f)$ and defined as:

$$SV_v(f) = \frac{1}{|V|} \sum_{S \subseteq V \setminus \{v\}} \frac{1}{\binom{|V|-1}{|S|}} \big(f(S \cup \{v\}) - f(S)\big). \tag{1}$$

Here, the expression $f(S \cup \{v\}) - f(S)$ is known as the *marginal contribution* of player v to coalition S; it is the difference that v makes when joining S. An alternative equivalent formula for the Shapley value is:

$$SV_v(f) = \frac{1}{|V|!} \sum_{\pi \in \Omega(V)} \big(f(S_v^\pi \cup \{v\}) - f(S_v^\pi)\big), \tag{2}$$

where $\Omega(V)$ is the set of all permutations of set V, i.e., all bijections $\pi : V \to \{1, \dots, |V|\}$, and S_v^π is the set of players that precede v in π, i.e., $S_v^\pi = \{u \in V : \pi(u) < \pi(v)\}$.

Myerson [43] considered a model under which the cooperation of players is restricted by a communication graph, G. Specifically in this model, only connected coalitions, i.e., coalitions in which all members can communicate (either directly or indirectly through intermediaries members) can generate value added from cooperation. As for any disconnected coalition, its value equals the sum of the values of its connected components. As such, Myerson's model is defined by a graph, G, and a function, f, that specifies the value of every connected subgraph of G. Over the past decades, this became widely accepted as the canonical model of restricted cooperation. Myerson also proposed a value—now known as *the Myerson value*—which is denoted for player v by $MV_v(f, G)$; this value is simply the Shapley value of the *restricted game* $(V, f/G)$, i.e.:

$$MV_v(f, G) = SV_v(f/G), \tag{3}$$

where f/G is defined as follows:

$$f/G(S) = \sum_{C \in K(G[S])} f(C), \quad \text{for every } S \subseteq V. \tag{4}$$

Myerson [43] proved that for every game, f, there exists a unique value that satisfies two axioms called *Component Efficiency* and *Fairness*. The former axiom can be thought of as a graph version of Shapley's *Efficiency* axiom, i.e.:

- *Component Efficiency (for game f):* $\sum_{v \in C} \varphi_v(G) = f(C)$ for every graph G, every connected component $C \in K(G)$ and every game f.

The later axiom—Fairness—requires that the addition of an edge equally affects the two nodes connected by that edge. We will define this axiom more formally in the following section, when we use it to characterize the importance of different nodes in a given graph.

4. Attachment centrality

Our aim in this section is to define a new centrality index that reflects a node's ability to enable communication in the network, by following an "axiomatic approach". That is, we want to identify a set of requirements, and then prove that there exists exactly one possible centrality index that satisfies all of those requirements, or "*axioms*". This way, the axioms would serve as a theoretical foundation for the centrality index that they uniquely define. Ideally, those axioms should be as intuitive and desirable as possible, to justify the use of the resulting index. In reality, however, any such set of axioms would probably be more suitable and intuitive for some settings, and less so for others. Still, identifying such a set of axioms would serve as a important step towards better understanding how a centrality index, F, can be tailored to reflect a node's ability to enable communication.

To this end, we propose five requirements, namely *Locality*, *Normalization*, *Fairness*, *Monotonicity*, and *Gain-loss*. Next, we explain each of these requirements, starting with Locality:

Locality: *For every graph $G = (V, E)$ and every node $v \in V$, the centrality of v depends solely on $G[K_v(G)]$. That is,*

$$F_v\big(G\big) = F_v\big(G[K_v(G)]\big).$$

As centrality indices are typically defined for connected graphs, the Locality requirement can be interpreted as a natural extension to disconnected graphs, whereby the index is independently applied to every connected component of the disconnected graph. Note that all three standard indices—Degree, Closeness and Betweenness—satisfy this requirement.

Moving on to the *Normalization* requirement, which is inspired by the observation that all three standard centralities—Degree, Closeness, and (normalized) Betweenness—return a *minimum* value of 0, and a *maximum* value of $n-1$. Moreover, they are all *minimized* when the node is isolated, and *maximized* when the node is the center of a star. Our Normalization requirement generalizes this observation as follows.

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

- $F_v(G) \in [0, n-1]$;
- $F_v(G) = 0$ *when v is isolated in G;*
- $F_v(G) = n-1$ *when G is a star, the center of which is v.*

The remaining three requirements are concerned with the impact of adding an edge; *Fairness* focuses on how this addition affects both ends of the edge, whereas *Monotonicity* and *Gain-loss* focus on how this addition affects every node other than the two ends of the edge. Next, we explain each requirement in more detail.

Fairness: *For every $G = (V, E)$ and every $v, u \in V$, adding the edge $\{v, u\}$ affects the centrality of v and u equally:*

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

This notion of Fairness was first proposed by Myerson [43]. Arguably, this seems to be a reasonable requirement when the two ends of the edge are considered to be equally responsible for it. This is perhaps more evident in settings where the formation of an edge requires the consent of both ends, and where the edge can be broken at any time by either end, such as friendship relationships for example. Interestingly, Closeness and Betweenness centralities do not satisfy the Fairness requirement. As mentioned earlier, this requirement is clearly not the only possible interpretation of a "fair" centrality index. However, we choose not to modify the name proposed by Myerson because, at least in some settings, it seems to be reasonably fair.

Moving on to the final two requirements, namely *Gain-loss* and *Monotonicity*; these reflect somewhat-opposing views of how adding an edge affects the indices of the remaining nodes. Next, we formally define these two requirements, before discussing the intuition behind each.

Gain-loss: *For every connected graph, $G = (V, E)$, and every pair of nodes, $u, w \in V$, adding the edge $\{u, w\}$ does not affect the sum of indices, i.e.:*

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

Monotonicity: *For every graph, $G = (V, E)$, adding an edge does not decrease the index of any node in V. That is, for every $v, u, w \in V$:*

$$F_v(G + \{u, w\}) \geq F_v(G).$$

Arguably, from the point of view of enabling communication, Gain-loss makes more sense compared to Monotonicity. To see why this is this case, consider a situation in which the removal of a node, v, breaks a connected graph, G, into two components, G_1 and G_2. Clearly, v plays an important role in terms of enabling communication, since its presence is necessary to connect G_1 with G_2. Now suppose that the edge $\{u, w\}$ was added to G, where u belongs to G_1 and w belongs to G_2. With this addition, it seems reasonable to claim that the role played by u and w grows more important, whereas the role played by v diminishes, since its presence is no longer necessary to connect G_1 with G_2. From this perspective, the Gain-loss requirement seems reasonable, whereas Monotonicity seems rather unintuitive, since it assumes that the role of v in enabling communication remains unchanged, or even grows more important, after the addition of $\{u, w\}$.

Note that the Gain-loss requirement only deals with the addition of an edge to an already-connected component. As such, if an edge is added that connects some of the connected components of a disconnected graph, the Gain-loss requirement places no assumptions or restriction on how this would affect the centrality of nodes across the graph.

Having defined the five requirements, we are now ready to introduce our centrality index, which we call *Attachment centrality*. Later on, we will prove that this centrality index is the unique index satisfying Locality, Normalization, Fairness and Gain-Loss (Theorem 2) and that the Degree centrality is the unique index satisfying Locality, Normalization, Fairness and Monotonicity (Theorem 3).

Definition 1. Attachment centrality is the centrality index defined for a graph, $G = (V, E)$, and a node, $v \in V$, as:

$$A_v(G) = \frac{1}{|V|!} \sum_{\pi \in \Omega(V)} mc^*_v(S^\pi_v), \tag{5}$$

where for any $S \subseteq V \setminus \{v\}$, the term $mc^*_v(S)$ equals the number of components in $G[S]$ that node v connects, multiplied by two, i.e.:

$$mc^*_v(S) = 2(|K(G[S])| - |K(G[S \cup \{v\}])| + 1). \tag{6}$$

The intuition behind the Attachment centrality is as follows. If we were to remove nodes from the graph one by one in a random order, then the Attachment centrality of $v \in V$ would be the expected number of components created from the removal of v, multiplied by 2 for normalization purposes.

Equivalently, the Attachment centrality can be defined as the Myerson value of the game $f^*(S) = 2(|S| - 1)$ restricted by the communication graph G, i.e.:

$$A_v(G) = MV_v(f^*, G). \tag{7}$$

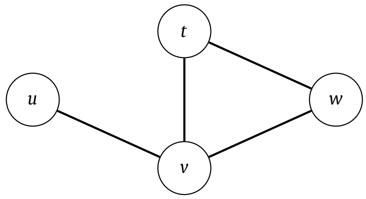

Fig. 2. The graph considered in Example 1.

To see why this is the case, first recall that f/G denotes the game f restricted by the graph G, defined in Equation (4). Thus:

$$f^*/G(S) = \sum_{C \in K(G[S])} f^*(C) = \sum_{C \in K(G[S])} 2(|C|-1) = 2(|S| - |K(G[S])|). \tag{8}$$

This equation implies that $f^*/G(S) \in \{0, 2, \dots, 2(|S|-1)\}$. It also implies that $f^*/G(S \cup \{v\}) - f^*/G(S) = mc_v^*(S)$. Based on this, as well as Equation (2), we find that the Attachment centrality is equivalent to the Shapley value of the game f^*/G, i.e.:

$$SV_v(f^*/G) = \frac{1}{|V|!} \sum_{\pi \in \Omega(V)} \left(f^*/G(S_v^\pi \cup \{v\}) - f^*/G(S_v^\pi) \right) = \frac{1}{|V|!} \sum_{\pi \in \Omega(V)} mc_v^*(S_v^\pi) \ = \ A_v(G) \tag{9}$$

Based on this equation, as well as Equation (3), we can derive the following alternative formula for the Attachment centrality:

$$A_v(G) = \frac{1}{|V|} \sum_{S \subseteq V \setminus \{v\}} \frac{mc_v^*(S)}{\binom{|V|-1}{|S|}}. \tag{10}$$

Example 1. Consider the sample graph, G, that is illustrated in Fig. 2. We know from Equation (8) that the coalitional game f^*/G assigns to every group of nodes, S, a value representing how well S is connected. For example, if there are no edges between the members of S, then its value is minimized and equals 0, e.g., for $S = \{u, w\}$ we have $f^*/G(S) = 0$. On the other hand, if S is connected[2] then its value is maximized and equals $2(|S|-1)$, e.g., for $S = \{v, u, w\}$ we have $f^*/G(S) = 2(3-1) = 4$. More generally, the values of all the groups in the game f^*/G are as follows:

S	$f^*/G(S)$	S	$f^*/G(S)$	S	$f^*/G(S)$	S	$f^*/G(S)$
$\emptyset$	0	$\{v\}$	0	$\{u, w\}$	0	$\{v, u, w\}$	4
$\{u\}$	0	$\{v, u\}$	2	$\{u, t\}$	0	$\{v, u, t\}$	4
$\{t\}$	0	$\{v, t\}$	2	$\{w, t\}$	2	$\{v, w, t\}$	4
$\{w\}$	0	$\{v, w\}$	2	$\{u, w, t\}$	2	$\{v, u, w, t\}$	6

We also know from Equation (9) that the Attachment centrality is defined as the Shapley value of the game f^*/G, i.e., it reflects the importance of each node in this game. Now, consider node v. The coalitions that do not contain v are in columns 1 and 3 of the table above, while the coalitions containing v are in columns 2 and 4. Using Equation (10), we have:

$$AV_v(G) = \frac{1}{4}\left(\frac{0}{1} + \frac{2}{3} + \frac{2}{3} + \frac{2}{3} + \frac{4}{3} + \frac{4}{3} + \frac{2}{3} + \frac{4}{1}\right) = \frac{7}{3}.$$

Performing the same calculations for other nodes we get:

$$AV_u(G) = 1, \qquad AV_t(G) = 4/3, \qquad AV_w(G) = 4/3.$$

In accordance with intuition, the only cut vertex, v, is the most important node in the graph according to the Attachment centrality. Nodes w and t are slightly more important than node u, as they are connected to each other. This concludes Example 1.

In order to prove that the Attachment centrality is the unique centrality index that satisfies Locality, Normalization, Fairness and Gain-loss, we will start by showing that it satisfies those axioms.

[2] Recall that a subset of nodes, $S \subseteq V$, is said to be connected if $G[S]$—the subgraph induced by S—is connected.

Lemma 1. *The Attachment centrality satisfies Locality, Normalization, Fairness, and Gain-loss.*

Proof. We consider all axioms one by one. In particular, Myerson's result, as well as Equation (7), imply that the Attachment centrality satisfies Fairness, and that for every graph G and every connected component $C \in K(G)$ we have:

$$\sum_{v \in C} A_v(G) = 2(|C| - 1). \tag{11}$$

This implies that adding an edge between two nodes in a connected component $C \in K(G)$ does not affect the sum of the Attachment centrality of every node in that component. This, in turn, implies that the Attachment centrality satisfies the Gain-loss requirement.

Moving on to Normalization; we need to show that the Attachment centrality satisfies the three conditions outlined in the definition of Normalization. To this end, for every $v \in V$:

- If v is isolated in G, then v forms a connected component by itself, and (11) implies that $A_v(G) = 0$.
- If G is a star the center of which is v, then we need to show that $A_v(G) = n - 1$. To this end, for every $S \subseteq V \setminus \{v\}$, the induced subgraph $G[S]$ consists of $|S|$ connected components, whereas the subgraph $G[S \cup \{v\}]$ consists of a single connected component. Thus, $mc^*_v(S) = 2|S|$. This fact, together with Equation (10), imply that:

$$A_v(G) = \frac{1}{|V|} \sum_{S \subseteq V \setminus \{v\}} \frac{2|S|}{\binom{|V|-1}{|S|}} = \frac{2}{n} \sum_{s=0}^{n-1} s \sum_{\substack{S \subseteq V \setminus \{v\} \\ |S|=s}} \frac{1}{\binom{n-1}{s}} = \frac{2}{n} \sum_{s=0}^{n-1} s = n - 1.$$

- Finally, we need to prove that $A_v(G) \in [0, n-1]$. Note that $0 \le mc^*_v(S) \le 2|S|$ simply because the number of components in $G[S]$ that are connected by v is between 0 and $|S|$. Furthermore, if $mc^*_v(S) = 2|S|$ for every $S \subseteq V \setminus \{v\}$, then $A_v(G) = n - 1$, as we just proved. In contrast, if $mc^*_v(S) = 0$ for every $S \subseteq V \setminus \{v\}$, then $A_v(G) = 0$. This implies that $A_v(G) \in [0, n - 1]$.

The above three points imply that the Attachment centrality satisfies the Normalization requirement.

Next, we prove that the Attachment centrality satisfies Locality. In other words, given a graph, $G = (V, E)$, and a node, $v \in V$, we prove that $A_v(G) = A_v(G[K_v(G)])$. Observe that $mc^*_v(S)$ is not influenced by any node lying outside $K_v(G)$. Thus, for every $S \subseteq V \setminus \{v\}$, we have: $mc^*_v(S) = mc^*_v(S \cap K_v(G))$. Next, we rewrite (5) as follows:

$$A_v(G) = \sum_{S \subseteq K_v(G) \setminus \{v\}} mc^*_v(S) \sum_{P \subseteq V \setminus K_v(G)} \frac{1}{|V|\binom{|V|-1}{|S \cup P|}}.$$

Simple calculations show that for every $S \subseteq K_v(G) \setminus \{v\}$ we have:

$$\sum_{P \subseteq V \setminus K_v(G)} \frac{1}{|V|\binom{|V|-1}{|S \cup P|}} = \frac{1}{|K_v(G)|\binom{|K_v(G)|-1}{|S|}}.$$

Thus, $A_v(G) = A_v(G[K_v(G)])$, i.e., the Attachment centrality satisfies Locality. This concludes the proof of Lemma 1. □

Theorem 2. *The Attachment centrality is the unique centrality index that satisfies Locality, Normalization, Fairness, and Gain-loss.*

Proof. In Lemma 1 we proved that the Attachment centrality satisfies Locality, Normalization, Fairness and Gain-loss. It remains to prove that the Attachment centrality is the only such possible centrality index. To put it differently, given a centrality index, F, that satisfies those requirements, it remains to prove that $F_v(G) = A_v(G)$ for any graph $G = (V, E)$ and any node $v \in V$. We will do so by showing that the sum of the F indices of all nodes belonging to the same connected component S equals: $f^*(S) = 2(|S| - 1)$. In so doing, we show that F satisfies *Component Efficiency* for game f^*. Since F also satisfies *Fairness*, then based on Myerson's result, this index is unique.

Let $G = (V, E)$ be a star with node v being the center, and let u be an arbitrary node in $V \setminus \{v\}$. Normalization implies that:

$$F_v(G) = n - 1.$$

Next, we show that $F_u(G) = 1$. To this end, consider the graph obtained from G by removing the edge $\{v, u\}$. Since u is isolated in $G - \{v, u\}$, Normalization implies that:

$$F_u(G - \{v, u\}) = 0,$$

and from Locality, we know that:

$$F_v(G - \{v, u\}) = F_v(G[V \setminus \{u\}]) = n - 2.$$

Now since the Fairness requirement implies that the removal of $\{v, u\}$ affects the centrality indices of both v and u equally, then:

$$F_u(G) = F_u(G - \{v, u\}) + (F_v(G) - F_v(G - \{v, u\})) = 1.$$

As node u was chosen arbitrarily from the set $V \setminus \{v\}$, we conclude that every node other than the center of the star has a centrality index of 1. Thus, the sum of indices in a star equals $2(n-1)$. As every connected graph can be obtained from a star by adding and removing edges, the Gain-loss requirement implies that the sum of indices in any connected graph with n nodes equals $2(n-1)$. Finally, Locality implies that the sum of indices in any connected component S equals $2(|S|-1)$. This concludes the proof of Theorem 2. □

Interestingly, replacing Gain-loss by Monotonicity in this axiomatization leads to the Degree centrality, as stated in the following theorem.

Theorem 3. *The Degree centrality is the unique centrality index that satisfies Locality, Normalization, Fairness, and Monotonicity.*

Proof. We begin by proving that the Degree centrality satisfies Locality, Normalization, Fairness, and Monotonicity. For any graph, $G = (V, E)$, and any node, $v \in V$:

- Locality: $D_v(G)$ depends solely on the connected component containing v in G, meaning that Locality is met;
- Normalization: $D_v(G) = 0$ when v is isolated, and $D_v(G) = n - 1$ when G is a star the center of which is v. It also holds that: $0 \le D_v(G) \le n - 1$. Thus, Normalization is met;
- Monotonicity: $D_v(G)$ does not decrease by adding the edge $\{u, w\}$ for any $u, w \in V : u \neq w$, meaning that Monotonicity is met;
- Fairness: adding an edge, $\{v, u\}$ for any $v, u \in V : v \neq u$, increases the Degree centrality of both v and u by 1. Therefore, Fairness is met.

It remains to prove that Degree centrality is the only possible centrality index satisfying the above four requirements, i.e., assuming that F is a centrality index that satisfies those requirements, it remains to prove that $F_v(G) = D_v(G)$ for any graph $G = (V, E)$ and any node $v \in V$. We will do so by first proving that $F_v(G) \ge D_v(G)$ and then proving that $F_v(G) \le D_v(G)$.

Step 1: In this step, we will prove that:

$$F_v(G) \ge D_v(G), \text{ for every } v \in V. \tag{12}$$

Let us fix a node, $v \in V$, and remove all the edges from G except those connecting v to its neighbors. In so doing, we obtain a new graph, G', such that $K_v(G')$ forms a star the center of which is v. Since v is the center of the star $G'[K_v(G')]$, then based on Locality and Normalization:

$$F_v(G') = F_v(G'[K_v(G')]) = |K_v(G')| - 1 = D_v(G).$$

Furthermore, from Monotonicity we know that: $F_v(G) \ge F_v(G')$, which concludes the proof of the correctness of (12).

Step 2: In this step, we will prove that:

$$F_v(G) \le D_v(G), \text{ for every } v \in V. \tag{13}$$

To this end, for any set of nodes, V, and any $v \in V$, let $\mathcal{G}^{\dagger}_v(V) \subseteq \mathcal{G}^V$ denote the set of all possible graphs such that all nodes in $V \setminus \{v\}$ form a clique, i.e.:

$$\mathcal{G}^{\dagger}_v(V) = \{(V, E) : \{u, w\} \in E \text{ for every } u, w \in V \setminus \{v\}, u \neq w\}.$$

We will prove by induction over the degree of v that the following holds:

$$F_v(G) = D_v(G), \text{ for every } G \in \mathcal{G}^{\dagger}_v(V). \tag{14}$$

Specifically, for any graph, $G \in \mathcal{G}^{\dagger}_v(V)$, if $|N_G(v)| = 0$, then v is isolated and (14) holds from Normalization. Now, assume that (14) holds for all $G \in \mathcal{G}^{\dagger}_v(V)$ such that $|N_G(v)| \le k$ for some integer $k < n - 1$. We will prove that (14) also holds for every graph $G \in \mathcal{G}^{\dagger}_v(G)$ such that $|N_G(v)| = k + 1$. Let G be one such graph, and let $u \in V$ be a neighbor of v in G. Then, if we remove the edge $\{v, u\}$ from G, we obtain a new graph, $G - \{v, u\} \in \mathcal{G}^{\dagger}_v(G)$, in which the degree of v equals k. From the inductive assumption we know that $F_v(G - \{v, u\}) = D_v(G - \{v, u\}) = k$. We also know from (12) that $F_v(G) \ge D_v(G) = k + 1$. This means that removing $\{v, u\}$ from G decreases the centrality of v by *at least* 1.

$$F_v(G) - F_v(G - \{v, u\}) \geq 1. \tag{15}$$

Now let us turn our attention back to node u. Observe that:

- $F_u(G) = n - 1$. This is because u is connected to v in G (since u and v are neighbors) and connected to every other node in G (since $V \setminus \{v\}$ forms a clique). Thus, based on Normalization and (12), it holds that: $F_u(G) = n - 1$;
- $F_u(G - \{v, u\}) \geq n - 2$. This follows immediately from (12) and the fact that node u has $n - 2$ neighbors in $G - \{v, u\}$.

The above two observations imply that, with the removal of $\{v, u\}$ from G, the centrality of u decreases by *at most* 1. That is,

$$F_u(G) - F_u(G - \{v, u\}) \leq 1. \tag{16}$$

Based on (15) and (16), as well as the Fairness requirement, we conclude that: $F_v(G) - F_v(G - \{v, u\}) = 1$, implying that $F_v(G) = D_v(G)$. This concludes our proof of the correctness of (14).

Let us now move back to our original goal—proving the correctness of (13), i.e., proving that $F_v(G) \leq D_v(G)$ for every $G \in \mathcal{G}^V$ and $v \in V$. To this end, we know from (14) that if $V \setminus \{v\}$ happened to form a clique in G, then $F_v(G) = D_v(G)$. If $V \setminus \{v\}$ did not form a clique in G, then we can add every missing edge between pairs of nodes in $V \setminus \{v\}$. In so doing, we end up with a new graph, G', in which $V \setminus \{v\}$ forms a clique and $D_v(G) = D_v(G')$. Thus, based on Monotonicity and (14), we have:

$$F_v(G) \leq F_v(G') = D_v(G') = D_v(G).$$

This concludes the proof of Theorem 3. □

We end this section by showing the independence of axioms for both axiomatizations. We use Iverson brackets: $[P] = 1$ if the logical condition P is true, and $[P] = 0$, otherwise.

- $F_v(G) = D_v(G)$ violates Gain-loss, but satisfies Locality, Normalization, Fairness, and Monotonicity;
- $F_v(G) = A_v(G)$ violates Monotonicity, but satisfies Locality, Normalization, Fairness, and Gain-loss;
- $F_v(G) = |K_v(G)| - 1$ violates Fairness, but satisfies Locality, Normalization, Monotonicity, and Gain-loss;
- $F_v(G) = 0$ violates Normalization, but satisfies Locality, Fairness, Monotonicity, and Gain-loss;
- $F_v(G) = (n - 1) \cdot [G \text{ is connected}]$ violates Locality, but satisfies Normalization, Fairness, Monotonicity, and Gain-loss.

Note that Theorems 2 and 3 imply that there exist no centrality index that satisfies all five axioms: Locality, Normalization, Fairness, Monotonicity, and Gain-loss.

5. Properties

In this section, we discuss the properties of the Attachment centrality. We begin by analyzing in more detail how adding an edge impacts the Attachment centrality. We show, in particular, that there exists a strong relation between the Attachment centrality and the notion of induced paths. Later on, building upon this analysis, we consider how the presence of *cut cliques*, *bridges*, and *leaves* in a graph affects the Attachment centrality.

5.1. Induced paths and the Attachment Delta

To better understand how the Attachment centrality works, let us introduce the concept of the *Attachment Delta*, which is a centrality index based on the Myerson value parametrized by two nodes, $s, t \in V$. Basically, for every node, v, this index reflects the role that v plays in *connecting* s with t. As we will show next, the Attachment Delta also reflects how the addition of the edge $\{s, t\}$ affects the Attachment centrality.

Definition 2. For two nodes $s, t \in V$, the *Attachment Delta* is the centrality index defined for every graph, $G = (V, E)$, and every node, $v \in V$, as:

$$\Delta_v^{s,t}(G) = 2 \cdot MV_v(u_{\{s,t\}}, G) = \frac{2 \cdot |\{\pi \in \Omega(V) : (s \sim_{G[S_v^\pi \cup \{v\}]} t) \wedge (s \not\sim_{G[S_v^\pi]} t)\}|}{|V|!}. \tag{17}$$

Recall that $(s \sim_G t)$ means that nodes s and t are connected in graph G.[3] From the definition, we see that the Attachment Delta assigns non-zero value only to the nodes for which there exists a permutation of nodes such that in the subgraph induced by the predecessors of v:

[3] Note that we slightly abuse this notation and use it also for nodes which are outside of graph $G = (V, E)$. This is because, if $s, t \notin V$, then obviously they are not connected in G: $s \not\sim_G t$.

- nodes s and t are not connected; but
- after the addition of node v and its edges, nodes s and t become connected.

Alternatively, we can say that v *connects* s and t in some (induced) subgraph. Formally, s and t are not connected in $G[S]$, but they are connected in $G[S \cup \{v\}]$ for some $S \subseteq V \setminus \{v\}$.

Before we move on to analyze the Attachment Delta in detail, let us motivate this analysis by showing the connection of the Attachment Delta to the Attachment centrality. In a nutshell, the Attachment Delta of node $v \in V \setminus \{s,t\}$ is equal to the decrease in the Attachment centrality of node v caused by the addition of edge $\{s,t\}$. To put it differently, the Attachment Delta represents the role that v plays in connecting nodes s and t which is lost when the direct connection between s and t is added to the graph.

Lemma 4. *For every graph $G = (V, E)$, edge $\{s,t\} \notin E$ and node $v \in V \setminus \{s,t\}$:*

$$\Delta_v^{s,t}(G) = A_v(G) - A_v(G + \{s,t\}).$$

Moreover, $\Delta_v^{s,t}(G) = A_v(G) - A_v(G + \{s,t\}) + 1$ if $v \in \{s,t\}$.

Proof. Let us denote $G + \{s,t\}$ by G'. Using the Iverson brackets, our formula can be rewritten as follows:

$$\Delta_v^{s,t}(G) - [v \in \{s,t\}] = A_v(G) - A_v(G'). \tag{18}$$

We begin by showing that:

$$\Delta_v^{s,t}(G') = [v \in \{s,t\}]. \tag{19}$$

Fix $v \in V$ and consider an arbitrary coalition $S \subseteq V \setminus \{v\}$. Since the edge $\{s,t\}$ is in G', the condition $(s \sim_{G'[S]} t)$ is satisfied if and only if $s,t \in S$. Thus, we get:

$$(s \sim_{G'[S \cup \{v\}]} t) \Leftrightarrow (s \sim_{G'[S]} t), \quad \text{if } v \in V \setminus \{s,t\},$$

which combined with (17) implies that $\Delta_v^{s,t}(G') = 0$. Next, assume otherwise that $v \in \{s,t\}$. Without loss of generality, let $v = s$. Since $s \notin S$, we have that $(s \not\sim_{G'[S]} t)$ and that:

$$(s \sim_{G'[S \cup \{v\}]} t) \wedge (s \not\sim_{G'[S]} t) \Leftrightarrow (t \in S), \quad \text{if } v = s.$$

By using (17) we get: $\Delta_s^{s,t}(G + \{s,t\}) = 2 \cdot |\{\pi \in \Omega(V) : t \in S_s^{\pi}\}|/|V|! = 1$, because t is before s in half of the permutations, and after s in the other half. Hence, we proved the correctness of (19). Now, by combining (18) and (19) we get that:

$$A_v(G) - A_v(G') = \Delta_v^{s,t}(G) - \Delta_v^{s,t}(G'). \tag{20}$$

From Definitions 1 and 2 and Additivity of the Shapley value, we conclude that, in order to prove (20), it is sufficient to show the equality of the corresponding coalitional games:

$$f^*/G - f^*/(G') = 2 \cdot u_{\{s,t\}}/G - 2 \cdot u_{\{s,t\}}/(G').$$

This, however, based on (4) and (7), is equivalent to:

$$K(G'[S]) - K(G[S]) = [s \sim_{G[S]} t] - [s \sim_{G'[S]} t], \quad \text{for every } S \subseteq V, \tag{21}$$

where $[P]$ denotes the Iverson brackets. If $s \notin S$ or $t \notin S$, then clearly both sides of the equation simplify to 0. Assume $s,t \in S$. From the fact that s and t are connected by an edge in G', we know that $(s \sim_{G'[S]} t)$. Also, $1 - [P] = [\neg P]$. Thus, we get that (21) is equivalent to:

$$K(G'[S]) = K(G[S]) - [s \not\sim_{G[S]} t], \quad \text{for every } S \subseteq V : s,t \in S.$$

To see why this is case, note that if s and t are not connected in $G[S]$, then the number of components in $G'[S]$ is smaller than in $G[S]$ by one, i.e., $K(G'[S]) = K(G[S]) - 1$; otherwise, the numbers of components in both graphs are the same: $K(G'[S]) = K(G[S])$. This concludes the proof of Lemma 4. □

The Attachment Delta can be considered as an alternative to assessing the role of a node in connecting two nodes by looking at either the shortest paths, as in the classical Betweenness centrality [21] or the flow, as in the Flow Betweenness centrality [23]. In the following example, we compare it with these alternatives.

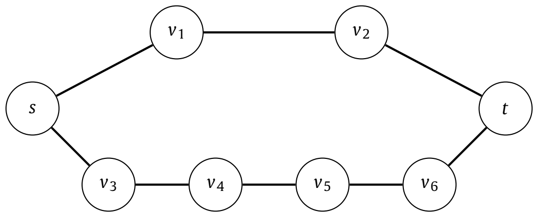

Fig. 3. A sample graph of eight nodes with two paths between s and t, which are (s, v_1, v_2, t) and $(s, v_3, v_4, v_5, v_6, t)$.

Example 2. Given the graph in Fig. 3, consider the following question: what is the role of nodes $v_1, \dots, v_6$ in connecting s and t? According to the Betweenness centrality, the role of a node v reflects the proportion of the shortest paths between s and t that v belongs to. More specifically:

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

This implies that only the nodes v_1 and v_2 are important in connecting s and t, since we have: $(B_{v_1}^{s,t}(G), \dots, B_{v_6}^{s,t}(G)) = (1, 1, 0, 0, 0, 0)$.

Now, consider the Flow Betweenness centrality [23]. According to this index, a node is as important as its (relative) marginal contribution to the *flow* between both nodes, i.e.:

$$FB_v^{s,t}(G) = \frac{\mathit{flow}_{s,t}(G) - \mathit{flow}_{s,t}(G[V \setminus \{v\}])}{\mathit{flow}_{s,t}(G)}.$$

For an unweighted graph, the amount of the flow is equivalent to the number of edge-independent paths from s and t. As there exist two paths between s and t in graph G, we have $\mathit{flow}_{s,t}(G) = 2$, and for every node v_i we have $\mathit{flow}_{s,t}(G[V \setminus \{v_i\}]) = 1$. This implies that the nodes $v_1, \dots, v_6$ are equally important in terms of connecting s and t, since we have: $(FB_{v_1}^{s,t}(G), \dots, FB_{v_6}^{s,t}(G)) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})$.

Arguably, both of these results are counterintuitive. On one hand, the nodes $v_3, \dots, v_6$ seem to have some non-zero role in connecting s and t. On the other hand, this role seems smaller than the role of nodes v_1, v_2 since those latter nodes form the shorter path between s and t. Viewed from a different perspective, node v_1 requires only one other node—v_2—in order to provide a connection between s and t, whereas node v_3 requires three other nodes—v_4, v_5, v_6—in order to provide such a connection. Thus, it seems counterintuitive to consider v_1 and v_3 indistinguishable in terms of the role they play in connecting s and t.

Compared to the outcomes of $B^{s,t}$ and $FB^{s,t}$, the outcome of the Attachment Delta seems consistent with this intuition, since we have: $(\Delta_{v_1}^{s,t}(G), \dots, \Delta_{v_6}^{s,t}(G)) = (\frac{1}{4}, \frac{1}{4}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12})$. For instance, let us explain where the value $\Delta_{v_1}^{s,t}(G) = \frac{1}{4}$ comes from. In any permutation, π, the node v_1 connects s and t in the subgraph induced by the predecessors of v_1 if (i) all the nodes s, t, v_2 are before v_1, and (ii) not all the nodes v_3, v_4, v_5, v_6 are before v_1. Condition (i) is satisfied in $\frac{1}{4}$ of all the permutations, and among those, condition (ii) is *not* satisfied only when all nodes are before v_1, i.e., v_1 is last in the permutation, which happens in $\frac{1}{8}$ of all permutations. Thus, $\Delta_{v_1}^{s,t}(G) = 2(\frac{1}{4} - \frac{1}{8}) = \frac{1}{4}$. Analogously, it can be shown that $\Delta_{v_3}^{s,t}(G) = 2(\frac{1}{6} - \frac{1}{8}) = \frac{1}{12}$. This concludes Example 2.

Observe that both paths (s, v_1, v_2, t) and $(s, v_3, v_4, v_5, v_6, t)$ in Fig. 3 are induced. The following lemma shows that the Attachment Delta of a node is non-zero if and only if it is on an induced path.

Lemma 5. *For every graph $G = (V, E)$ and nodes $v, s, t \in V$, $\Delta_v^{s,t}(G) \neq 0$ if and only if there exists an induced path, $p \in \Pi_i(s,t)$, such that $v \in p$.*

Proof. From Definition 2, we have that $\Delta_v^{s,t}(G) \geq 0$ and $\Delta_v^{s,t}(G) > 0$ if and only if there exists a coalition $S \subseteq V \setminus \{v\}$ such that $(s \not\sim_{G[S]} t)$, but $(s \sim_{G[S \cup \{v\}]} t)$. Thus, we need to prove that such a coalition exists if and only if there exists an induced path from s to t that traverses v:

$$(\exists p \in \Pi_i(s,t) : v \in p) \Leftrightarrow \left(\exists S \subseteq V : (s \not\sim_{G[S]} t) \wedge (s \sim_{G[S \cup \{v\}]} t)\right).$$

$\Rightarrow$: First, observe that if there exists an induced path $p = (u_1, \dots, u_m) \in \Pi_i(s,t)$ such that $u_1 = s$, $u_m = t$ and $u_k = v$ for some $k \in \{1, \dots, m\}$, then $S = \{u_1, \dots, u_m\} \setminus \{u_k\}$ is the coalition that we are looking for: since p is an induced path, there exists no edge between $u_1, \dots, u_{k-1}$ and $u_{k+1}, \dots, u_m$; thus, $(s \not\sim_{G[S]} t)$, but $(s \sim_{G[S \cup \{v\}]} t)$.

$\Leftarrow$: Now, assume there exists $S \subseteq V$ such that $(s \not\sim_{G[S]} t)$ and $(s \sim_{G[S \cup \{v\}]} t)$. We will show that v is on some induced path between s and t. To this end, consider the shortest path from s to v and from v to t in $G[S \cup \{v\}]$: $p_s = (u_1, \dots, u_k)$, $p_t = (u_k, \dots, u_m)$ such that $u_1 = s$, $u_k = v$ and $u_m = t$. We will prove by contradiction that $p^* = (u_1, \dots, u_m)$ is an induced path.

To this end, consider two non-consecutive nodes u_i, u_j, $1 \le i < j \le m$, $j \ne i+1$ and assume $\{u_i, u_j\} \in E$. If $j \le k$, then $u_i, u_j \in p_s$ which contradicts the assumption that p_s is the shortest path between s and v. Analogously, if $i \ge k$, then $u_i, u_j \in p_t$ which implies that p_t is not the shortest path between v and t. If $i < k$ and $j > k$, then $(u_1, \dots, u_i, u_j, \dots, u_m)$ is a path in $G[S]$ which contradicts the fact that $(s \not\sim_{G[S]} t)$. This shows that p^* is an induced path and concludes the proof of Lemma 5. □

An immediate conclusion from Lemmas 4 and 5 is the following theorem.

Theorem 6. *Adding an edge $\{s, t\}$ to a graph G affects only the Attachment centrality of s, t, and the nodes belonging to any of the induced paths between s and t.*

Proof. Fix $v \in V \setminus \{s, t\}$. From Lemma 4, we know that $A_v(G) - A_v(G + \{s, t\}) = \Delta_v^{s,t}(G)$. Thus, from Lemma 5 we get that $A_v(G) - A_v(G + \{s, t\}) \ne 0$ if and only if v lies on an induced path between s and t. Now, if $v \in \{s, t\}$, then from Lemma 4 we have that $A_v(G) - A_v(G + \{s, t\}) = \Delta_v^{s,t}(G) - 1$. Since $\Delta_v^{s,t}(G) = 1$ if and only if edge $\{s, t\}$ is already in G, we get that $A_v(G) - A_v(G + \{s, t\}) \ne 0$. □

Let us summarize our analysis of the effects of adding edge $\{s, t\}$ to graph G on the role of the nodes in connecting s and t:

- if node v is not on any induced path between s and t (did not *connect* them), then its Attachment centrality does not change;
- if node v, other than s and t, is on an induced path between s and t, then v's role in connecting these nodes (equal to $\Delta_v^{s,t}(G)$) is deduced from its Attachment centrality;
- if $v \in \{s, t\}$, then v's role in connecting s and t (equal to $\Delta_v^{s,t}(G)$) is deduced from its Attachment centrality and 1 (which represents the role in connecting s and t in graph $G + \{s, t\}$) is added.

Building upon our analysis, the Attachment centrality can be constructed from the Attachment Delta using the following scheme: (1) start with an empty graph $G^* = (V, \emptyset)$ and set the centrality index of each node to zero, i.e., $A_v(G^*) = 0$ for every $v \in V$; (2) add to G^* edges from E, one by one, and for every such edge update the centralities according to Lemmas 4 and 5. Moreover, the order in which edges are added would not affect the final result. In particular, by adding edges of v at the very end of the process, we get the following reformulation of the Attachment centrality:

Theorem 7. *For every graph $G = (V, E)$, node $v \in V$ and $N_G(v) = (u_1, \dots, u_k)$:*

$$A_v(G) = |N_G(v)| - \sum_{u_i \in N_G(v)} \Delta_v^{v,u_i}(G - \{\{v, u_j\} : j \le i\}).$$

Proof. For $i \in \{0, \dots, k\}$, let us denote by G_i the graph obtained from G by removing i edges of node v to nodes $u_1, \dots, u_i$. Formally: $G_i = G - \{\{v, u_j\} : j \le i\}$. We need to prove that:

$$A_v(G) = |N_G(v)| - \sum_{u_i \in N_G(v)} \Delta_v^{v,u_i}(G_i).$$

Observe that $G_0 = G$. On the other hand, in graph G_k, node v does not have any edges, i.e., it is isolated. From Normalization we have that the Attachment centrality of an isolated node equals zero, so $A_v(G_k) = 0$.

From Lemma 4 and the fact that $G_{i-1} = G_i + \{v, u_i\}$, we have that:

$$A_v(G_{i-1}) - A_v(G_i) = -\Delta_v^{v,u_i}(G_i) + 1, \quad \text{for every } i \in \{1, \dots, k\}. \tag{22}$$

Summing (22) over all $i \in \{1, \dots, k\}$ we get that:

$$A_v(G_0) - A_v(G_k) = -\sum_{i \in \{1,\dots,k\}} \Delta_v^{v,u_i}(G_i) + k.$$

Since we know that $|N_G(v)| = k$, $i \in \{1, \dots, k\}$ is equivalent to $u_i \in N_G(v)$, $G_0 = G$ and $A_v(G_k) = 0$ this concludes the proof of Theorem 7. □

In the next subsection, we show how the relation between the Attachment centrality and induced paths can be used to study how cut cliques, bridges and leaves in a graph affect the Attachment centrality.

5.2. Cut cliques, bridges and leaves

The following result plays the crucial role in our analysis how the presence of cut cliques in the graph affects the Attachment centrality.

Theorem 8. *If $G=(V,E)$ is a connected graph such that $C \subseteq V$ forms a clique and $K(G[V \setminus C])=\{S_1,\dots,S_k\}$, then:*

$$A_v(G)=\begin{cases} A_v(G[S_i \cup C]) & \text{if } v \in S_i, \\ \left(\sum_{i\in\{1,\dots,k\}} A_v(G[S_i \cup C])\right) - \frac{2(|C|-1)}{|C|}(k-1) & \text{if } v \in C. \end{cases} \quad (23)$$

Proof. First, let us focus on an arbitrary node, $v \in S_i$ for some $i \in \{1,\dots,k\}$. We need to show that $A_v(G)=A_v(G[S_i \cup C])$. From Locality, we know that $A_v(G[S_i \cup C])=A_v(V,E[S_i \cup C])$, so equivalently we need to show that $A_v(G)=A_v(V,E[S_i \cup C])$. To put it in words, our goal is to prove that the Attachment centrality of node v is not affected by the removal of any of the edges from the components other than $S_i \ni v$ that emerge after removing cut clique C. To this end, we begin by proving that removing *one* such edge does not affect the Attachment centrality:

(*) *For $G=(V,E)$, if C forms a clique and $S_i \in K(G[V \setminus C])$, then $A_v(G)=A_v(G-\{u,w\})$ for every $\{u,w\} \in E \setminus E[S_i \cup C]$.*

Since S_i is a component in $G[V \setminus C]$, then edges from $E \setminus E[S_i \cup C]$ are not incident to nodes from S_i, i.e., $u,w \notin S_i$. Now recall that based on Theorem 6 it is sufficient to show that v is not on any induced path between u and w in $G-\{u,w\}$. Assume otherwise: let $p \in \Pi_i(u,w)$ be an induced path such that $v \in p$. This path consists of two parts: a path from u to v and a path from v to w. Since $v \in S_i$ and $u,w \notin S_i$, we know that both of these paths contain a node from C. This implies that p is not induced, as it contains the same node from C twice or two different non-consecutive nodes from C which—by the definition—are connected by an edge. Thus, we proved that (*) holds.

To show that *all* edges from $E \setminus E[S_i \cup C]$ can be removed, it is sufficient to prove that after removing one edge, $\{u,w\}$, the graph $G'=G-\{u,w\}$ also satisfies the assumptions of (*), i.e., C forms a clique in G' and $S_i \in K(G'[V \setminus C])$. To see why this is the case, observe that the removed edge is not from $E[C]$ (i.e., C still forms a clique in G') and the edges of the nodes from S_i have remained intact (i.e., S_i is a component in $K(G'[V \setminus C])$). This implies that, starting from graph G, we can remove edges from $E \setminus E[S_i \cup C]$ one by one and (*) implies that the Attachment centrality of node v remains intact. Thus, after removing all such edges we get the graph $(V,E[S_i \cup C])$ and we get $A_v(G)=A_v(V,E[S_i \cup C])$, which concludes the first part of the proof.

Now let us turn our attention to an arbitrary node $v \in C$ and denote the graph $(V,E[S_i \cup C])$ by G_i. From Locality we know that $A_v(G[S_i \cup C])=A_v(G_i)$. Thus, we need to prove that:

$$A_v(G)=\sum_{i\in\{1,\dots,k\}} A_v(G_i)-\frac{2(|C|-1)}{|C|}\cdot(k-1). \quad (24)$$

To this end, let us analyze the marginal contribution of node v to an arbitrary coalition $S \subseteq V \setminus \{v\}$ in game f^*/G and in game f^*/G_i for every $i \in \{1,\dots,k\}$. Recall that based on Definition 1, $f^*/G(S \cup \{v\})-f^*/G(S)=|K(G[S])|-|K(G[S \cup \{v\}])|+1$, which is equivalent to the number of components of $G[S]$ that are connected to v.

Without loss of generality, let $K(G[S])=\{T_1,\dots,T_l\}$ be the components of S and assume that v is connected to the first m components: $T_1,\dots,T_m$ for $m \le l$. Observe that if component T_j contains a node from C, then it contains all nodes from $S \cap C$. This implies that there exists at most one component that contains nodes from C (and if it exists, then it must be connected to v) and all other components are subsets of components $S_1,\dots,S_k$. This can be formalized as follows:

$$m=|\{j \in \{1,\dots,m\}: \exists_i T_j \subseteq S_i\}|+[S \cap C \neq \emptyset]. \quad (25)$$

Recall that in graph G_i there are only edges between nodes from $S_i \cup C$. So, node $v \in C$ in graph G_i is connected only to components T_j that are subsets of component S_i and to nodes from $S \cap C$, if there exists:

$$K(G_i[S])-K(G_i[S \cup \{v\}])+1=|\{j \in \{1,\dots,m\}: T_j \subseteq S_i\}|+[S \cap C \neq \emptyset].$$

Summing over all $i \in \{1,\dots,k\}$ and using (25) we get:

$$\begin{aligned}\sum_{i\in\{1,\dots,k\}} (K(G_i[S])-K(G_i[S \cup \{v\}])+1) &= m+(k-1)\cdot[S \cap C \neq \emptyset] \\ &= (K(G[S])-K(G[S \cup \{v\}])+1)+(k-1)\cdot[S \cap C \neq \emptyset].\end{aligned}$$

Thus, using Definition 1, we get that

$$\sum_{i\in\{1,\dots,k\}} A_v(G_i)=A_v(G)+(k-1)\cdot\frac{2|\{\pi \in \Omega(V): S_v^\pi \cap C \neq \emptyset\}|}{|V|!}.$$

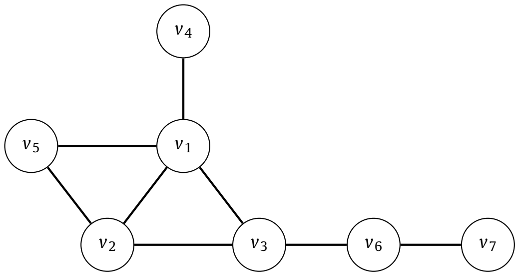

Fig. 4. A sample graph of 7 nodes. For a cut clique $C = \{v_1, v_2, v_3\}$, we have $K(G[V \setminus C]) = \{\{v_4\}, \{v_5\}, \{v_6, v_7\}\}$. Theorem 8 implies, in particular, that $A_{v_4}(G) = A_{v_4}(G[C \cup \{v_4\}])$ and $A_u(G) = A_u(G[C \cup \{v_4\}]) + A_u(G[C \cup \{v_5\}]) + A_u(G[C \cup \{v_6, v_7\}]) - 8/3$ for every $u \in C$.

In permutation π, it holds that $S^{\pi}_{v} \cap C \neq \emptyset$ if and only if v is not the first node from C. This happens with probability $(|C| - 1)/|C|$, which concludes the proof of Theorem 8. □

See Fig. 4 for an illustration of the implications of Theorem 8. From this theorem, we know that if there exists a cut clique, then the Attachment centrality of a graph can be computed based on the Attachment centrality of subgraphs determined by the cut clique. We will use this fact as one of the foundations of our algorithmic analysis of the Attachment Centrality in Section 6.

Now we will analyze how the removal of a *cut vertex* or a *bridge* influences the Attachment centrality of other nodes.

Corollary 9. *The Attachment centrality satisfies the following properties:*

(a) (Leaf) The centrality of a leaf equals 1. *Moreover, removing a leaf from the graph decreases the centrality of its only neighbor by* 1 *and does not affect the centrality of other nodes.*
(b) (Cut vertex) The centrality of a cut vertex equals the sum of its centralities in the graphs induced by itself and the nodes from each component created by its removal.
(c) (Bridge) Removing a bridge decreases the centrality of both its ends by 1 and does not affect the centrality of other nodes.

Proof. We will prove that the Attachment centrality satisfies these properties one by one.

(a) Assume v is a leaf and $N(v) = \{u\}$. Clearly, u is a cut clique of size 1 and Theorem 8 implies that $A_v(G) = A_v(\{v, u\}, \{\{v, u\}\}) = 1$ (from Normalization), $A_u(G) = A_u(G[V \setminus \{v\}]) + 1$, and $A_w(G) = A_w(G[V \setminus \{v\}])$ for $w \in V \setminus \{v, u\}$.
(b) Assume G is connected and $K(G[V \setminus \{v\}]) = \{S_1, \dots, S_k\}$. From Theorem 8 for $C = \{v\}$ we get that $A_v(G) = \sum_{i \in \{1,\dots,k\}} A_v(G[S_i \cup \{v\}])$.
(c) Assume $\{v, u\}$ is a bridge and $K(G - \{v, u\}) = \{S_v, S_u\}$. Since both v and u are cut vertices, from Locality and the analysis of cut vertices we get that $A_w(G - \{v, u\}) = A_w(G[S_v]) = A_w(G)$ for every $w \in S_v$ and, analogously, $A_w(G[S_u]) = A_w(G)$ for every $w \in S_u$. On the other hand, $A_v(G) = A_v(G[S_v]) + A_v(G[S_u \cup \{v\}])$. Since v is a leaf in $G[S_u \cup \{v\}]$, from the analysis of leaves we get that $A_v(G) = A_v(G - \{v, u\}) + 1$. In the same way we can prove that $A_u(G) = A_u(G - \{v, u\}) + 1$.

This concludes the proof of Corollary 9. □

Building upon the above analysis, in the next section we study the computational complexity of the Attachment centrality.

6. Complexity results

In this section, we study the computational complexity of the Attachment centrality. First, we show that computing the Attachment centrality is #**P**-complete. Then, we propose a polynomial algorithm for chordal graphs and, building upon this algorithm, we propose a general-purpose algorithm with the exponential pessimistic time complexity. Finally, we test our algorithm on the terrorist network responsible for the Madrid train bombing in 2004.

6.1. Hardness result

To show that computing the Attachment centrality is #**P**-complete, we first show that computing the Attachment Delta is #**P**-complete. To this end, we will use the reduction from #3-SAT—the problem of counting the number of satisfying assignments of a given Boolean formula in the 3-CNF form. This result, combined with Lemma 4, will imply that computing the Attachment centrality is also #**P**-complete.

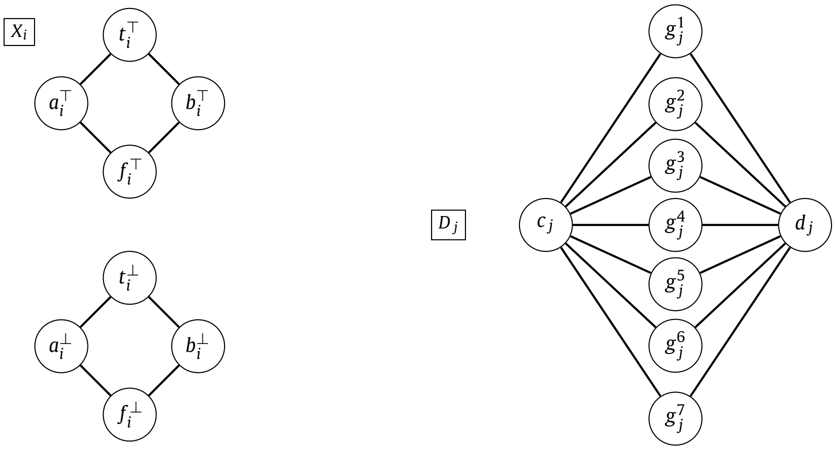

Fig. 5. Graphs X_i and D_j used in the proof of Theorem 10.

Theorem 10. *Computing the Attachment Delta is #**P**-complete.*

Proof. The value $n!\Delta_v^{s,t}(G)$ can be considered as the number of accepting paths of nondeterministic Turing machine, so the problem is in #**P**.

To show that the problem is #**P**-complete, we use the reduction from the #3-SAT problem. Fix logical variables $\mathcal{X} = \{x_1, \dots, x_n\}$. Let $\mathcal{C} = \{C_1, \dots, C_m\}$ be m clauses, each of the form $C_j = z_j^1 \vee z_j^2 \vee z_j^3$, where $z_j^k \in \{x_1, \neg x_1, \dots, x_n, \neg x_n\}$ for $k \in \{1, 2, 3\}$. The *(truth) assignment* is function $\mathcal{X} \to \{0, 1\}$ that, for every variable x_i, assigns a logical value 1 (true) or 0 (false). Now, the #3-SAT problem is to determine, given clauses $C_1, \dots, C_m$, how many assignments satisfy $C_1 \wedge \dots \wedge C_m$?

We will construct a graph, G, with nodes s, t, and v such that the Attachment Delta, $\Delta_v^{s,t}(G)$, is equal to the number of assignments satisfying $C_1 \wedge \dots \wedge C_m$ multiplied by some weight, $\beta(n, m)$. Specifically, we will show that the number of truth assignments is equal to the size of set $\mathcal{S}(G)$ defined as follows:

$$\mathcal{S}(G) = \{S \subseteq V \setminus \{v\} : (s \not\sim_{G[S]} t) \wedge (s \sim_{G[S \cup \{v\}]} t)\}, \tag{26}$$

and that $\Delta_v^{s,t}(G) = \beta(n, m) \cdot |\mathcal{S}(G)|$. The construction of graph G is inspired by the work of Bienstock [8]. In more detail, G consists of:

- n subgraphs corresponding to n logical variables;
- m subgraphs corresponding to m clauses; and
- additional nodes: s, t and v.

For a variable, $x_i \in \mathcal{X}$, we define graph $X_i = (V(X_i), E(X_i))$ as follows:

$$X_i = (\{a_i^\top, b_i^\top, t_i^\top, f_i^\top, a_i^\bot, b_i^\bot, t_i^\bot, f_i^\bot\}, \{\{a_i^\circ, t_i^\circ\}, \{t_i^\circ, b_i^\circ\}, \{a_i^\circ, f_i^\circ\}, \{f_i^\circ, b_i^\circ\} : \circ \in \{\top, \bot\}\}).$$

See Fig. 5 for an illustration. Intuitively, path $(a_i^\circ, t_i^\circ, b_i^\circ)$ corresponds to $x_i = 1$, and $(a_i^\circ, f_i^\circ, b_i^\circ)$—to $x_i = 0$.

For a clause, $C_j \in \mathcal{C}$, we define graph $D_j = (V(D_j), E(D_j))$ as follows:

$$D_j = (\{c_j, d_j, g_j^1, \dots, g_j^7\}, \{\{c_j, g_j^k\}, \{g_j^k, d_j\} : k \in \{1, \dots, 7\}\}).$$

See Fig. 5 for an illustration. Intuitively, for $C_j = z_j^1 \vee z_j^2 \vee z_j^3$, if z_j^1, z_j^2, z_j^3 are literals of $x_{i_1}, x_{i_2}, x_{i_3}$, then each of the 7 paths (c_j, g_j^k, d_j) corresponds to one truth assignment $\{x_{i_1}, x_{i_2}, x_{i_3}\} \to \{0, 1\}$ that satisfies C_j. Specifically, if $k = (k_1 k_2 k_3)_2$ in a binary notation, then path (c_j, g_j^k, d_j) corresponds to $z_j^1 = k_1$, $z_j^2 = k_2$ and $z_j^3 = k_3$. This truth assignment can be equivalently characterized by the set of variables with value "true" and the set of variables with value "false", defined as follows: $T(g_j^k) = \{x_i \in \mathcal{X} : k = (k_1 k_2 k_3)_2 \wedge (\exists_{1 \le l \le 3}\ (k_l = 1 \wedge z_j^l = x_i) \vee (k_l = 0 \wedge z_j^l = \neg x_i))\}$ and $F(g_j^k) = \{x_i \in \mathcal{X} : \exists_{1 \le l \le 3}\ (z_j^l \in \{x_i, \neg x_i\})\} \setminus T(g_j^k)$.

For example, consider clause $C_1 = x_1 \vee x_2 \vee x_3$ and $k = 1 = (001)_2$. Now, path (c_1, g_1^1, d_1) corresponds to assignments in which $x_1 = 0$, $x_2 = 0$, and $x_3 = 1$ (here, $T(g_1^1) = \{x_3\}$ and $F(g_1^1) = \{x_1, x_2\}$). On the other hand, for $C_2 = \neg x_1 \vee x_2 \vee x_4$ and $k = 3 = (011)_2$, path (c_2, g_2^3, d_2) corresponds to assignments in which $\neg x_1 = 0$ (i.e., $x_1 = 1$), $x_2 = 1$, and $x_4 = 1$ (here, $T(g_2^3) = \{x_1, x_2, x_4\}$ and $F(g_2^3) = \emptyset$).

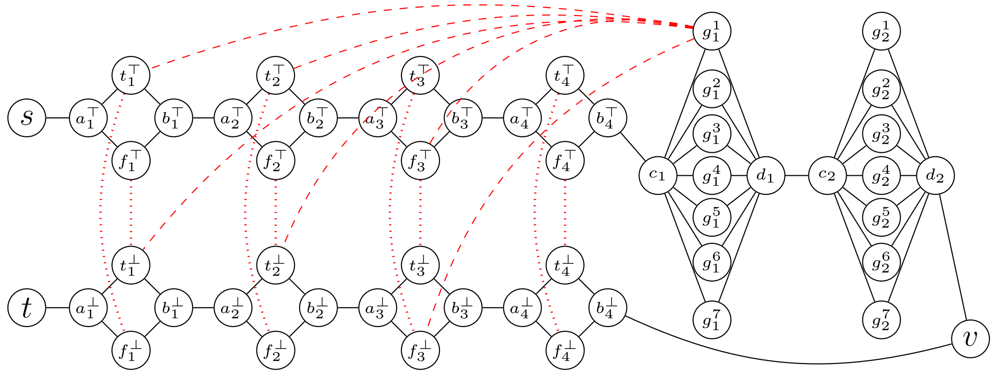

Fig. 6. Graph G for variables $x_1, \dots, x_4$ and clauses $C_1 = x_1 \vee x_2 \vee \neg x_3$ and $C_2 = \neg x_1 \vee x_2 \vee x_4$. For clarity of presentation, red edges incident to nodes $g_1^2, \dots, g_1^7$ and $g_2^1, \dots, g_2^7$ have been omitted; other red edges have been drawn with dotted lines.

We are now ready to formally define graph $G = (V, E)$. In G, the set of vertices is:

$$V = \{s, t, v\} \cup \bigcup_{x_i \in \mathcal{X}} V(X_i) \cup \bigcup_{C_j \in \mathcal{C}} V(D_j).$$

The set of edges will consists of black (*regular*) edges and red (*forbidden*) edges: $E = E_B \cup E_R$. The black (regular) edges are defined as follows:

$$E_B = \{\{s, a_1^\top\}, \{t, a_1^\perp\}\} \cup \bigcup_{x_i \in \mathcal{X}} E(X_i) \cup \{\{b_i^\top, a_{i+1}^\top\}, \{b_i^\perp, a_{i+1}^\perp\} : 1 \le i < n\} \cup$$
$$\{\{b_n^\top, c_1\}\} \cup \bigcup_{C_j \in \mathcal{C}} E(D_j) \cup \{\{d_j, c_{j+1}\} : 1 \le j < m\} \cup \{\{d_m, v\}, \{b_n^\perp, v\}\}.$$

See Fig. 6 for an illustration. The red (forbidden) edges will connect nodes that are not allowed to be together in the same set:

$$E_R = \{\{g_j^k, t_i^\top\}, \{g_j^k, t_i^\perp\}\} : C_j \in \mathcal{C}, x_i \in F(g_j^k)\} \cup$$
$$\{\{g_j^k, f_i^\top\}, \{g_j^k, f_i^\perp\}\} : C_j \in \mathcal{C}, x_i \in T(g_j^k)\} \cup \{\{t_i^\top, f_i^\perp\}, \{t_i^\perp, f_i^\top\} : x_i \in \mathcal{X}\}.$$

For every variable $x_i \in \mathcal{X}$, node $t_i^\top$ is connected by a forbidden red edges to $f_i^\perp$, and $t_i^\perp$ to $f_i^\top$. Moreover, for every clause C_j and $1 \le k \le 7$, we add 6 edges between g_j^k and nodes $t_i^\top, t_i^\perp$ for $x_i \in F(g_j^k)$, and $f_i^\top, f_i^\perp$ for $x_i \in T(g_j^k)$. For example, for $C_j = x_1 \vee x_2 \vee \neg x_3$, we already argued that $k = 1 = (001)_2$ corresponds to assignment $x_1 = 0$, $x_2 = 0$, and $x_3 = 1$. Thus, we connect node g_j^k by red edges with $t_1^\top$, $t_1^\perp$, $t_2^\top$, $t_2^\perp$, $f_3^\top$, and $f_3^\perp$. See Fig. 6 for an illustration.

In the remainder of the proof we will show a one-to-one mapping between assignments satisfying $C_1 \wedge \dots \wedge C_m$ and subsets of nodes $S \in \mathcal{S}(G)$ (see (26) for the definition). Lemma 11 maps every assignment to a subset from $\mathcal{S}(G)$. In Lemma 12, a mapping from a subset from $\mathcal{S}(G)$ to an assignment will be defined.

Lemma 11. *For every assignment $x_1, \dots, x_n$ satisfying $C_1 \wedge \dots \wedge C_m$, the set:*

$$S = \{s, t\} \cup \{a_i^\top, b_i^\top, a_i^\perp, b_i^\perp : x_i \in \mathcal{X}\} \cup \{c_j, d_j : C_j \in \mathcal{C}\} \cup$$
$$\{t_i^\top, t_i^\perp : x_i = 1\} \cup \{f_i^\top, f_i^\perp : x_i = 0\} \cup \{g_j^k : C_j \in \mathcal{C}, k = (z_j^1 z_j^2 z_j^3)_2\}$$

is in $\mathcal{S}(G)$, i.e., satisfies $(s \not\sim_{G[S]} t)$ and $(s \sim_{G[S \cup \{v\}]} t)$.

Proof. Let $tf_i^\circ = t_i^\circ$ if $x_i = 1$, and $tf_i^\circ = f_i^\circ$, otherwise, for $\circ \in \{\top, \perp\}$. Also, let $k_j = (z_j^1 z_j^2 z_j^3)_2$. First, observe that all nodes from $S \cup \{v\}$ form a path:

$$p = (s, a_1^\top, tf_1^\top, b_1^\top, \dots, a_n^\top, tf_n^\top, b_n^\top, c_1, g_1^{k_1}, d_1, \dots, c_m, g_m^{k_m}, d_m, v, b_n^\perp, tf_n^\perp, a_n^\perp, \dots, t).$$

Thus, $(s \sim_{G[S \cup \{v\}]} t)$. It remains to prove that $(s \not\sim_{G[S]} t)$. To this end, it suffices to show that p is an induced path, i.e., there exists no edge between two non-consecutive edges. Clearly, there is no such black (regular) edge. Consider red (forbidden) edges. Since either $S \cap \{t_i^\top, t_i^\perp\} = \emptyset$ or $S \cap \{f_i^\top, f_i^\perp\} = \emptyset$, we get that $\{t_i^\top, f_i^\perp\}, \{f_i^\top, t_i^\perp\} \notin E[S]$ for every $x_i \in \mathcal{X}$. Now, consider red edges from $g_j^{k_j}$ for an arbitrary $C_j \in \mathcal{C}$. We know that $k_j = (z_j^1 z_j^2, z_j^3)_2$. Without loss of generality, let us focus on z_j^1. We know that $z_j^1 = x_i$ or $z_j^1 = \neg x_i$ for some $x_i \in \mathcal{X}$. Assume $z_j^1 = 1$. It means that if $z_j^1 = x_i$, then $x_i = 1$ and there are red edges to $f_i^\top, f_i^\perp$, but since $x_i = 1$ these nodes are not in S. On the other hand, if $z_j^1 = \neg x_i$, then $x_i = 0$ and red edges connect $g_j^{k_j}$ with $t_i^\top, t_i^\perp$, but these nodes are not in S. Analogously, we can prove that if $z_j^1 = 0$, then nodes incident to red edges from $g_j^{k_j}$ are not in S. Thus, p is induced and $(s \not\sim_{G[S]} t)$. This concludes the proof of Lemma 11. □

Note that for different assignments the set S is different—if two assignments differ with respect to variable x_i, then for one assignment we have $\{t_i^\top, t_i^\perp\} \subseteq S$, and for the other one $\{t_i^\top, t_i^\perp\} \not\subseteq S$. Thus, from Lemma 11, we get that:

$$|\mathcal{S}(G)| \geq \text{number of assignments } x_1, \ldots, x_n \text{ satisfying } C_1 \wedge \cdots \wedge C_m. \tag{27}$$

In the following lemma, we show that every set from $\mathcal{S}(G)$ has the form described in Lemma 11 for some assignment; this will imply that the number of sets in $\mathcal{S}(G)$ is indeed equal to the number of assignments.

Lemma 12. *For every $S \in \mathcal{S}(G)$, assignment $x_i = 1$ if $t_i^\top \in S$, $x_i = 0$, otherwise, satisfies $C_1 \wedge \cdots \wedge C_m$; moreover, $|S| = 6n + 3m + 2$ and for two different sets $S, S' \in \mathcal{S}(G)$ these assignments are different.*

Proof. Fix $S \in \mathcal{S}(G)$, i.e., $S \subseteq V \setminus \{v\}$ such that $(s \not\sim_{G[S]} t)$ and $(s \sim_{G[S \cup \{v\}]} t)$. We begin by proving that there are no red edges in $G[S]$, i.e., $E[S] \cap E_R = \emptyset$. To this end, consider the following property:

$$P(i) = \text{ there are no red edges in } G[S] \text{ incident to nodes from } X_1, \ldots, X_i.$$

We will prove that $P(i)$ holds by induction over $i \in \{1, \ldots, n\}$. For $i = 0$, $P(0)$ trivially holds. Fix $i \geq 1$ and assume $P(i-1)$ holds. We will prove that $P(i)$ holds.

From $P(i-1)$, assumption $(s \sim_{G[S \cup \{v\}]} t)$ and the construction of the graph we get that $a_i^\top, a_i^\perp$ are on the path from s and t in $G[S \cup \{v\}]$. So, we know that:

$$((t_i^\top \in S) \vee (f_i^\top \in S)) \text{ and } ((t_i^\perp \in S) \vee (f_i^\perp \in S)). \tag{28}$$

Furthermore, from the assumption that s and t are not connected in $G[S]$, i.e., $(s \not\sim_{G[S]} t)$, we get that red edges $\{t_i^\top, f_i^\perp\}$ and $\{t_i^\perp, f_i^\top\}$ are not in $E[S]$:

$$((t_i^\top \notin S) \vee (f_i^\perp \notin S)) \text{ and } ((t_i^\perp \notin S) \vee (f_i^\top \notin S)). \tag{29}$$

By combining (28) and (29) we get: $(t_i^\top \in S) \Rightarrow (f_i^\perp \notin S) \Rightarrow (t_i^\perp \in S) \Rightarrow (f_i^\top \notin S)$. Analogously, $(t_i^\top \notin S) \Rightarrow (f_i^\top \in S) \Rightarrow (t_i^\perp \notin S) \Rightarrow (f_i^\perp \in S)$. As a result, we have that:

$$S \cap \{t_i^\top, t_i^\perp, f_i^\top, f_i^\perp\} \in \{\{t_i^\top, t_i^\perp\}, \{f_i^\top, f_i^\perp\}\}. \tag{30}$$

This implies that other red edges also cannot be in $G[S]$: we have that $\{t_i^\top, g_j^k\} \in E[S]$ is equivalent to $\{t_i^\perp, g_j^k\} \in E[S]$ for every $C_j \in \mathcal{C}$, $k \in \{1, \ldots, 7\}$, but if both conditions are satisfied, then the assumption $(s \not\sim_{G[S]} t)$ is violated. Analogously, $\{f_i^\top, g_j^k\} \in E[S]$ is equivalent to $\{f_i^\perp, g_j^k\} \in E[S]$ which also leads to the contradiction. This concludes the proof of $P(i)$.

Now observe that every edge is incident to a node from X_i. Thus, $P(n)$ implies that there are no red edges in $G[S]$. This fact combined with $(s \sim_{G[S \cup \{v\}]} t)$ imply:

- $\{a_i^\top, a_i^\perp, b_i^\top, b_i^\perp\} \in S$ for every $x_i \in \mathcal{X}$, $\{c_j, d_j\} \in S$ for every $C_j \in \mathcal{C}$ and $s, t \in S$ (every path in G between s and t that does not use red edges goes through these nodes);
- $S \cap \{t_i^\top, t_i^\perp, f_i^\top, f_i^\perp\} \in \{\{t_i^\top, t_i^\perp\}, \{f_i^\top, f_i^\perp\}\}$ for every $x_i \in \mathcal{X}$ (from (30));
- $|S \cap \{g_j^1, \ldots, g_j^7\}| = 1$ for every $C_j \in \mathcal{C}$ (since c_j and d_j are connected, we have $|S \cap \{g_j^1, \ldots, g_j^7\}| \geq 1$; if $g_j^k, g_j^{k'} \in S$ for some C_j and $1 \leq k < k' \leq 7$, then there must exist $x_i \in \mathcal{X}$ such that g_j^k is connected to $t_i^\top$ and $t_i^\perp$ by red edges, and $g_j^{k'}$ is connected to $f_i^\top$ and $f_i^\perp$ by red edges; (30) leads to the contradiction).

Thus, we know that S contains: 6 nodes from X_i for every $x_i \in \mathcal{X}$, 3 nodes from D_j for every $C_j \in \mathcal{C}$, s and t; overall $|S| = 6n + 3m + 2$.

Now, for $S \in \mathcal{S}(G)$ let us consider an assignment $x_i = [t_i^\top \in S]$. First, let us argue that a different set will result in different assignments. Assume otherwise that there exists $S, S' \in \mathcal{S}(G)$ such that $[t_i^\top \in S] = [t_i^\top \in S']$ for every $x_i \in \mathcal{X}$. Immediately from the characterization of sets from $\mathcal{S}(G)$ we get that $S \cap \{t_i^\top, t_i^\perp, f_i^\top, f_i^\perp\} = S' \cap \{t_i^\top, t_i^\perp, f_i^\top, f_i^\perp\}$ for every $x_i \in \mathcal{X}$, which implies $S \cap \{g_j^1, \ldots, g_j^7\} = S' \cap \{g_j^1, \ldots, g_j^7\}$ for every $C_j \in \mathcal{C}$. Thus, we get that $S = S'$.

It remains to prove that for assignment $x_i = [t_i^\top \in S]$ formula $C_1 \wedge \cdots \wedge C_m$ is satisfied. Fix $C_j = z_j^1 \vee z_j^2 \vee z_j^3$ and assume $g_j^k \in S$ for some $1 \leq k = (k_1 k_2 k_3)_2 \leq 7$. Since $k \geq 1$ we have $k_l = 1$ for some $l \in \{1, 2, 3\}$. We know that $z_j^l \in \{x_i, \neg x_i\}$ for some

$x_i \in \mathcal{X}$. If $z_j^l = x_i$, then node g_j^k has red edges to nodes $f_i^\top, f_i^\perp$, so they cannot be in S. From this and (30), we get that $t_i^\top, t_i^\perp \in S$. In result, based on the assignment, $x_i = 1$, $z_j^l = x_i = 1$ and C_j is satisfied. On the other hand, if $z_j^l = \neg x_i$, then node g_j^k has red edges to nodes $t_i^\top, t_i^\perp$. Hence, again, from (30) we get that $f_i^\top, f_i^\perp \in S$. In result, based on the assignment, $x_i = 0$, $z_j^l = \neg x_i = 1$ and C_j is satisfied. This concludes the proof of Lemma 12. □

Based on Lemma 12 we get that:

$$|\mathcal{S}(G)| \leq \text{number of assignments } x_1, \dots, x_n \text{ satisfying } C_1 \wedge \cdots \wedge C_m. \tag{31}$$

Combining (27) and (31) we get that the number of assignments that satisfies $C_1 \wedge \cdots \wedge C_m$ is equal to the number of coalitions $S \in \mathcal{S}(G)$. Furthermore, in Lemma 12 we proved that all sets from $\mathcal{S}(G)$ have $(6n + 3m + 2)$ nodes. Thus, from Definition 2 we have that:

$$\Delta_v^{s,t}(G) = \frac{2 \cdot |\{\pi \in \Omega(V) : (s \not\sim_{G[S_v^\pi]} t) \wedge (s \sim_{G[S_v^\pi \cup \{v\}]} t)\}|}{|V|!} = \frac{2(6n+3m+2)!(2n+6m)!}{(8n+9m+3)!} \cdot |\mathcal{S}(G)|,$$

where $(8n+9m+3)$ is the number of nodes in G and $(2n+6m)$ is the number of nodes outside $S \cup \{v\}$ for every $S \in \mathcal{S}(G)$. The equality follows the fact that for every subset, $S \subseteq V \setminus \{v\}$, there are $|S|!(|V| - |S| - 1)!$ permutations in which $S_v^\pi = S$; here, $S!$ is the number of possible orderings of nodes from S before v, and $(|V| - |S| - 1)!$ is the number of possible orderings of nodes from $V \setminus (S \cup \{v\})$ after v.

Thus, computing the Attachment Delta $\Delta_v^{s,t}(G)$ gives us the number of assignments that satisfy $C_1 \wedge \cdots \wedge C_m$. Since the problem of computing the number of such assignments (i.e., #3-SAT) is #**P**-complete and computing the Attachment Delta is in #**P**, it proves that computing the Attachment Delta is also #**P**-complete. This concludes the proof of Theorem 10. □

In result, we get that computing the Attachment centrality itself is also #**P**-complete.

Corollary 13. *Computing the Attachment centrality is #**P**-complete.*

Proof. The value $n!A_v(G)$ can be considered as the number of accepting paths of nondeterministic Turing machine, so the problem is in #**P**. Now, from Lemma 4, we know that $\Delta_v^{s,t}(G) = A_v(G) - A_v(G + \{s,t\})$ for every $v \in V \setminus \{s,t\}$. Thus, computing the Attachment centrality would allow us to compute the Attachment Delta. Since based on Theorem 10 computing the Attachment Delta is #**P**-complete we get that computing the Attachment centrality is also #**P**-complete. □

Our results show that computing the Attachment centrality is challenging and, in general, requires an exponential number of steps (unless **P** = **NP**). In the next sections, we will present our algorithms that try to deal with the exponential time complexity for chordal graphs.

6.2. Algorithm for chordal graphs

In this section, we present a polynomial time algorithm for chordal graphs. To this end, we start with some additional terminology.

The *elimination ordering* $\pi = (v_1, v_2, \dots, v_n)$ is a permutation of nodes, i.e., a bijection $\pi : V \to \{1, \dots, n\}$. The *fill-in* caused by ordering π, denoted by E_π, is the set of edges defined as follows:

$$E_\pi = \{\{u, w\} \notin E : \pi(u) < \pi(w), u \sim_{G[S_u^\pi \cup \{u,w\}]} w.\}$$

Recall that S_v^π is the set of predecessors of v in π. To put it in words, the fill-in contains edges between nodes which are connected by a path that consists of nodes that are earlier in the permutation than both of them. Elimination ordering π is *perfect* if the fill-in E_π is empty and it is *minimal* if there is no ordering π' such that $E_{\pi'} \subsetneq E_\pi$. The graph that results from adding the edges in E_π to G is denoted by $G_\pi = (V, E \cup E_\pi)$ and is called a *fill-in* graph. See Fig. 7 for an illustration.

Our first algorithm will be dedicated to *chordal graphs*. A graph is *chordal* if it contains no induced cycle of length four or more. In other words, every cycle longer than three has a *chord*, i.e., an edge joining two non-consecutive nodes on a cycle. Importantly, Fulkerson and Gross [24] proved that a graph is chordal if and only if it has a perfect elimination ordering.

Let us analyze in more detail a perfect elimination ordering, $\pi = (v_1, \dots, v_n)$. Fix v_k and consider two of its neighbors, u, w, which are in π after v_k, i.e., $u, w \in N_G(v)$ such that $\pi(u), \pi(w) > k$. Since $p = (u, v_k, w)$ is a path in $G[S_u^\pi \cup \{u, w\}]$, we get that $\{u, w\} \in E$ or $\{u, w\} \in E_\pi$. We assumed that π is a perfect elimination ordering, so $E_\pi = \emptyset$ which implies $\{u, w\} \in E$. This shows that every two neighbors of node v_k in graph $G[\{v_k, \dots, v_n\}]$ are connected by an edge which means that node v_k is *simplicial* in graph $G[\{v_k, \dots, v_n\}]$—a node is called *simplicial* if its neighbors form a clique.

Building upon this analysis and our result about cut cliques from Corollary 9, we show that the Attachment centrality can be calculated in polynomial time in chordal graphs.

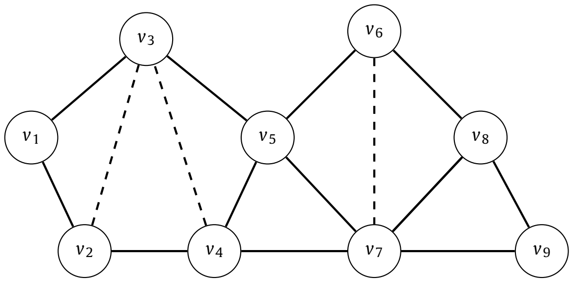

Fig. 7. An elimination ordering $\pi = (v_1, v_2, \dots, v_9)$ is a minimal elimination ordering of a graph $G = (V, E)$. Dashed lines represent the fill-in E_π caused by ordering π. In result, $G_\pi = (V, E \cup E_\pi)$ is a chordal graph and π is its perfect elimination.

Theorem 14. *The Attachment centrality can be calculated in chordal graphs in polynomial time.*

Proof. We begin by considering simplicial nodes. Fix $v \in V$ and let $C = N_G(v)$ be the set of neighbors of v that forms a clique. We will prove that the Attachment centrality in G can be easily determined based on the Attachment centrality in graph $G[V \setminus \{v\}]$. Specifically:

$$A_u(G) = \begin{cases} \frac{2|C|}{|C|+1} & \text{if } u = v, \\ A_u(G[V \setminus \{v\}]) + \frac{2}{|C|(|C|+1)} & \text{if } u \in C, \\ A_u(G[V \setminus \{v\}]) & \text{otherwise.} \end{cases} \tag{32}$$

To this end, observe that $C \cup \{v\}$ is a clique. From the fact that all nodes in graph $G[C \cup \{v\}]$ are symmetrical we get that:

$$A_v(G[C \cup \{v\}]) = \frac{2|C|}{|C|+1} = A_u(G[C \cup \{v\}]) \text{ for every } u \in C.$$

Analogously, $A_u(G[C]) = \frac{2(|C|-1)}{|C|}$ and $A_u(G[C \cup \{v\}]) - A_u(G[C]) = \frac{2}{|C|(|C|+1)}$ for every $u \in C$. Thus, if $V = C \cup \{v\}$, then (32) holds.

Assume otherwise. In such a case, we get that C is a cut clique and, since all neighbors of v are in C, we also get that $\{v\} \in K(G[V \setminus C])$. Now assume that $K(G[V \setminus C]) = \{\{v\}, S_2, \dots, S_k\}$ for some $k > 1$. From Theorem 8 we have: $A_v(G) = A_v(G[C \cup \{v\}]) = \frac{2|C|}{|C|+1}$. Now, consider $u \in C$. We will use Theorem 8 twice. For C in G we get:

$$A_u(G) = \frac{2|C|}{|C|+1} + \sum_{i \in \{2,\dots,k\}} A_u(G[S_i \cup C]) - \frac{2(|C|-1)}{|C|}(k-1), \tag{33}$$

and for C in $G[V \setminus \{v\}]$ we get:

$$A_u(G[V \setminus \{v\}]) = \sum_{i \in \{2,\dots,k\}} A_u(G[S_i \cup C]) - \frac{2(|C|-1)}{|C|}(k-2). \tag{34}$$

Combining (33) and (34) we get that $A_u(G) = A_u(G[V \setminus \{v\}]) + \frac{2}{|C|(|C|+1)}$. Now, fix $w \in V \setminus (C \cup \{v\})$ and assume $w \in S_j$ for some $j \in \{2, \dots, k\}$. Again, by considering cut clique C in graphs G and $G[V \setminus \{v\}]$ from Theorem 8 we get:

$$A_w(G) = A_w(G[S_j \cup C]), \text{ and } A_w(G[V \setminus \{v\}]) = A_w(G[S_j \cup C]).$$

Thus, $A_w(G) = A_w(G[V \setminus \{v\}])$. This concludes the proof of (32).

Now, it is known that a perfect elimination ordering can be found in polynomial time [61]—we will discuss one algorithm for this purpose after the proof. The Attachment centrality in graph $G[\{v_n\}]$ is trivial. Now, assume we know the Attachment centrality in graph $G[\{v_{k+1}, \dots, v_n\}]$. We already argued that in a perfect elimination ordering node v_k is simplicial in graph $G[\{v_k, \dots, v_n\}]$. So from (32) we get that the Attachment centrality of every node in $G[\{v_k, \dots, v_n\}]$ can be calculated in polynomial time based on the Attachment centrality in graph $G[\{v_{k+1}, \dots, v_n\}]$. After $n-1$ such steps we will get the Attachment centrality in graph $G[\{v_1, \dots, v_n\}] = G$. This concludes the proof of Theorem 14. □

Let us discuss how to find a perfect elimination ordering. To this end, we use the *maximum cardinality search* (MCS), proposed by Tarjan and Yannakakis [61]. In this method, we create a perfect elimination ordering in a reverse order, by iterating backwards from n to 1. In the first step, we choose an arbitrary node and put it at the end of the ordering. Then, in each step, we add to the ordering a node with the largest number of neighbors with assigned positions. From [61,

Algorithm 1: Computing the Attachment centrality in chordal graphs

Input: Chordal graph $G = (V, E)$
Output: Attachment centrality $A_v(G)$ for every $v \in V$
1 **foreach** $v \in V$ **do**
2 $\quad A_v \leftarrow 0$; $N_v \leftarrow \emptyset$; $\pi_v \leftarrow 0$;
`// maximum cardinality search`
3 **for** $k \leftarrow |V|$ **downto** 1 **do**
4 $\quad v \leftarrow$ node with $\pi_v = 0$ and the largest $|N_v|$;
5 $\quad \pi_v \leftarrow k$;
6 $\quad$ **foreach** $u \in V : (\pi_u = 0)$ **do**
7 $\quad\quad$ **if** $\{v, u\} \in E$ **then**
8 $\quad\quad\quad N_u \leftarrow N_u \cup \{v\}$;
9 $\quad$ **foreach** $u \in N_v$ **do**
10 $\quad\quad A_u \leftarrow A_u + 2/(|N_v| \cdot (|N_v| + 1))$;
11 $\quad A_v \leftarrow 2 \cdot |N_v|/(|N_v| + 1)$;
`// π is a perfect elimination ordering`
12 **return** A_v for every $v \in V$;

Theorem 2] we know that the resulting ordering is a perfect elimination ordering. Thus, we know that all these neighbors form a clique and we can update the Attachment centrality in accordance with (32) from Theorem 14.

Algorithm 1 presents the pseudocode of our algorithm. Lines 1 and 2 initialize all variables. Here, the Attachment centrality of every node v is A_v (initialized to 0 at the beginning), N_v is the set of neighbors with assigned position in the ordering (empty at the beginning), and π_v is its position in the perfect elimination ordering, i.e., $\pi(v)$, or 0 if its not yet assigned (0 at the beginning). Lines 3–11 contain the main loop. In line 4, the node v with the largest set N_v is selected and assigned to position k in the ordering π. After that, in lines 6–8, for every neighbor u of v without assigned position, the set N_u is updated. Lines 9–11 update the Attachment centrality. In result, at the end of each iteration of the main loop, A_u for every $u \in V$ stores the Attachment centrality of a graph that consists of all nodes with assigned position, i.e., nodes at the last $|V| - k + 1$ positions in a perfect elimination ordering π. Finally, line 12 returns the Attachment centrality.

The time complexity of Algorithm 1 is $O(|V| + |E|)$. To see why this is the case, observe that lines 6–8 can be performed one time for each edge when the first node incident to this edge has the position assigned. Furthermore, lines 9–10 are also performed one time for each edge when the second node incident to this edge has the position assigned. Therefore, the time to compute lines 6–10 is $O(|E|)$ and to compute the other lines is $O(|V|)$. The space complexity is also $O(|V| + |E|)$, because the sum of sizes of sets N_v for every v is equal to the number of edges in the graph.

Example 3. Consider Algorithm 1 applied to graph G_π from Fig. 7 (the graph defined by both solid and dashed lines). Consider the first iteration of the main loop. At the beginning (line 4), all sets N_u are empty and any node can be selected; assume that node v_9 is chosen. Then, in lines 6–8 node v_9 is added to sets N_u for its neighbors—v_7 and v_8. In result, after the first iteration, we have: $N_{v_7} = N_{v_8} = \{v_9\}$ and $N_u = \emptyset$ for every other u. Since $N_{v_9} = \emptyset$, in line 9–11 variables A_v does not change. In the second iteration, either v_7 or v_8 is chosen in line 4; assume node v_8 is chosen. Both v_6 and v_7—neighbors of v_8 without assigned number—have v_8 added to sets N_u in lines 6–8. Since $N_{v_8} = \{v_9\}$, in lines 9-11 we assign $A_{v_9} = 2/(1 \cdot 2) = 1$ and $A_{v_8} = 2 \cdot 1/2 = 1$. These values are the Attachment centralities in graph $G[\{v_8, v_9\}]$. The algorithm continues and systematically finds the Attachment centrality in graphs $G[\{v_7, v_8, v_9\}]$, ..., $G[\{v_2, \dots, v_9\}]$ and finally in $G[\{v_1, \dots, v_9\}] = G$. We conclude Example 3 by outlining all the updates to variables A_v performed by Algorithm 1:

	A_{v_1}	A_{v_2}	A_{v_3}	A_{v_4}	A_{v_5}	A_{v_6}	A_{v_7}	A_{v_8}	A_{v_9}
k=8, lines 10-11								1.00	1.00
k=7, lines 10-11							1.33	0.33	0.33
k=6, lines 10-11						1.33	0.33	0.33	
k=5, lines 10-11					1.33	0.33	0.33		
k=4, lines 10-11				1.33	0.33		0.33		
k=3, lines 10-11			1.33	0.33	0.33				
k=2, lines 10-11		1.33	0.33	0.33					
k=1, lines 10-11	1.33	0.33	0.33						
	1.33	1.67	2.00	2.00	2.00	1.67	2.33	1.66	1.33

6.3. Algorithm for general graphs

In this section, we present our general-purpose algorithm inspired by the algorithm for chordal graphs from the previous subsection. Recall that the algorithm for chordal graphs was based on the observation that if we consider nodes in a chordal graph in increasing order with respect to a perfect elimination ordering, then neighbors of every node form a clique in the subgraph induced by the remaining nodes. In our general algorithm we will also look for cut cliques and apply Theorem 8; by doing so, we will limit the size of a graph for which an extensive calculation of the Myerson value is required. To decompose the graph using cut cliques we will use the algorithm proposed by Tarjan [60]. This algorithm is based on a minimal elimination ordering of a graph. We begin by discussing how to find such an ordering.

To find a minimal elimination ordering we use an extension of the maximum cardinality search proposed by Berry et al. [6], called MCS-M. As in a maximal cardinality search, we create an ordering in a reverse order. In the first step, we choose an arbitrary node and put it at the end of the ordering. Then, in each step, we add to the ordering a node with the largest number of nodes with assigned numbers that are connected to it directly (i.e., its neighbors) or indirectly through nodes earlier in the ordering. More precisely, a node u without assigned number is considered to be indirectly connected to node v if at the time v has a number assigned there exists a path from u to v such that every node in the middle of this path, w, has no number assigned and is connected to a smaller number of nodes added earlier than u. Thus, the difference between MCS-M and maximal cardinality search lies in considering also indirectly connected nodes. In fact, edges between indirectly connected nodes constitute the fill-in caused by the constructed ordering.

Let us go back to the description of the algorithm for cut clique decomposition by Tarjan [60]. First, in addition to finding a minimal elimination ordering π, the algorithm for each node v computes set N_v, defined as follows:

$$N_v = \{u \in V : \pi(u) > \pi(v), \{v, u\} \in E \cup E_\pi\}.$$

These sets can be found using MCS-M along with a minimal elimination ordering. Now, the algorithm considers nodes in the increasing order with respect to π and for each checks whether the set N_v is a cut clique. If it is, then node v and its component in the graph without cut clique N_v are removed from the graph and the procedure continues. Importantly, this procedure ensures that there exists no cut clique in the removed part of the graph, i.e., any found part cannot be further decomposed.

Algorithm 2: Computing the Attachment centrality

Input: Graph $G = (V, E)$
Output: Attachment centrality $A_v(G)$ for every $v \in V$

1 **foreach** $v \in V$ **do**
2     $A_v \leftarrow 0$; $N_v \leftarrow \emptyset$; $\pi_v \leftarrow 0$;

// MCS-M [6]
3 **for** $k \leftarrow |V|$ **downto** 1 **do**
4     $v \leftarrow$ node with $\pi_v = 0$ and the largest $|N_v|$;
5     $\pi_v \leftarrow k$;
6   **foreach** $u \in V : (\pi_u = 0)$ **do**
7   **if** $(v \sim u)$ *in* $G[\{w \in V : (\pi_w = 0) \wedge (|N_w| < |N_u|)\} \cup \{v, u\}]$ **then**
8             $N_u \leftarrow N_u \cup \{v\}$;

// π is a minimal elimination ordering
9 $U \leftarrow V$;
10 **for** $k \leftarrow 1$ **to** $|V|$ **do**
11     $v \leftarrow$ node with $\pi_v = k$;
12   **if** $v \notin U$ **then continue**;
13     $S_v \leftarrow K_v(G[U \setminus N_v])$;
14   **if** (N_v *is a clique*) $\wedge$ $(S_v \neq U \setminus N_v)$ **then**
15   calculate $MV(f, G[S_v \cup N_v])$ for $f(C) = 2(|C| - 1)$;
16   **foreach** $u \in S_v \cup N_v$ **do**
17             $A_u \leftarrow A_u + MV_u(f, G[S_v \cup N_v])$;
18   **if** $u \in N_v$ **then** $A_u \leftarrow A_u - 2 \cdot (|N_v| - 1)/|N_v|$;
19         $U \leftarrow U \setminus S_v$;

20 calculate $MV(f, G[U])$ for $f(C) = 2(|C| - 1)$;
21 **foreach** $u \in U$ **do**
22     $A_u \leftarrow A_u + MV_u(f, G[U])$;

23 **return** A_v *for every* $v \in V$;

The pseudocode can be found in Algorithm 2. Here, lines 1 and 2 initialize variables, as in Algorithm 1. Lines 3–8 correspond to the MCS-M algorithm and find a minimal elimination ordering as well as the sets N_v. Here, whenever a node, v, has a position in the ordering assigned, it is added to the set N_u for every unassigned node for which there exists a path $p = (v, w_1, \dots, w_k, u)$ such that $\pi_{w_i} = 0$ and $|N_{w_i}| < |N_u|$ for every $w_i \in p$. Equivalently, node u is such that v and u are connected in a subgraph induced by v, u and nodes defined as follows: $\{w \in V : (\pi_w = 0) \wedge (|N_w| < |N_u|)\}$—these are the nodes that do not have a number assigned and have smaller sets N_w than node N_u. Note that this is the only line that differs between lines 1–8 in Algorithms 1 and 2 (i.e., between MCS and MCS-M).

Furthermore, lines 9–22 find a cut clique decomposition and compute the Attachment centrality. We model removing nodes from the graph with variable U that contains all nodes which are still in the graph (at the beginning initialized to

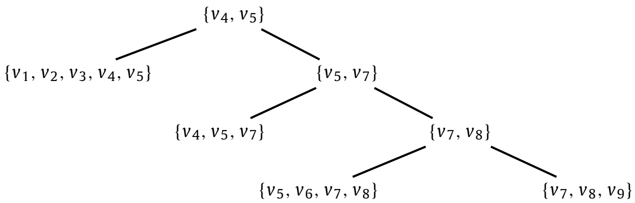

Fig. 8. A cut clique decomposition of a graph from Fig. 7.

set V in line 9). In each iteration of the loop (lines 10–19) node v from the k-th position in the ordering is considered (line 11). If it was already removed from the graph we move to the next iteration (line 12). Set S_v is the connected component in the graph with nodes from N_v removed from the graph. If nodes from N_v form a clique and S_v does not contain all nodes in the graph, then N_v is a cut clique in graph $G[U]$. In this case, from Theorem 8 we know that:

- For every node in S_v, the Attachment centrality of that node in graph G is equal to its Attachment centrality in graph $G[S_v \cup N_v]$;
- For every node in N_v, the Attachment centrality of that node in graph G is equal to its Attachment centrality in graph $G[S_v \cup N_v]$ plus its Attachment centrality in graph $G[U \setminus S_v]$ minus $2(|N_v| - 1)/|N_v|$;
- For every node in $U \setminus (N_v \cup S_v)$, the Attachment centrality of that node in graph G is equal to its Attachment centrality in graph $G[U \setminus S_v]$.

Thus, we calculate the Attachment centrality of graph $G[S_v \cup N_v]$ using the algorithm for the Myerson value of the game $f(C) = 2(|C| - 1)$ proposed by Skibski et al. [55] (line 15), add the correct values to nodes from $S_v \cup N_v$ (lines 16–18) and remove the nodes in S_v from the graph (line 19). After all iterations of the loop, we are left with the remaining part of the graph that cannot be further decompose. Thus, we again compute the Attachment centrality using the algorithm for the Myerson value of the game $f(C) = 2(|C| - 1)$.

The running time of our algorithm highly depends on the topology of the graph. In the worst case, none of the decomposition steps in line 14 will succeed and our algorithm will calculate the Myerson value for the entire graph in line 20. As shown by Skibski et al. [55], this will require enumerating all induced connected subgraphs of the graph, the number of which is usually exponential.[4] The space complexity is also linear in respect to the number of induced connected subgraphs.

Example 4. Consider Algorithm 2 applied to graph G from Fig. 7. First, consider finding the minimal ordering (lines 3–8). After two iterations of the loop we have nodes v_8 and v_9 at the end of the ordering (i.e., $\pi_{v_8} = 8$, $\pi_{v_9} = 9$) and $N_{v_6} = \{v_8\}$, $N_{v_7} = \{v_8, v_9\}$ and $N_u = \emptyset$ for other nodes without assigned position. In the next iteration, node v_7 has a position assigned (i.e., $\pi_{v_7} = 7$ in line 5) and it is added to sets N_{v_3}, N_{v_4} and N_{v_6}. This is because nodes v_3, v_4 and v_6 are all unnumbered, the former two are directly connected to v_7 and the last one, v_6, is connected undirectly with a path (v_6, v_4, v_7) on which $|N_{v_4}| < |N_{v_6}|$. As a result, our algorithm finds the ordering $\pi = (v_1, \dots, v_9)$ with the fill-in $E_\pi = \{\{v_2, v_3\}, \{v_3, v_5\}, \{v_6, v_7\}\}$ (represented by dashed edges on Fig. 7).

Now, consider lines 9–22 that find the Attachment centrality by performing the cut clique decomposition of a graph. In the 1st iteration, for v_1, we have $N_{v_1} = \{v_2, v_3\}$. Since N_{v_1} is not a clique, the set U and values A_v remain unchanged. In the 2nd iteration, for v_2, we have $N_{v_2} = \{v_3, v_4\}$. Again, N_{v_2} is not a clique. In the 3rd iteration, for v_3, we have $N_{v_3} = \{v_4, v_5\}$. This time, N_{v_3} is a clique. Moreover, removing N_{v_3} from the graph results in more than one connected component: $S_v = \{v_1, v_2, v_3\}$ and $S_v \neq U \setminus N_{v_3}$; thus, the if-condition from line 14 is satisfied and the first cut clique is found. The algorithm calculates the Myerson value for graph $G[\{v_1, \dots, v_5\}]$ (which—from the symmetry of all nodes—is equal to 1.6 for every node), updates the Attachment centrality, removes nodes $\{v_1, v_2, v_3\}$ from the set U (i.e., from the consideration) and moves forward with the procedure. The algorithm succeeds in finding a cut clique in the 4th and 6th iterations (for $N_{v_4} = \{v_5, v_7\}$ and $N_{v_6} = \{v_7, v_8\}$). All updates of variables A_v performed by Algorithm 2 are as follows:

	A_{v_1}	A_{v_2}	A_{v_3}	A_{v_4}	A_{v_5}	A_{v_6}	A_{v_7}	A_{v_8}	A_{v_9}
k=3, line 17	1.60	1.60	1.60	1.60	1.60				
k=3, line 18				−1.00	−1.00				
k=4, line 17				1.33	1.33		1.33		
k=4, line 18					−1.00		−1.00		
k=6, line 17					1.50	1.50	1.50	1.50	
k=6, line 18							−1.00	−1.00	
k=9, line 22							1.33	1.33	1.33
	1.60	1.60	1.60	1.93	2.43	1.50	2.16	1.83	1.33

[4] See the work by Elkind [19] for the analysis of the case where the number of induced connected subgraphs in a graph is polynomial.

Table 1
The ten highest ranked nodes in the Madrid network, according to different centrality indices. The numbers in columns 2–5 correspond to the nodes from Fig. 9.

Rank	Attachment		Betweenness	Closeness	Degree
1st	7	(Imad Eddin Barakat)	63	1	1
2nd	63	(Semaan Gaby Eid)	1	3	3
3rd	19	(Abderrahim Zbakh)	3	41	7
4th	61	(Mohamed El Egipcio)	40	7	11
5th	24	(Naima Oulad Akcha)	7	31	41
6th	11	(Amer Azizi)	31	40	18
7th	6	(Mohamed Chedadi)	24	24	24
8th	31	(Jamal Ahmidan)	19	11	19
9th	21	(Jose Emilio Suaez)	61	19	31
10th	1	(Jamal Zougam)	25	30	61

The resulting cut clique decomposition is presented in Fig. 8. Consequently, instead of performing the extensive Myerson value calculation for a graph of size 9, our algorithm only performs this calculation for 4 smaller subgraphs of sizes ranging from 3 to 5: $\{v_1, v_2, v_3, v_4, v_5\}$, $\{v_4, v_5, v_7\}$, $\{v_5, v_6, v_7, v_8\}$ and $\{v_7, v_8, v_9\}$. This concludes Example 4.

6.4. A case study: 2004 Madrid train bombings network

As a sample application, we focus on the identification of key terrorists in covert organizations. To this end, we analyze the terrorist network responsible for the 2004 attacks on Madrid trains. The reasons behind our choice of the application and the network are twofold. Firstly, it has been recently argued that, in order to identify key members in a terrorist networks, one should focus on nodes that enable communication within that network [37,39]. Secondly, the Madrid network is relatively big and, thus far, has never been analyzed with a centrality index of this kind.

The *Madrid network* consists of 70 nodes and 243 edges. The size of the network makes it impractical to compute the existing centrality indices focused on nodes that enable communication. In more detail, the computation involves enumerating all induced connected subgraphs of the network. Unfortunately, even the state-of-the-art algorithm for this purpose [55] takes over 100 seconds to compute the Myerson value for a sparse network with only 36 nodes. Furthermore, the running time grows exponentially with the size of the network; every additional node nearly doubles it. To address this challenge, we use Algorithm 2 introduced in the previous subsection to narrow down the set of nodes for which the Myerson value has to be calculated.

The original Madrid network [48] contains 6 isolated nodes. From Normalization, we know that the Attachment centrality of each of these nodes is 0. We also know from Locality that those 6 nodes can be removed without affecting the Attachment centrality of others. Furthermore, we observe that the Madrid network contains 8 leaf nodes. From Corollary 9, we know that every such node has an Attachment centrality of 1, and can easily be removed from the network (since the corollary specifies the impact of this removal). Moreover, from Theorem 14 we know that every node whose set of neighbors, K, forms a clique has an Attachment centrality of $\frac{2|K|}{|K|+1}$, and can easily be removed from the network (since the theorem specifies the impact of this removal on the Attachment centrality of other nodes). Note that removing nodes results in a chain reaction, meaning that the above rules can be applied repeatedly (e.g., by removing a leaf, some other node might become a leaf, which can then be removed, and so on).

Algorithm 2 carries out the above process systematically, by finding the cut clique decomposition of a graph. As we have already mentioned, the running time of our algorithm depends on the topology of the graph. In the case of the Madrid network, the largest subgraph for which the Myerson value had to be calculated in line (11) was a subgraph consisting of 26 nodes. The running time on the entire network was 15.01 seconds on a standard desktop PC.

The results of our analysis are summarized in Table 1. As can be seen, the Attachment centrality significantly differs from the standard centrality indices. For instance, let us consider the two nodes with the highest number of edges—nodes 1 (Jamal Zougam) and 3 (Mohamed Chaoui)—which are positioned top by the Degree and Closeness centralities. Interestingly, the Betweenness centrality also gives them very high (the second and the third) positions while the Attachment centrality ranks node 1 as the tenth and node 3 even lower. Such a substantial difference between the Betweenness and Attachment centralities is quite surprising, given that nodes that connect different parts of the network are more likely to belong to shortest paths than other nodes. However, nodes 1 and 3 have so many neighbors that they are very often parts of the shortest paths in the network, more often than nodes important from the perspective of enabling communication.

Our analysis revealed a previously unknown aspect of the Madrid network. In particular, we identified several overlapping parts of the network that are almost fully connected, i.e., each part is almost a clique, except for very few missing edges. This new insight confirms the existing belief, that *terrorist networks consist of rather sparsely-connected, highly-dense parts* [63].

7. Conclusions

In this paper, we proposed the *Attachment centrality*—a new centrality index that belongs to the recently-proposed class of indices focusing on identifying nodes that *enable communication altogether*, rather than identifying nodes that contribute

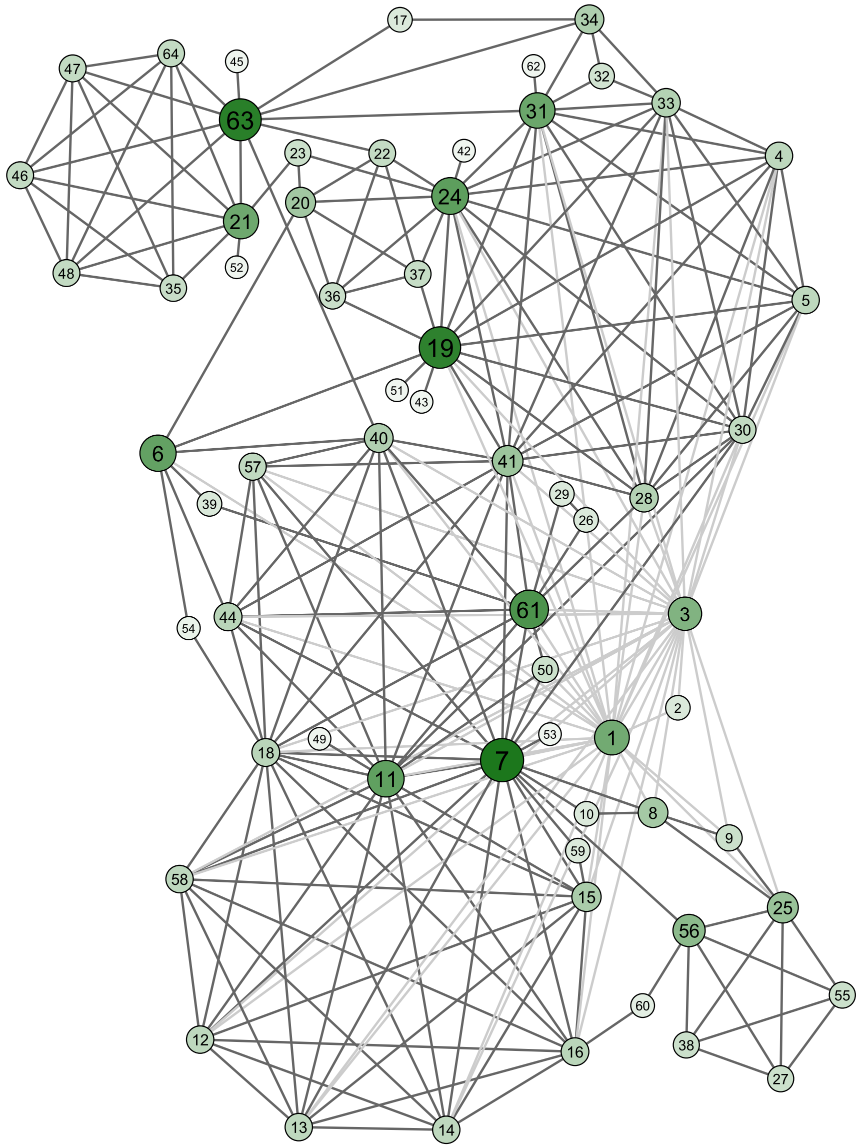

Fig. 9. *Madrid network*. The node size reflects the Attachment centrality (the larger the node the greater its centrality). To highlight the differences even further, the node color is set to reflect the node size (the larger the node the darker the color).

towards the *quality or speed of the communication* within the network, as it is typically the case with the classical centrality indices such as Degree, Closeness, Betweenness, and Eigenvector. In other words, when using the Attachment centrality, a node is evaluated based on *whether or not the network (or part of it) becomes disconnected if this node is removed*.

The Attachment centrality, is equivalent to the Myerson value of the graph-restricted game defined by the characteristic function $f^*(S) = 2(|S| - 1)$. The new index has a simple and elegant interpretation: If we were to remove nodes from the network one by one in a random order, then the Attachment centrality of a node is the expected number of components created immediately after the removal of this node from the network, multiplied by 2 for the normalization purposes.

We proved that the Attachment centrality is the only centrality measure that satisfies a number of intuitive axioms. While there were a number of attempts in the literature to provide theoretical foundations to various centrality indices [51, 33,14], our analysis is the first in the literature that proposes an axiomatization of an index focusing on the aforementioned notion of connectivity.

Building upon our theoretical analysis we show that, while computing the Attachment centrality is #P-complete, it has certain computational properties that are more attractive than the Myerson value for an arbitrary game. In particular, it can be computed in chordal graphs in polynomial time.

There are various interesting directions in which our work can be extended. In particular, while most work on centrality measures considers unweighted graphs, many real-life networks are weighted. Sosnowska and Skibski [58] extended the

Attachment centrality to weighted graphs, but the computational complexity of this extension has not been studied to date. Yet another direction is the axiomatic analysis of other game-theoretic centralities. In this context, Skibski et al. [53] recently proved that, while every centrality measure can be obtained using the game-theoretic approach, natural subclasses of game-theoretic centralities can be characterized by Fairness and its modifications. Finally, we plan to consider approximation algorithms for the Attachment centrality.

Acknowledgements

This article is an extended version of a paper originally presented at AAMAS-16 [56] with a number of new technical results. Specifically, all the results regarding the Attachment Delta are new (Lemmas 4–5 and Theorem 7). Furthermore, most of the complexity results are new: all hardness results (Theorem 10, Lemmas 11–12, Theorem 13), as well as the complexity result for chordal graphs (Theorem 14). Finally, we significantly extended the related work section (Section 2), added several new examples and illustrations.

Oskar Skibski was supported by the Foundation for Polish Science within the Homing programme (Project title: "Centrality Measures: from Theory to Applications"). Oskar Skibski and Makoto Yokoo were supported by JSPS KAKENHI Grant (24220003). Tomasz Michalak was supported by the Polish National Science Centre grant 2015/19/D/ST6/03113 and by the European Research Council under Advanced Grant 291528 ("RACE"). Makoto Yokoo was supported by JSPS KAKENHI Grant (17H00761). The early version of this work was also supported by the Polish National Science Centre grant 2013/09/D/ST6/03920.

References

[1] E. Algaba, J.M. Bilbao, J.R. Fernández, N. Jiménez, J.J. López, Algorithms for computing the Myerson value by dividends, in: Discrete Mathematics Research Progress, chap. 6, Nova Science Publishers, 2007, pp. 126–137.
[2] A. Altman, M. Tennenholtz, Ranking systems: the PageRank axioms, in: Proceedings of the 6th ACM Conference on Electronic Commerce (ACM-EC), 2005, pp. 1–8.
[3] R. Amer, J.M. Giménez, A connectivity game for graphs, Math. Methods Oper. Res. 60 (3) (2004) 453–470.
[4] J. Belau, Consequences of connection failure - centrality and the importance for cohesion, in: Public Choice Society, 2014.
[5] S. Benati, R. Rizzi, C. Tovey, The complexity of power indexes with graph restricted coalitions, Math. Soc. Sci. 76 (2015) 53–63.
[6] A. Berry, J.R. Blair, P. Heggernes, B.W. Peyton, Maximum cardinality search for computing minimal triangulations of graphs, Algorithmica 39 (4) (2004) 287–298.
[7] A.P. Bianzino, C. Chaudet, D. Rossi, J.-L. Rougier, S. Moretti, The green-game: striking a balance between QoS and energy saving, in: Proceeding of the 23rd International Teletraffic Congress (ITC), 2011, pp. 262–269.
[8] D. Bienstock, On the complexity of testing for odd holes and induced odd paths, Discrete Math. 90 (1) (1991) 85–92.
[9] J.-M. Bilbao, Values and potential of games with cooperation structure, Int. J. Game Theory 27 (1) (1998) 131–145.
[10] F. Bistaffa, A. Farinelli, J. Cerquides, J. Rodríguez-Aguilar, S.D. Ramchurn, Algorithms for graph-constrained coalition formation in the real world, ACM Trans. Intell. Syst. Technol. 8 (4) (2017) 60.
[11] F. Bloch, M.O. Jackson, P. Tebaldi, Centrality Measures in Networks, 2016, available at SSRN 2749124.
[12] P. Boldi, S. Vigna, Axioms for centrality, Internet Math. 10 (3–4) (2014) 222–262.
[13] P. Bonacich, Factoring and weighting approaches to status scores and clique identification, J. Math. Sociol. 2 (1) (1972) 113–120.
[14] S.P. Borgatti, M.G. Everett, A graph-theoretic perspective on centrality, Soc. Netw. 28 (4) (2006) 466–484.
[15] U. Brandes, T. Erlebach, Network Analysis: Methodological Foundations, Springer-Verlag, ISBN 3540249796, 2005.
[16] M. Dehmer, S.C. Basak, Statistical and Machine Learning Approaches for Network Analysis, vol. 707, John Wiley & Sons, 2012.
[17] V. Dequiedt, Y. Zenou, Local and Consistent Centrality Measures in Networks, cEPR Discussion Paper No. DP10031, 2014.
[18] Z. Dezső, A.-L. Barabási, Halting viruses in scale-free networks, Phys. Rev. E 65 (2002) 055103.
[19] E. Elkind, Coalitional games on sparse social networks, in: Proceedings of the 10th Conference on Web and Internet Economics (WINE), Springer, 2014, pp. 308–321.
[20] J.R. Fernández, E. Algaba, J.M. Bilbao, A. Jiménez, N. Jiménez, J.J. López, Generating functions for computing the Myerson value, Ann. Oper. Res. 109 (1–4) (2002) 143–158.
[21] L.C. Freeman, A set of measures of centrality based on betweenness, Sociometry (1977) 35–41.
[22] L.C. Freeman, Centrality in social networks: conceptual clarification, Soc. Netw. 1 (3) (1979) 215–239.
[23] L.C. Freeman, S.P. Borgatti, D.R. White, Centrality in valued graphs: a measure of betweenness based on network flow, Soc. Netw. 13 (2) (1991) 141–154.
[24] D. Fulkerson, O. Gross, Incidence matrices and interval graphs, Pac. J. Math. 15 (3) (1965) 835–855.
[25] M. Garg, Axiomatic foundations of centrality in networks, available at SSRN 1372441, 2009.
[26] M. Gerlach, T.P. Peixoto, E.G. Altmann, A network approach to topic models, Sci. Adv. 4 (7) (2018) eaaq1360.
[27] D. Gómez, E. González-Arangüena, C. Manuel, G. Owen, M. del Pozo, J. Tejada, Centrality and power in social networks: a game theoretic approach, Math. Soc. Sci. 46 (1) (2003) 27–54.
[28] D. Gómez, E. González-Arangüena, C. Manuel, G. Owen, M. del Pozo, J. Tejada, Splitting graphs when calculating Myerson value for pure overhead games, Math. Methods Oper. Res. 59 (3) (2004) 479–489.
[29] N. Himaja, G. Murali, Social network analysis for efficient cloud data access, in: Proceedings of the International Conference on Communication and Electronics Systems (ICCES), IEEE, 2016, pp. 1–4.
[30] R. Kaewpuang, D. Niyato, P.-S. Tan, P. Wang, Cooperative management in full-truckload and less-than-truckload vehicle system, IEEE Trans. Veh. Technol. 66 (7) (2017) 5707–5722.
[31] A. Keinan, C.C. Hilgetag, I. Meilijson, E. Ruppin, Casual localization of neural function: the Shapley value method, Neurocomputing 58–60 (2004) 215–222.
[32] M. Kitti, Axioms for centrality scoring with principal eigenvectors, Soc. Choice Welf. 46 (3) (2016) 639–653.
[33] D. Koschützki, K.A. Lehmann, L. Peeters, S. Richter, D. Tenfelde-Podehl, O. Zlotowski, Centrality indices, in: Network Analysis, in: Lecture Notes in Computer Science, vol. 3418, Springer, 2005, pp. 16–61.
[34] D. Koschützki, K.A. Lehmann, D. Tenfelde-Podehl, O. Zlotowski, Advanced centrality concepts, in: Network Analysis, in: Lecture Notes in Computer Science, vol. 3418, Springer, 2005, pp. 83–111.

[35] R. Kötter, A.T. Reid, A. Krumnack, E. Wanke, O. Sporns, Shapley ratings in brain networks, Front. Neuroinformatics 1 (2007) 2.
[36] C.A. Lareau, B.C. White, A.L. Oberg, B.A. McKinney, Differential co-expression network centrality and machine learning feature selection for identifying susceptibility hubs in networks with scale-free structure, BioData Min. 8 (1) (2015) 5.
[37] R. Lindelauf, H. Hamers, B. Husslage, Cooperative game theoretic centrality analysis of terrorist networks: the cases of Jemaah Islamiyah and Al Qaeda, Eur. J. Oper. Res. 229 (1) (2013) 230–238.
[38] T.P. Michalak, K.V. Aadithya, P.L. Szczepański, B. Ravindran, N.R. Jennings, Efficient computation of the Shapley value for game-theoretic network centrality, J. Artif. Intell. Res. 46 (2013) 607–650.
[39] T.P. Michalak, T. Rahwan, P.L. Szczepański, O. Skibski, R. Narayanam, M. Wooldridge, N.R. Jennings, Computational analysis of connectivity games with applications to the investigation of terrorist networks, in: Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI), 2013, pp. 293–301.
[40] D.C. Mocanu, H.B. Ammar, D. Lowet, K. Driessens, A. Liotta, G. Weiss, K. Tuyls, Factored four way conditional restricted Boltzmann machines for activity recognition, Pattern Recognit. Lett. 66 (2015) 100–108.
[41] D.C. Mocanu, G. Exarchakos, H.B. Ammar, A. Liotta, Reduced reference image quality assessment via Boltzmann machines, in: Proceedings of the IFIP/IEEE International Symposium on Integrated Network Management (IM), IEEE, 2015, pp. 1278–1281.
[42] S. Moretti, V. Fragnelli, F. Patrone, S. Bonassi, Using coalitional games on biological networks to measure centrality and power of genes, Bioinformatics 26 (21) (2010) 2721–2730.
[43] R.B. Myerson, Graphs and cooperation in games, Math. Methods Oper. Res. 2 (3) (1977) 225–229.
[44] M. Newman, Networks: An Introduction, Oxford University Press, 2010.
[45] L. Page, S. Brin, R. Motwani, T. Winograd, The PageRank citation ranking: bringing order to the web, Stanford InfoLab.
[46] S. Pal, Y. Dong, B. Thapa, N.V. Chawla, A. Swami, R. Ramanathan, Deep learning for network analysis: problems, approaches and challenges, in: Military Communications Conference (MILCOM), IEEE, 2016, pp. 588–593.
[47] I. Palacios-Huerta, O. Volij, The measurement of intellectual influence, Econometrica 72 (3) (2004) 963–977.
[48] J.A. Rodríguez, The March 11th Terrorist Network: in Its Weakness Lies Its Strength, Ph.D. thesis, Universitat de Barcelona, 2005.
[49] W. Saad, Z. Han, M. Debbah, A. Hjørungnes, T. Başar, Coalitional game theory for communication networks, IEEE Signal Process. Mag. 26 (5) (2009) 77–97.
[50] J. Sabater, C. Sierra, Reputation and social network analysis in multi-agent systems, in: Proceedings of the 1st International Conference on Autonomous Agents and Multiagent Systems (AAMAS), ACM, 2002, pp. 475–482.
[51] G. Sabidussi, The centrality index of a graph, Psychometrika 31 (4) (1966) 581–603.
[52] L.S. Shapley, A value for n-person games, in: H. Kuhn, A. Tucker (Eds.), Contributions to the Theory of Games, vol. II, Princeton University Press, 1953, pp. 307–317.
[53] O. Skibski, T.P. Michalak, T. Rahwan, Axiomatic characterization of game-theoretic centrality, J. Artif. Intell. Res. 62 (2018) 33–68.
[54] O. Skibski, T.P. Michalak, Y. Sakurai, M. Yokoo, A pseudo-polynomial algorithm for computing power indices in graph-restricted weighted voting games, in: Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI), 2015, pp. 631–636.
[55] O. Skibski, T. Rahwan, T.P. Michalak, M. Wooldridge, Enumerating connected subgraphs and computing the Myerson and Shapley values in graph-restricted games, ACM Trans. Intell. Syst. Technol. 10 (2) (2019) 15:1–15:25.
[56] O. Skibski, T. Rahwan, T.P. Michalak, M. Yokoo, Attachment centrality: an axiomatic approach to connectivity in networks, in: Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), 2016, pp. 168–176.
[57] O. Skibski, J. Sosnowska, Axioms for distance-based centralities, in: Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI), 2018, pp. 1218–1225.
[58] J. Sosnowska, O. Skibski, Attachment centrality for weighted graphs, in: Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI), 2017, pp. 416–422.
[59] J. Sosnowska, O. Skibski, Path evaluation and centralities in weighted graphs - an axiomatic approach, in: Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI), 2018, pp. 3856–3862.
[60] R.E. Tarjan, Decomposition by clique separators, Discrete Math. 55 (2) (1985) 221–232.
[61] R.E. Tarjan, M. Yannakakis, Simple linear-time algorithms to test chordality of graphs, test acyclicity of hypergraphs, and selectively reduce acyclic hypergraphs, SIAM J. Comput. 13 (3) (1984) 566–579.
[62] M.K. Tarkowski, P.L. Szczepański, T.P. Michalak, P. Harrenstein, M. Wooldridge, Efficient computation of semivalues for game-theoretic network centrality, J. Artif. Intell. Res. 63 (2018) 145–189.
[63] M. Tsvetovat, K.M. Carley, Generation of Realistic Social Network Datasets for Testing of Analysis and Simulation Tools, Tech. Rep., DTIC Document, 2005.
[64] R. van den Brink, P. Borm, R. Hendrickx, G. Owen, Characterizations of the β-and the degree network power measure, Theory Decis. 64 (4) (2008) 519–536.
[65] R. van den Brink, R.P. Gilles, Measuring domination in directed networks, Soc. Netw. 22 (2) (2000) 141–157.
[66] T. Wąs, O. Skibski, An axiomatization of the eigenvector and Katz centralities, in: Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI), 2018, pp. 1258–1265.
[67] T. Wąs, O. Skibski, Axiomatization of the PageRank centrality, in: Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI), 2018, pp. 3898–3904.
[68] K. Zhou, T. Liu, L. Zhou, Industry 4.0: towards future industrial opportunities and challenges, in: Proceedings of the 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), IEEE, 2015, pp. 2147–2152.