Diffusion Delay Centrality: decelerating diffusion processes across networks

Valerio Leone Sciabolazza[1] and Luca Riccetti*[2]

[1]University of Naples Parthenope, via Gen. Parisi 20, 80132, Naples, Italy
[2]University of Macerata, via Crescimbeni 14, 62100 Macerata, Italy

Abstract

This paper presents a new measure (the Diffusion Delay Centrality - DDC) to identify agents who should be put into isolation to decelerate a diffusion process spreading throughout a network. We show that DDC assigns a high rank to agents acting as the gatekeepers of the fringe of the network. We also show that the ranking of nodes obtained from the DDC is predicted by the difference in the values of betweenness and eigenvector centrality of network agents. The findings presented might constitute a useful tool to reduce diffusion processes both for policy makers and for corporate managers in the organization of production.

Keywords: network, contagion, betweenness centrality, eigenvector centrality.

JEL classification codes: D85

Funding: Not applicable.

Conflicts of interest/Competing interests: Not applicable.

Availability of data and material: Not applicable.

Code availability: from the authors, upon request.

---

*Corresponding author.
E-mail address: luca.riccetti@unimc.it

1 Introduction

Social networks play a key role in several facets of our life. The contagion of diseases, the spreading of behavior, the transmission of ideas and information, the adoption of a technology or a product, and propagation of a financial crisis are all network-based phenomena. A crucial aspect when studying such issues is the identification of who holds a vital position in the network for enhancing or impeding a diffusion process. This recognition has sparked a growing literature in investigating the individual prominence of an agent as a channel of diffusion in terms of network centrality measures (see Jackson et al., 2017, for a recent review).[1]

The present paper introduces a new measure of centrality which assesses the importance of an agent in decelerating a diffusion process mediated through a network. We refer to this measure as to Diffusion Delay Centrality (DDC). In constructing this measure, we assume no prior knowledge about the origin of the diffusion process and the way it is transmitted. In this setting then, each area of the network is equally likely to be the source of contagion, and transmission can be caused even by a single contact between two nodes. Formally, this implies a simple deterministic contagion process, where diffusion is transmitted by a direct contact between two agents. Simple contagion processes are a well-documented phenomena from the empirical literature, and they govern the adoption of different technologies (Conley and Udry, 2010; Fogli and Veldkamp, 2021; Foster and Rosenzweig, 1995; Munshi, 2004), the circulation of a disease (Centola, 2013, 2018), the diffusion of many behaviors within online communities (Dow et al., 2013; Karampourniotis et al., 2015; State and Adamic, 2015), the spreading of traffic jams in urban networks (Saberi et al., 2020), and the transmission of a simple information (see e.g. Kim et al., 2015), among other things. Moreover, simple contagion can trigger and coexist with complex contagion processes (Min and San Miguel, 2018), where transmission is caused by the contact with multiple nodes (for a discussion see Guilbeault et al., 2018).

To the best of our knowledge, this is the first network metric which allows for investigation of how to slow down diffusion when the only information known is the structure of the network. On the contrary, existing network centrality measures are used to study agents as starting points of a diffusion

---

[1] A different strand of investigation instead has been devoted to understand how diffusion is facilitated by the characteristics of the population and the way in which people connect with each other. Seminal contributions on this issue were provided by the literature on epidemiology (Pastor-Satorras and Vespignani, 2000, 2001), game theory (Jackson and Yariv, 2005, 2007; Jackson and López-Pintado, 2013; López-Pintado, 2006, 2008, 2012), and physics (Newman, 2002).

process, and predict whether the diffusion starting from an agent will or will not remain confined within certain areas of the network, or the time needed to reach other parts of the network (see e.g. Ballester et al., 2006; Banerjee et al., 2014; Borgatti, 2005; Katz and Lazarsfeld, 1955; Kempe et al., 2003, 2005; Rogers, 1995). Other measures instead, require additional information about diffusion dynamics or about the infected status of the agents to identify how to decelerate the diffusion (Klemm et al., 2012; Liu et al., 2016; Lü et al., 2016; Piraveenan et al., 2013).

The analysis of network centrality measures has relevant policy implications. Indeed, by targeting most central agents, it is possible to take full advantage of the social interaction effects, and boost or react against the diffusion process. Empirical studies on simple contagion processes show the effectiveness of this approach when trying to sustain cooperative behavior and community social capital (Jackson, 2020), accelerate the participation to a microfinance program (Banerjee et al., 2013), and reduce the spread of juvenile delinquency behavior (Lee et al., 2020). This is also the case when dealing with complex contagion processes, when one wants to facilitate the sharing of the knowledge base to develop an emerging industrial sector (Choi et al., 2011), stimulate the adoption of a new insurance product (Cai et al., 2015), reduce the financial contagion (see for extensive reviews of this large literature Hüser, 2015, and Neveu, 2018, among others), or decelerate the diffusion of a severe acute respiratory syndrome (SARS) across the world (Colizza et al., 2006; Colizza and Vespignani, 2007).[2]

When studying simple contagion processes, network centrality measures adopt different criteria to rank the importance of an agent in facilitating a diffusion process (see among others Wasserman and Faust, 1994; Jackson, 2010). By looking at the paths of a network originating from, terminating at, or mediated by an agent, these measures identify better-connected individuals who can accelerate the diffusion and potentially determine a rapid change in the network. At the network-local level, degree centrality has been shown to be a useful predictor of diffusion processes (Christakis and Fowler, 2010). This records the maximum number of people who can be reached by direct contact from the agent. An extension of degree centrality is closeness (Bavelas, 1950), which is a weighted sum of the number of people who can be reached by a node by direct or indirect contact (Bloch et al., 2021), with weights given by the distance separating the node from each agent. This measure can be used to assess how rapidly an agent can propagate diffusion. At the network-global level, the most reliable measure of diffusion is provided

---

[2] See for further references and background Jackson and Yariv (2011) and Lamberson (2016).

by the family of eigenvector centralities (Jackson et al., 2017), which also represents an extension of degree centrality. Specifically, this set of metrics tracks in various forms the ability of the agent to act as a social multiplier of the diffusion process, by measuring how he or she can directly and indirectly reach high degree agents in different local areas of the network (Gould, 1967). Finally, another popular measure is betweenness, which proxies the ability of the agent in being a connector or an intermediary of diffusion, by measuring the agent's involvement in the shortest diffusion paths between every pair of agents in the network. Recent literature (Loepfe et al., 2013) found that diffusion can be slowed down by removing some of the nodes with the highest betweenness centrality.

With respect to extant measures used to investigate a simple contagion process, the DDC ranks the importance of an agent by assessing how this is able to alter the average speed of diffusion in the network by putting him or herself in isolation, causing a discontinuity in the chain of interactions existing in the network, and forcing the diffusion process to spread through different and potentially longer paths. In this setting then, diffusion spreads through direct contact, and the speed of contagion between two agents is equal to the minimum number of interactions separating one agent from the other: i.e. the shortest path length indirectly connecting a pair of agents in the network.

A crucial feature of the DDC is that diffusion is analyzed regardless of where the contagion process started. We discuss the theoretical details under a broad set of hypotheses. The insights provided by the DDC on diffusion are investigated by implementing a Monte Carlo experiment. Specifically, we analyze the situation in which diffusion can reach any other agent only throughout direct contact. We further assume that when a direct contact occurs between two agents, diffusion can flow in either direction, and hence all agents represent at the same time a potential starting point and a terminal point of diffusion. From a theoretical standpoint, this represents the case of a population embedded in an undirected and unweighted network, where diffusion is equally likely to reach all nodes without following any specific direction. However, we will also show that DDC can be equally used to examine directed and weighted networks, as well as single and multiple component networks. The Monte Carlo experiment is conducted by simulating a number of undirected and unweighted networks with different densities. Simulated networks are then used to compare how DDC ranks the relative importance of the agents in spreading diffusion with respect to most common measures of centrality. Specifically, we consider degree, betweenness, closeness, and eigenvector centrality, given their prominent relevance in the literature.

The use of DDC is well suited to investigate situations in which a simple diffusion process propagates from the periphery to the core of the network, and vice versa. In fact, DDC assigns a high rank to agents with two specific characteristics: first, they act as a major bridge for the diffusion process, meaning that they feature a high betweenness centrality; second, they are in contact with low-connected agents located at the periphery of the network: i.e. they feature a low eigenvector centrality. In other words, high DDC agents act as the gatekeepers of the fringe of the network. Once these are isolated, the spreading of diffusion from the periphery to the core of the network (or vice cersa) is prevented or at least decelerated. In line with this finding, we also show that the DDC of an agent is strongly predicted by the difference in the value of his or her betweenness and eigenvector centrality.

The contribution of this paper is twofold. First, we introduce a new metric, the DDC, which easily allows the identification of agents to be isolated in order to decelerate a simple diffusion process when this spreads from the periphery to the core of the network, or the other way around. Second, we show that, under a broad set of circumstances, DDC has a simple interpretation in terms of standard centrality measures, because it has a straightforward relationship with betweenness and eigenvector centrality: i.e., it is closely approximated by the difference between the two. This also implies that one can easily proxy the DDC ranking by simply taking the difference in the value of betweenness and eigenvector centrality.

The results presented are relevant and easily implementable by a policy maker interested in preventing the risks related to undesired diffusion processes, like the Covid-19 pandemic of 2020-21. For instance, the DDC can be used to identify those workers in a firm or a hospital who should be first put into isolation or equipped with specific medical devices (e.g. high performing respiratory protection masks) or vaccinated to prevent the diffusion of the contagion. Of course, in order to do that it is crucial to have knowledge about the structure of the network under analysis. However, this should not be a primary concern given that it is now possible to retrieve this information, for instance, from contact-tracing apps.

The remainder of the paper is organized as follows. Section 2 details the steps needed to compute DDC. Section 3 presents the results of the Monte Carlo experiment implemented to compare DDC to other network centrality measures. Finally, section 4 concludes.

2 Diffusion Delay Centrality

A diffusion process can be studied as a network-based phenomenon, where nodes represent agents, and links signal a direct contact between two agents. In what follows, we detail a method to quantify the importance of a node in decelerating a diffusion process in a network by using the DDC.

We describe our measure assuming that contagion does not follow a predetermined direction, but can spread from i to j and vice versa, i.e. the network is undirected, and that the probability of receiving the diffusion by direct contact is equal to everyone, that is the network is unweighted. However, we will show in the following that the DDC can be easily extended to other circumstances, and applied also when the network is directed and/or weighted.

We measure the speed of diffusion originating from agent i and reaching agent j by counting the minimum number of contacts required among network agents for the process to spread from i to j. We refer to this measure as to *shortest path length* (SPL) or geodesic distance, and we denote it with the symbol $d(i, j)$. The sum of the length of shortest paths originating from i are referred as to *farness* of agent i, and we denote it with $f(i)$.[3]

$$f(i) = \sum_{j} d(i, j) \tag{1}$$

We assume now that diffusion can start from any node. Since the starting point of the process is unknown, our goal is to slow down the speed of diffusion process regardless of where it initially started, by increasing the distance between all nodes: i.e. incrementing the farness between each pair of nodes. Our aim is then to identify the node who, once isolated, will maximally increase the farness of all other nodes. In order to do that, we first need to calculate the overall speed of diffusion in the network, that is the sum of all agents' farness:

$$f(N) = \sum_{i} f(i) \tag{2}$$

Then, we need to identify the node to be isolated in order to maximally increase $f(N)$. This is done iteratively.

First, we isolate the generic node i, thereby interrupting all the shortest paths on which i was lying over. This has the effect of altering the speed of the diffusion process from any node pair connected by one of the interrupted paths. In fact, all the other alternative paths starting from one node and

---

[3] Observe that by definition the inverse value of the farness is equal to the closeness centrality of agent i.

reaching another will be by definition equal or longer with respect to the shortest path passing through i.

Second, we calculate the farness of nodes now that i has been removed, and diffusion flows through different paths. As a result, we obtain a new value for the overall speed of diffusion in the network: i.e., $f(N_i)$.

Third, we compare the speed of diffusion in the network when i is isolated, i.e. $f(N_i)$, with that obtained when i is included in the network. In order to correctly make a comparison, we need to confront the overall speed in the two networks considering the same number of nodes, and exclude from the computation of $f(N)$ the value of farness relative to i, that is:

$$f(N) - 2f(i)$$

The reason why we remove the farness of i twice, i.e. $f(i)$, from $f(N)$ is straightforward. Since we are using an undirected network, the same shortest path exists between agent i and j, and between agent j and agent i. Hence, we need to discard both paths starting from and reaching agent i from the computation in order to correcly account for the removal of this node. Once this is done, it is possible to calculate the decrease in the diffusion speed of the network after i has been isolated, that is the delay diffusion centrality of node i:

$$DDC(i) = f(N_i) - [f(N) - 2 \cdot f(i)] \tag{3}$$

This process is repeated for every node contained in the network, so that each one is associated to a given DDC value. Then, DDC is used to rank nodes and identify those who, once removed, maximally increases the farness of all other nodes, and decelerate diffusion. When DDC is high, then node i is very important to slow down the diffusion. Hence, a policy aiming at decelerating the diffusion process will put this node into isolation. By contrast, if DDC is close or equal to zero, then node i is not relevant for determining the speed of diffusion. Consequently, a policy will not include this node in the list of those to be put into isolation.

Intuitively, high DDC nodes are those involved in many shortest paths, but embedded in a relatively isolated area of the network, characterized by small path redundancies. When removed, many shortest paths connecting the area of the network where they are located will be cut off, and this will have a major impact on the diffusion process (at least in this area). On the contrary, low DDC nodes are either poorly connected nodes, or well connected to well-connected nodes. Thus, by isolating them, we will not observe any significant change in the speed of the diffusion process. Taken together, these two observations suggest that DDC assigns a high rank to agents with two specific characteristics: first, they act as a major bridge for the diffusion process, meaning that they feature a high betweenness centrality; second, they are in contact with low-connected agents located at the periphery of the network: i.e. they feature a low eigenvector centrality. We will further explore this relation between DDC, eigenvector and betweenness centrality in Section 3.

In what follows, we detail how DDC works in practice, we discuss a number of interesting properties featured by this measure, and we present some examples to show how this measure can be computed using different network structures.

2.1 Diffusion Delay Centrality and Betweenness centrality

An interesting property highlighted by the DDC is that the DDC of a node with betweenness centrality equal to zero is zero by construction. In fact, betweenness centrality is equal to the number of shortest paths passing through a node. When a node features a betweenness centrality equal to zero, the node is not involved in any shortest paths, hence its removal from the network will not increase the farness of the other nodes, and its DDC will be equal to zero.

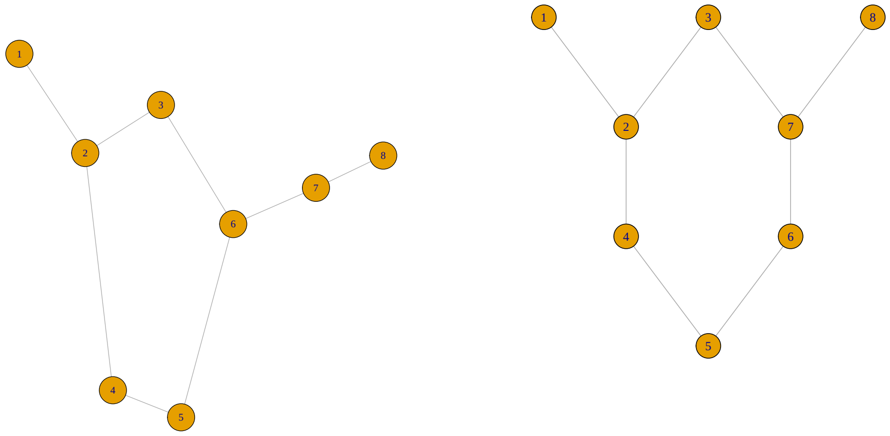
It follows from this argument that nodes without connections with any other nodes, i.e. isolated nodes, have DDC equal to zero. In fact, they do not lie on any shortest paths contained in the network, and by definition their betweenness centrality is equal to zero. As a consequence, one can disregard the presence of isolated nodes when computing DDC. Since these are already unable to get in contact with other agents and spread diffusion, it is not necessary to evaluate whether they have to be put into isolation or not.

At the same time, DDC has notable differences with betweenness centrality, because it includes a different set of information. The reason is that betweenness simply counts the number of shortest paths through a node, and it disregards the increase in the distances among agents resulting from removing that node. We clarify the main difference between these two measures using the example contained in Figure 1 . In both networks, node 3 has the same betweenness centrality. However, node 3 has a higher DDC in the network represented on the right compared to the network on the left. This is because, when removing node 3 in the network on the right, the distance between nodes 2 and 7 increases by 2 steps.[4] By the same reasoning, one can see that the distance between the node pairs 2 and 8, 1 and 7, and 1 and

---

[4] I.e. it takes alternatively 2 or 4 steps to reach node 7 from node 2, depending on whether node 3 is present or not.

Figure 1: Example 1.



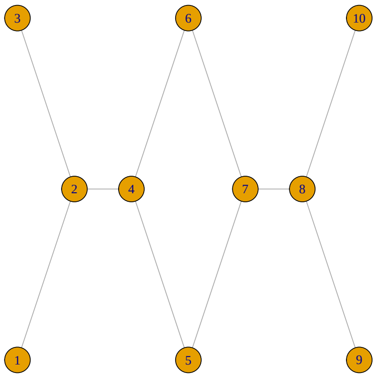
Note: two toy networks used for illustrative purposes. In each network, a node represents an agent. Two agents are connected if they interact with each other. Observe that Node 3 has the same betweenness in both networks, but it has a higher DDC in the network on the right.

8 increases by 2 when removing node 3. Of course, the same applies when considering distances from the opposite direction (i.e. from 7 and 8 to 1 and 2). It follows, that the DDC of node 3 is equal to 16.
On the contrary, the removal of node 3 in the network on the left has a smaller impact on the diffusion process, because the distance between the aforementioned node pairs increases only by 1. As a result, the DDC of node 3 is only 12 in this case. To sum up, DDC highlights that node 3 is more relevant in slowing down the contagion in the network represented on the right compared to that on the left.

Moreover, a node with betweenness centrality different from zero can have a DDC equal to zero, given the presence of path redundancies. This is the case of nodes 5 and 6 in Figure 2.

For these reasons, we argue that DDC is different from betweenness and other centrality measures which rely on the count of the number of shortest paths passing through a node, such as the Percolation Centrality (Piraveenan et al., 2013). We will further analyze this latter measure in Sections 2.4 and 2.5.

Figure 2: Example 2.



Note: A toy network used for illustrative purposes. Node 5 and 6 have a positive betweenness centrality but DDC equal to zero.

2.2 Diffusion Delay Centrality: dealing with different components

Until now we have implicitly assumed that the network is composed of a single component: i.e. a set of nodes which can directly or indirectly reach each other. However, it is often the case that a network features multiple components, and not all nodes can reach all the other nodes in the network. In theory, this does not represent a source of concern to us. In fact, since nodes belonging to two different components cannot get in contact with each other, one can study diffusion in the two components separately, and calculate the DDC of the nodes in one component without considering the presence of the other components. From a formal standpoint, the reason is that the distance between two nodes belonging to different components is infinite. Hence, when any node is put into isolation in one component, the distance between this node and the nodes located in another component remains unchanged: i.e. infinite.

However, dealing with the distance between nodes located in separate components is an issue that should not be overlooked. The reason is that a one-component network can sometimes be split into two or more components when isolating a node. To understand why, consider a network where a node has only one connection: i.e. its degree centrality is equal to one. When node j has only one link with node i, and we put i into isolation, also j becomes automatically isolated: i.e. the network splits into two components, one containing only j, and the other containing all other nodes.

By definition, the distance between two components is equal to infinity. As a result, when calculating the SPL reaching and leaving from j when i is removed might become cumbersome. Of course, different choices can be made depending on the definition of contagion that is applied. A possibility is to give an infinite value to the DDC of a node that, if isolated, protects from diffusion at least one more node compared to the baseline setting. However, this choice implies that many nodes may potentially have DDC equal to infinity, making it difficult to create a usable ranking of nodes. In fact, a node able to isolate many nodes, and a node able to isolate only one node, will have both DDC equal to infinity, and they will be ranked in the same way even if their contribution to slowing down diffusion is different.

Another approach is to consider the maximum theoretical SPL from two nodes in the one component network composed of n nodes, that is $n - 1$. Then, one can use a value above $n - 1$ to measure the distance between nodes located in two different components after the network is split. In other words, when the distance $d(i, j)$ between two nodes is infinite, one can replace it with a finite large value $d*$ with $n \leq d* < \infty$. This allows to measure the distance i and j with a metric that is larger than the maximum theoretical SPL in the network, thus mimicking a very large distance. As a result, computation becomes tractable and the ranking emerging with this choice will be unaffected in most practical situations, even if the DDC value might be considered arbitrary. In our examples, we adopt this approach and specifically we set $d = n$.

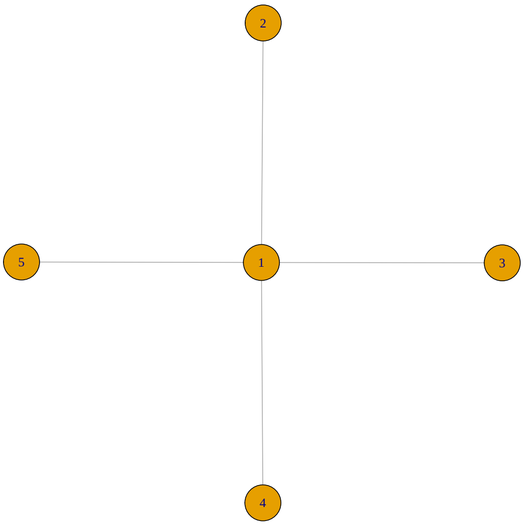
For instance, in the star network composed of 5 nodes represented in Figure 3, we can assume that an isolated node has a distance equal to 5 from all the other nodes, that is the total number of nodes of the network (i.e. $d = n$). Then, we can compute the DDC of each node:

- node 1 has a DDC equal to 36. When isolating node 1, the distance between nodes 2 and 3 increases from 2 to 5. The same happens for the distance between nodes 2 and 4 or 5. As a result, when isolating node 1, the distance of node 2 from the other nodes increases by 9 steps. The same metric is obtained when considering nodes 3, 4 and 5. Hence, the total DDC of node 1 is 36;

- nodes 2, 3, 4, and 5 have a DDC equal to 0, confirming what is stated in Section 2.1, that is the DDC of a node with betweenness centrality equal to zero is zero by construction.

Consequently, node 1 ranks first while all the other nodes ranks second in terms of DDC, and this result does not change for every $d*$ in the proposed range $n \leq d* < \infty$.

Figure 3: Example 3.



Note: A toy network used for illustrative purposes. Assuming that an isolated node has distance=5 from all the other nodes, node 1 has a DDC=36, while other nodes have a DDC=0.

2.3 Diffusion Delay Centrality using directed and weighted networks

When dealing with the potential threat of a diffusion that must be stopped, but the origin of the diffusion process is unknown, and it is still unclear how contagion is transmitted, a safe choice is to assume that any potential area of the network may be the cause of the diffusion spread, and all nodes can be equally infected.

While we assume no prior knowledge over the diffusion process, we can still hypothesize to have information about the direction of the links in the network (this is for instance the case of international trade networks). For instance, if links register the movements of agents, and we have knowledge of such movements, it is possible to leverage this information to predict the potential direction of diffusion across network areas, and identify those nodes responsible for spreading diffusion along a specific direction. Formally, this requires calculating the DDC of agents using a directed network. The computation of the DDC can be easily extended to the case of directed networks using Equation 4:

$$DDC(i) = f(N_i) - [f(N) - f(i) - \sum_j d(j,i)] \tag{4}$$

where $d(j,i)$ is the SPL from all nodes js to node i.

Another interesting case is when the interactions of agents have different intensities, that is the case when the network under study is weighted. In this situation, one might suspect that only frequent or long interactions are responsible for contagion, while sporadic contacts between agents cannot be responsible for spreading the diffusion. Under these circumstances, we can assume that the contagion is transmitted only if the weight of the link is above a certain threshold. This approach can be used when the network is composed of both strong and weak links, according to their weight, and contagion is likely to pass only through strong (weak) links. Another application of this approach is when diffusion is governed by a stochastic process, and the probability of contagion is associated to the weight of the links of the network. We shall leave a detailed study of these applications to future work.

2.4 Conditional Diffusion Delay Centrality

Until now we have assumed no prior knowledge on the source of diffusion, and we have used the DDC to rank nodes so as to maximally increase the farness between nodes, and slow down diffusion regardless of where it began. However, when the source of diffusion is known, nodes can be ranked *conditional* to the information given, i.e. knowing the starting point of the contagion, one can focus only on the set of nodes directly exposed to the source of diffusion. We discuss this case in example 4 reported in Figure 4.

When the source of diffusion is unknown (*unconditional* choice), the DDC is computed as usual. By definition, nodes 1 and 5 have a DDC equal to 0, because their betweenness centrality is zero. Nodes 2 and 4 have a DDC equal to 12. Computation of the DDC of nodes 2 and 4 is straightforward.

Take for instance node 4. When this is put into isolation, node 5 becomes isolated. As suggested in Section 2.2, we can set the distance of node 5 from all other nodes equal to 5: i.e. the maximum theoretical SPL in the network plus 1. Consequently, the distance of node 5 from all other nodes increases by 6.[5] At the same time, when node 4 is isolated, the distance of node 1 from all other nodes except for node 5 remains unchanged, while its distance from node 5 increases from 4 to 5. The same applies for nodes 2 and 3: i.e. their distance from all other nodes remains unchanged, except for node 5. As a result, the distance from this node increases respectively by 2 and 3. Consequently, the overall DDC of node 4 is: 6+1+2+3 = 12. By the same reasoning, node 3 has a DDC equal to 16. Consequently, when the source of diffusion is unknown and only one node can be isolated, the DDC suggests to isolate node 3.

By contrast, when the source of contagion is known (*conditional* choice), the decision of who should be isolated can change. For instance, if the source of contagion is node 1, the best choice is to isolate node 2, so that diffusion cannot spread throughout the rest of the network. Formally, conditional DDC of generic node i can be computed by looking at the increase in the farness of the infected node when i is removed from the network, and disregarding the impact of its removal on the farness of all other nodes. This is done with the formula:

$$CDDC(i) = f(j)_{N_i} - [f(j) - d(j,i)] \tag{5}$$

Where j is the infected node, $f(j)_{N_i}$ and $f(j)$ are respectively the farness of j when i is removed from or included in the network, and the term $d(j,i)$ is the distance between i and j. As discussed in Section (2), the term $d(j,i)$ is included in the formula to remove from the calculation the shortest path existing between nodes i and j.[6]

It follows from equation 5 that when confronted with a conditional choice, that is when the source of contagion is known, CDDC assigns more importance to the node(s) most relevant at the beginnning of the diffusion process: i.e. the node(s) which can maximally increase the farness of the rest of the network from the source of contagion.

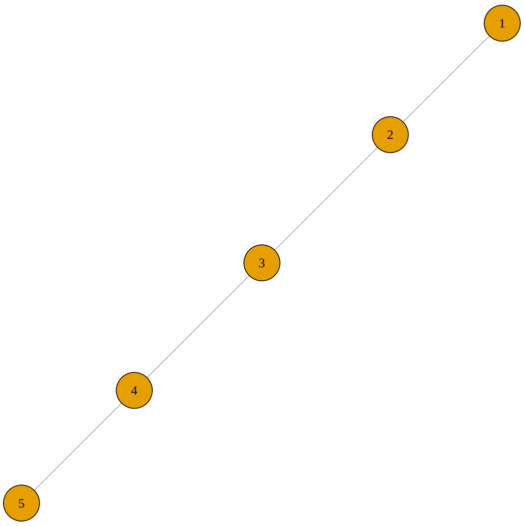
The extant literature provides us with a large array of tools for dealing with a conditional choice, complementing with the information provided by the DDC. These are the measures derived from percolation methods (Del

---

[5] Specifically, distance from node 1 increases by 1, distance from node 2 increases by 2, and distance from node 3 increases by 3.

[6] Observe that Equation 5 can be easily extended to the case of multiple infected nodes, by summing the CDDC of i with respect to each infected node j.

Figure 4: Example 4.



Note: A toy network used for illustrative purposes. Observe that node 3 has the highest DDC, hence the isolation of this node is the best *unconditional* choice to slow down the contagion. The same does not necessarily apply when the source of diffusion is known, and nodes can be ranked *conditional* to the information given.

Ferraro et al., 2018; Pei et al., 2017, 2018; Teng et al., 2016). One that recently emerged is the *percolation centrality* by Piraveenan et al. (2013),[7] henceforth PC, which measures the proportion of "percolated" paths in a network passing through a node: i.e. the paths beginning from an infected (percolated) node. The advantage of PC over other centrality measures is that it allows to rank the impact of all nodes over the contagion transmission at any given time, not just at the beginning.[8] Moreover, PC allows to con-

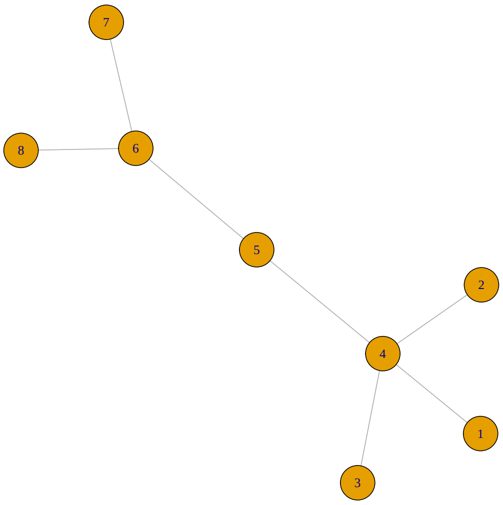
---

[7] Observe that this measure is different from that discussed by Pei et al. (2017), even though the two metrics have the same name.

[8] PC also allows to consider that nodes might feature different "viral loads" (i.e. they have a different ability to transmit the contagion). However, this situation violates the

stantly update the ranking of nodes while collecting new information about the diffusion spread. That being said, it is easy to see that there are cases when CDDC and PC provide the same information. This is for instance the case of the network presented in Figure 5.

Figure 5: Example 5.



Note: A toy network used for illustrative purposes. Here, PC and Conditional DDC returns the same rankings for nodes that should be first isolated to reduce the spread of contagions starting from different nodes.

Suppose that node 5 is infected. In Table 1, we present how nodes 4 and 6 are ranked by PC and CDDC in this situation.[9] Table 1 shows that when node 5 is infected, both metrics provide the same ranking. Node 4 should be the first to be isolated, because it is the node which can maximally spread contagion from node 5. Node 6 instead is considered of less importance, because its impact over the contagion process is relatively minor compared to that of node 4.

The consistency in the ranking between the two metrics emerges also when choosing a different source of contagion. Consider for instance the case when node 1 is infected, and we want to investigate the impact of nodes 4, 5, and 6 over the diffusion process. In Table 2, we present the rankings obtained in this case. Table 2 again shows that PC and conditional DDC agree on

---

assumption underlying this work, where one node is either infected or not. Therefore, it falls outside the scope of this paper to discuss this case.

[9] We exclude from the discussion nodes 1, 2, 3, 7, and 8 because their betweenness centrality is equal to zero, and thus by definition they have PC and CDDC equal to zero.

Table 1: PC and DDC ranking of nodes in the network of Figure 5 when node 5 is infected.

Node	PC ranking	Conditional DDC ranking
4	1	1
6	2	2

Note: for each considered node in the network of Figure 5, indicated in column 1, we report its ranking according to PC (column 2), and CDDC (column 3) when node 5 is infected.

Table 2: PC and DDC ranking of nodes in the network of Figure 5 when node 1 is infected.

Node	PC ranking	Conditional DDC ranking
4	1	1
5	2	2
6	3	3

Note: for each considered node in the network of Figure 5, indicated in column 1, we report its ranking according to PC (column 2), and CDDC (column 3) when node 1 is infected.

the ranking of nodes to be isolated to decelerate contagion. This implies that DDC and PC might provide consistent information in the presence of a conditional choice. Of course, this is not always the case.[10]

2.5 Additional comments on the difference between DDC and PC

It is important to stress that the consistency in the ranking provided by DDC and PC is not necessarily verified when dealing with an unconditional choice. Indeed, when the source of contagion is unknown, that is all nodes are in the same “percolated state”, PC is by definition equal to betweenness centrality.[11] In this case, the DDC can provide additional information to

---

[10] We shall leave to future research a detailed comparison of the differences between CDDC and PC, because this is out of the scope of this work.

[11] When the source of contagion is unknown, PC is obtained as follows. First, attribute the source of contagion to one node and compute the PC of all other nodes. Second, iterate the procedure over all other possible infected nodes. Third, assign to a node a PC equal to the average PC obtained in all the considered scenarios. It is easy to show that this “average percolation centrality” (APC) is equal to the betweenness centrality of a node.

policy makers compared to that obtained from PC. Consider for instance the case of the network in left panel of Figure 1. In Table 3, we report the ranking of the nodes embedded in this network provided by PC and the unconditional DDC. In this context, while PC (and betweenness centrality) assigns the same importance to nodes 3 and 7, the unconditional DDC indicates that node 7 might have a potentially higher impact over the contagion process compared to node 3. The reason is that node 7 plays a major role in diffusing contagion from node 8 (and viceversa). Once removed, if node 8 is infected, contagion is stopped. Hence, at least in one case, node 7 may have a crucial impact over the diffusion process, and in all other cases it is able to save node 8 from contagion. On the contrary, regardless of where the contagion starts, when isolating node 3, diffusion can still spread on a different route. Following this logic, when the source of contagion is unknown, DDC signals that node 7 should receive more attention than node 3. Coherently, we observe that nodes 3 and 7 have the same betweenness centrality, but node 3 features a higher eigenvector centrality due to the path redundancies in which it is embedded.

Table 3: PC and DDC ranking of nodes in the network of left panel of Figure 1 when nodes are not infected (or they are all equally infected).

Node	PC ranking	Unconditional DDC ranking
2	2	2
3	3	4
4	5	6
5	4	5
6	1	1
7	3	3

Note: for each considered node in the network of left panel of Figure 1, indicated in column 1, we report its ranking according to PC (column 2), and unconditional DDC (column 3) when nodes are not infected (or they are all equally infected).

The reason for this discrepancy between PC and DDC is straightforward, and it follows from our discussion of example 1 in section 2.1. The former measure focuses on the role of nodes in connecting different paths, and in this context it cannot leverage any information about the percolation status of the agents. Instead, the latter evaluates the extent to which a node can decelerate contagion by maximally increasing the path distance between all other nodes.

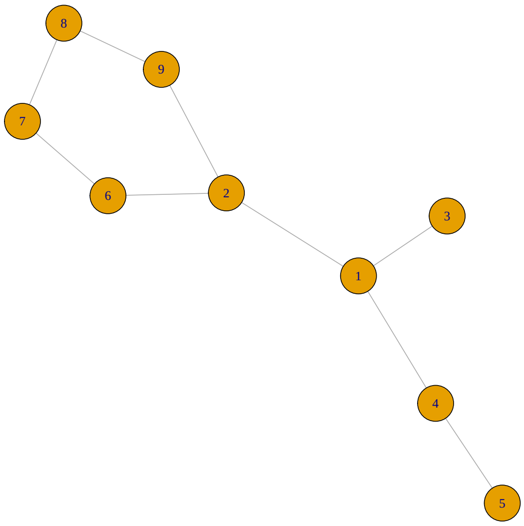
For additional details see Piraveenan et al. (2013).

2.6 Diffusion Delay and Speed

Additional insights on the relation between slowing down contagion and DDC can be observed by looking at the network in Figure 6. In order to do that, for each node *i*, we calculate the average speed of contagion diffusion of all other nodes before and after *i* is isolated. We report the results in Table 4. To understand what we do in practice, take for instance node 1. We compute how many steps are needed for a contagion starting from node 2 to reach, on average, all other nodes except for node 1 (nodes from 3 to 9). We then do the same considering a contagion starting from node 3 to reach all other nodes except for node 1 (i.e node 2, and node from 4 to 9). The same calculation is then replicated considering a contagion starting from each remaining node (i.e. nodes from 4 to 9). Finally, we take the average value obtained from all these calculations. This is the average speed of contagion when considering node 1, and we report it under the column "Avg Speed" (first line, referring to node 1). Then, we replicate the whole procedure in a network where node 1 has been removed, and we report the value obtained in this way under the column "Avg Speed after node isolation" (observe that the speed to reach nodes that are not reachable is set to 9, which is the number of nodes in the network). Then we take the difference between these two metrics: i.e. the average *difference in speed* generated from isolating node *i*. Additionally, we plot the difference in speed relative to *i* against its DDC in Figure 7. It is easy to see that a linear relation exists between the two measures: i.e. difference in speed increases for higher values of DDC.

An interesting feature that should also be observed is that the node which maximally delays diffusion in the network (i.e. the node with the highest DDC ranking) is not necessarily the node which maximally spreads contagion if it is the first infected node. This feature has been already discussed in extant literature (see among others Restrepo et al., 2006; Klemm et al., 2012). We show this again using example 6 reported in Figure 6. In this example, the node with the highest DDC ranking is node 1. When diffusion starts from this node, contagion reaches 3 nodes in 1 step, other 3 nodes in 2 steps, and the remaining nodes in 3 steps. Now consider when diffusion starts from node 2. In one step, diffusion reaches 3 nodes (1, 6 and 9). In two steps, it reaches other 4 nodes (3, 4, 7, and 8). In three steps, diffusion has reached all nodes. As one can see, after 2 steps, node 2 has infected more nodes than node 1: i.e. 7 instead of 6. Therefore, node 1 is more relevant for delaying the spreading of the contagion, while node 2 is more relevant for speeding up the contagion. This difference is not surprising, because the network is split into three different components when node 1 is isolated. By contrast, only two components are generated when node 2 is put into isolation.

Figure 6: Example 6.



Note: A toy network used for illustrative purposes. In this network, the node (node 1) which maximally delays diffusion is different from the node (node 2) able to infect most rapidly other nodes.

2.7 Higher order Diffusion Delay Centrality

Until now, we have analyzed the impact on the diffusion process of putting into isolation one node, this is to say that we have considered a *first order* delaying mechanism of the diffusion process. However, one can consider the situation when the policy maker has already isolated the node with the highest DDC ranking, thus generating a new network which modifies the DDC value of the remaining nodes. In this situation, one should consider a *higher order* delaying mechanism. In practice, this implies calculating the DDC of the nodes which have not been put into isolation. We consider this case by discussing example 7 in Figure 8. In the figure, only node 2 has DDC different from zero. Therefore, before diffusion starts, a policy maker should

Table 4: Speed of diffusion and DDC of nodes in the network of Figure 6

Node	DDC	Avg Speed	Avg Speed after node isolation	Difference in Speed
1	190	2.6429	6.0357	-3.3928
2	178	2.6786	5.8571	-3.1785
4	74	2.4643	3.7857	-1.3214
6	10	2.5357	2.7143	-0.1786
9	10	2.5357	2.7143	-0.1786
7	2	2.3929	2.4286	-0.0357
8	2	2.3929	2.4286	-0.0357
3	0	2.3929	2.3929	0
5	0	2.2143	2.2143	0

Note: for each node in the network of Figure 6, indicated in column 1, we report its DDC (column 2), the average speed of diffusion for a contagion starting from all the other nodes to all the other nodes, except for the node indicated in column 1, when this is included in the network (column 3), or removed from it (column 4), and the difference between these two metrics (column 5).

put into isolation node 2. Once this is done, node 3 becomes the only node relevant for delaying diffusion among the other nodes: i.e. it has the highest DDC ranking. As a result, the higher (second) order delaying mechanism indicates that node 3 should be the next one to be put into isolation.

3 Centrality measures to find nodes relevant for the contagion

In this Section, we compare the ranking obtained using the DDC with that produced by other common measures of network centrality, namely degree, eigenvector, betweenness and closeness centrality.[12]

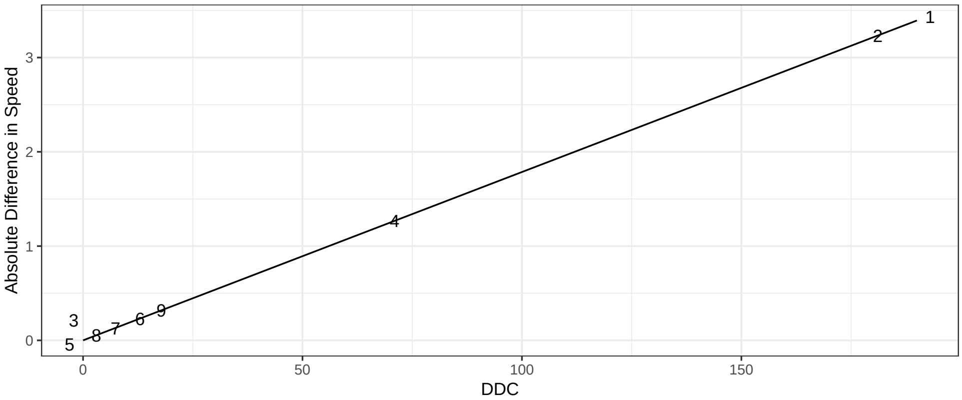
In what follows, we will show that DDC ranking can be easily proxied by a combination of the ranking produced by eigenvector and betweenness centrality. This can be done in two different steps:

1. assign lowest ranking to all nodes with betweenness centrality equal to zero;
2. let BE be the difference in the value of betweenness centrality and eigenvector centrality. Use BE to rank the remaining nodes, so that

---

[12] Computations are performed using the R package IGRAPH.

Figure 7: DDC vs Difference in Speed for the nodes in the network of Figure 6



Note: on the x-axis, we report the DDC of the nodes in the network in Figure 6. On the y-axis, we record the absolute difference in the speed of contagion when removing a node in the network in Figure 6, as reported in Table 4. Labels indicate to which value each node corresponds.

the node with highest ranking is that with the maximum value of BE, and the node with lowest ranking is that with the minimum value of BE.

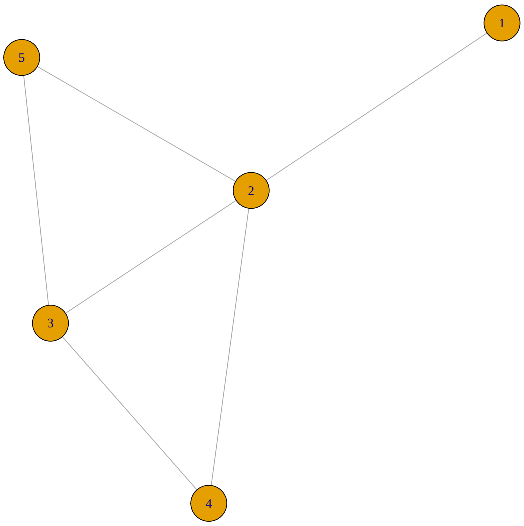
Observe that the support of eigenvector centrality is included between zero and one. For this reason, one needs also to rescale the support of betweenness centrality between zero and one in order to compute BE. This is done with the formula: $\frac{betweenness\ centrality_i}{max(betweenness\ centrality)}$.

The relation between BE and DDC provides a simple intuition on the role played by nodes featuring the highest DDC in delaying the network process.[13] The DDC assigns a high rank to agents with a high betweenness centrality, meaning that they act as a central hub of the diffusion process, and with a low eigenvector centrality, which implies that they are in contact with low-connected agents located at the periphery of the network. For this reason, we claim that high DDC agents act as the gatekeepers of the fringe of the network. In fact, when these are isolated, the spreading of diffusion

---

[13] The BE rule also features important benefits in terms of computational complexity compared to DDC. In fact, the computational complexity to calculate the DDC of all nodes in a network is 2n*O(m+n). On the contrary, computational complexity to obtain the betweenness and eigenvector centrality of all nodes in a network is O(m+n)+O(mn), which is considerably lower with respect to 2n*O(m+n).

Figure 8: Example 7.



Note: A toy network used for illustrative purposes. The BE rule ranks node 3 as more important than nodes 1, 4 and 5, whereas the DDC assigns the same importance to all these nodes. Considering the second order DDC, node 3 becomes more relevant than nodes 1, 4 and 5.

from the periphery to the core of the network (or vice cersa) is prevented or at least decelerated.

The extent to which the BE and DDC are related is easy to verify by looking at the examples already presented. We summarize the ranking of the nodes returned by the DDC in some given examples in Table 5 along with the ranking produced by the BE rule. For completeness, we also report the centrality values obtained from DDC, betweenness, eigenvector, and BE.
In the first two reported examples, DDC and the BE rule produce the same ranking. Example 7 instead presents two slightly different rankings. Still both rankings indicate the same node (i.e. node 2) to be the first which has to be put into isolation. However, the BE rule ranks node 3 as more

important than nodes 1, 4 and 5, while the DDC assigns the same importance to these nodes. To understand this difference, we observe that node 3 is the most important node in spreading diffusion when node 2 is already isolated, and we hypothesize that the BE rule is able to control for higher (second) order mechanism as explained in Section 2.7. That said, what is important to stress again is that the two rankings are consistent in pointing out what is the node to be first put into isolation.

Of course, one can object that the examples provided so far are not exhaustive and a more thorough investigation of the DDC properties is required. For this reason, in the next Subsection we perform a large Monte Carlo analysis to show the robustness of the relation between the ranking produced by the two measures. Here, we also discuss the situations in which the BE rule is less effective in reproducing the DDC ranking.

3.1 Monte Carlo analysis

We perform a Monte Carlo experiment by generating 49000 random networks (Erdős and Rènyi, 1959, 1960, 1961) for different random network configurations. In practice, we use a network with 100 nodes, and a variable number of edges is randomly generated. We perform 1000 simulations for each given number of edges included from 100 to 4900 with steps of 100 edges (therefore, we perform 49 steps). We do not to consider networks with a lower or higher number of edges to avoid degenerate situations, where it is unlikely to find a relevant pattern of diffusion (Anderson and May, 1992; Rogers, 1995; Colizza and Vespignani, 2007).[14]

For each simulation we compute the DDC of each node. In doing so, we do not consider simulations where all nodes have a DDC equal to zero. As already explained, in these cases, the DDC ranking is meaningless because there are no nodes able to slow down the contagion.[15] As a result, we focus

---

[14] In the simulations, we generate networks with a single component where all nodes have degree centrality higher than 1 so as to make the computation easier (see Section 2.2 for additional details). This implies a reduction of the number of nodes in simulations with an initial number of edges equal to 100 and 200 (and very rarely also in network up to 600 edges): simulations with 100 edges feature on average 47.9 nodes, and simulations with 200 edges feature on average 91.1 nodes. Consequently, in these simulations, also the final number of edges is reduced compared to the initial amount.

[15] From a formal standpoint, it means that we have to discard a number of simulations producing networks with a density higher than 25% and below 40%, and all simulations returning networks with a density higher than 40% (practically, this implies that we discard all simulations producing networks with a number of edges above 2000) Observe that network density is equal to the ratio of the actual edges over the potential number of edges in the network.

Table 5: Comparison between DDC ranking and the ranking obtained with the proposed BE rule.

	DDC	DDC rank	Bet. Cent.	Eig. Cent.	Difference	Bet-Eig rank
Example 4						
Node 1	0	4	0.00	0.50		4
Node 2	12	2	0.75	0.87	-0.12	2
Node 3	16	1	1	1	0.00	1
Node 4	12	2	0.75	0.87	-0.12	2
Node 5	0	4	0.00	0.50		4
Example 6						
Node 1	190	1	1	0.43	0.57	1
Node 2	178	2	1	0.52	0.48	2
Node 3	0	8	0	0.19		8
Node 4	74	3	0.41	0.24	0.17	3
Node 5	0	8	0	0.11		8
Node 6	10	4	0.30	0.37	-0.07	4
Node 7	2	6	0.07	0.3	-0.23	6
Node 8	2	6	0.07	0.3	-0.23	6
Node 9	10	4	0.30	0.37	-0.07	4
Example 7						
Node 1	0	2	0.00	0.22		3
Node 2	18	1	1	0.58	0.42	1
Node 3	0	2	0.14	0.52	-0.38	2
Node 4	0	2	0.00	0.41		3
Node 5	0	2	0.00	0.41		3

Note: we report for each node in a considered network, indicated in column (1): its DDC and how DDC ranks it with respect to all other nodes in the network (columns 2 and 3, respectively); its betweenness and eigenvector centrality (columns 4 and 5, respectively); the difference between these measures (column 6); how the node is ranked using the metric reported in column 6 compared to the other nodes in the network (column 7).

the analysis on simulations producing networks where the sum of agents' DDC is above zero, that is 15,557 simulations. Here, we use the DDC in order to rank the nodes according to their importance in slowing down diffusion. Then, we compare the DDC ranking with the ranking obtained with the BE rule, and with the rankings obtained with degree, closeness, betweenness, and eigenvector centrality.[16] In order to compare rankings, we use the Spearman's Rho. In Table 6, we report the number of times in which the BE rule obtains a higher value of the Spearman's Rho compared to one of the other centrality measures: i.e. it produces the ranking that most resembles the one returned by the DDC. Complementing with this, Figure 9 shows the average Spearman's Rho comparing the rankings between the DDC and the other measures as obtained by all the simulations performed for a given (initial) number of edges.[17] Both Table 6 and Figure 9 show that the BE rule clearly outperforms all the other centrality measures in replicating the DDC ranking. In particular, the Spearman's Rho appears to be very low for the ranking obtained with the eigenvector centrality. Also the Spearman's Rho comparing DDC and closeness centrality rankings is surprisingly low, even though the DDC is obtained from the computation of farness, which is the inverse of closeness.

With respect to the BE rule, only betweenness centrality manages to better replicate the ranking of the DDC for a given number of times. To dig deeper into this finding, we compare the performance of the BE rule versus the performance of the betweenness centrality in replicating the DDC ranking in different situations (i.e. by considering networks with a different number of edges). We present the results of the analysis in Table 7. This highlights that the BE rule almost always obtains the best performance when the number of edges is included between 200 and 2000. Only when the number of edges is very low (up to 100), the BE rule is outperformed by the betweenness centrality. Hence, we can conclude that the BE rule best reproduces the ranking of the DDC when the network is not extremely sparse. When the network is very sparse instead, betweenness produces a ranking most similar to that of the DDC compared to the BE rule. Still, the BE rule has a high degree of correlation with the DDC, with an average Spearman's Rho above 0.9. We show this in Figure 9. In the next subsection, we discuss these results in more detail.

---

[16] We provide additional insights on the relation between DDC and other centrality measures in Appendix A.

[17] We stress again that in each simulation we generate a network with a fixed number of edges. This number is what we refer to as the initial number of edges. However, the reader should remember that after a network is produced by a simulation, some edges and nodes might be removed to comply with our theoretical setting (i.e. analyzing a network composed of a single component where all nodes have degree centrality higher than one). As a result, the number of edges contained in a network after this procedure is completed might be different from that declared at the beginning of the simulation. This is especially the case of simulations with an initial number of edges equal to 100 or 200, as explained above.

Table 6: Performance of the BE rule compared to other centrality measures in the simulated networks.

Higher Spearman's Rho	Degree	Closeness	Betweenness	Eigenvector
BE rule	15238	15239	13545	15186
Other centrality measure	319	318	1341	317
Same result	0	0	671	54

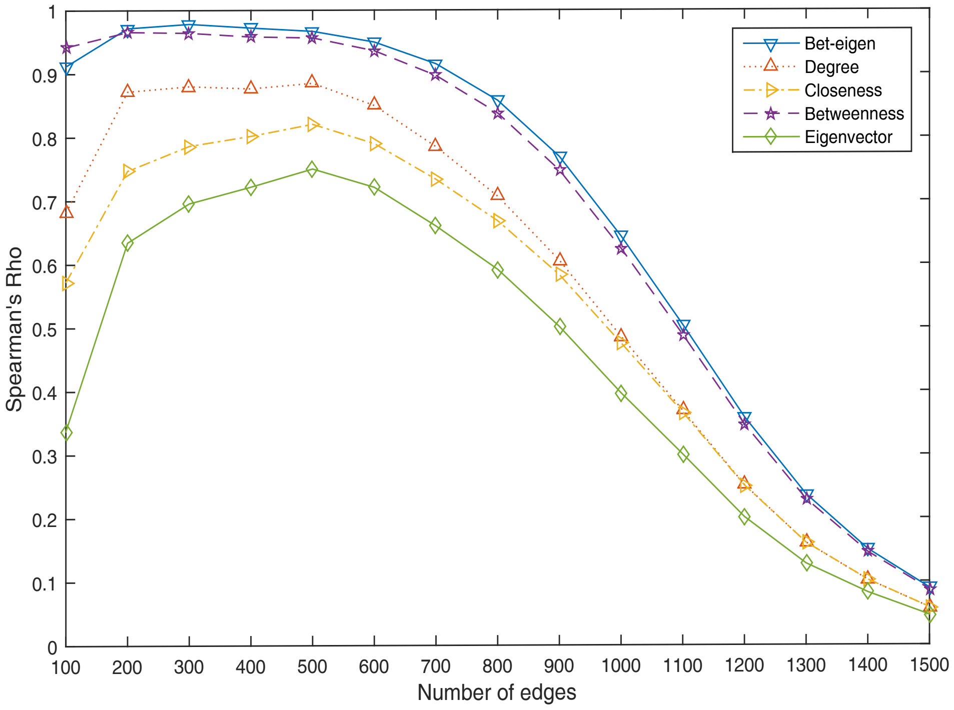
Note: row "BE rule" reports the number of times in which the BE rule obtains a higher value of the Spearman's Rho compared to the other centrality measure in the simulated networks. Row "Other centrality measure" reports the number of times in which the degree, closeness, betweenness, and eigenvector centralities respectively obtain a higher value in the simulated networks. Row "Same result" compare the performance of the considered metrics, by reporting the difference in the value between the first and the second row.

3.1.1 When the BE rule works well

Using the statistical properties of the simulated networks (the first three central moments of the distribution of node degree, closeness, betweenness, and eigenvector centrality, the network density, the number of edges and the number of nodes), we perform a number of regression analyses in order to understand what are the characteristics of the network explaining a high Spearman's Rho between the DDC ranking and that produced by the BE rule. Our main result is that a non-linear relationship exists between the density of the network and the Spearman's Rho, which is summarized in Figure 10. When increasing the density of the network from 0.05 to 0.1, the BE increases its ability to predict the DDC ranking, which is consistent with our claim in Section (3.1) that too sparse networks might generate degenerate situations which are unfit to our analysis, and the correlation between the BE and the DDC increases when we sufficiently increase the density of the network. When the network density is around 0.1, the BE rule performs best in recovering the DDC ranking (Spearman's Rho is always above 90%), and it performs very well up to a network density equal to 20%. From that point on, we observe the presence of a negative relationship between the Spearman's Rho and the network density. Consequently, we conclude that the ranking obtained with the BE rule works better in the situations that are of most interest to us: when the network is not too sparse, i.e. diffusion is not locked within certain areas of the network, and it is not too dense, that is diffusion cannot spread across an almost fully-connected population.

We delve more into our findings by considering three different sets of

Figure 9: Mean Spearman's Rho computed between DDC ranking and rankings obtained with a different measure.



Note: for every batch of simulated networks with the same (starting) number of edges, x-axis, we report the Mean Spearman's Rho, y-axis, computed between DDC ranking and the rankings obtained with BE rule, degree/closeness/betweenness/eigenvector centralities.

networks: those where the number of edges is low (200 or less and, therefore density is always below 10%), intermediate, and large. Consistent with previous finding, when density is low, we see from figures 9 and 10 that the average Spearman's Rho between the DDC and the BE rule ranking increases as density increases.

When the number of edges is intermediate, that is the situation of most interest to us, the BE rule performs well and the Spearman's Rho between the ranking of this metric and the DDC is over 0.9 in networks with edges between 300 and 700. Then, as network density significantly increases, we observe a decrease of the average Spearman's Rho. As already said, this is

Table 7: Performance of the BE rule compared to the betweenness centrality in obtaining a higher value of the Spearman's Rho.

Edges	BE rule	Betweenness	Same result	Valid simulations
100	330	670	0	1000
200	763	237	0	1000
300	981	19	0	1000
400	999	1	0	1000
500	998	2	0	1000
600	999	1	0	1000
700	1000	0	0	1000
800	1000	0	0	1000
900	1000	0	0	1000
1000	999	1	0	1000
1100	995	5	0	1000
1200	987	12	1	1000
1300	922	66	12	1000
1400	767	144	88	999
1500	502	114	270	886
1600	228	54	201	483
1700	57	15	74	146
1800	13	0	20	33
1900	4	0	5	9
2000	1	0	0	1

Note: for every batch of simulated networks with the same (starting) number of edges, indicated in column 1, we report: the number of times in which the BE rule obtains a higher value of the Spearman's Rho compared to the betweenness centrality (column 2); the number of times in which the betweenness centrality obtains a higher value of the Spearman's Rho compared to the BE rule (column 3); the number of times the two metrics report the same Spearman's Rho (column 4); the number of simulated networks considered (column 5).

due to the fact that nodes become less and less relevant in slowing down the contagion, since almost all of them are connected and the diffusion process would be unaffected by putting one node into isolation (i.e. the DDC ranking becomes increasingly useless). It is not surprising that the relation between the DDC and the BE rule is even less evident when the number of edges is very high, since DDC ranking is meaningless in these situations (because all agents' DDC approaches or reaches zero).[18] As a result, we can conclude that the BE rule is useful when the density of the network is relatively low, i.e. below 20%, that is also when the results returned by the DDC are most informative.

In order to investigate whether other factors can be relevant beside density, we separately investigate specific groups of simulations returning networks with the same number of edges: i.e. from 400 to 900 edges. We choose these simulations because networks from these groups feature on average a high value of Spearman's Rho, meaning that we are examining situations when the BE rule is effective in recovering the DDC ranking, but they also display relevant differences, since the standard deviation of the Spearman's Rho is markedly high. Hence, it is possible to investigate what drives the effectiveness of the BE rule in recovering more or less efficiently the DDC ranking holding network density constant.

Unsurprisingly, given the relation between the BE rule and the betweenness centrality, we find that an important driver in explaining the relation between the BE rule and the DDC is the standard deviation of the betweenness centrality in a network: i.e. an increase of the standard deviation increases the Spearman's Rho value. A reason for that might be that a high standard deviation of the betweenness centrality signals situations in which there are few nodes that are important for spreading the contagion, and a trail of nodes that are less relevant. This confirms that a strict relation between the BE rule and the DDC exists when the network is neither too sparse nor too connected. These are the circumstances under which it is more desirable to use DDC, because we are dealing with not-degenerate networks where contagion can be actually slowed down by isolating specific nodes.

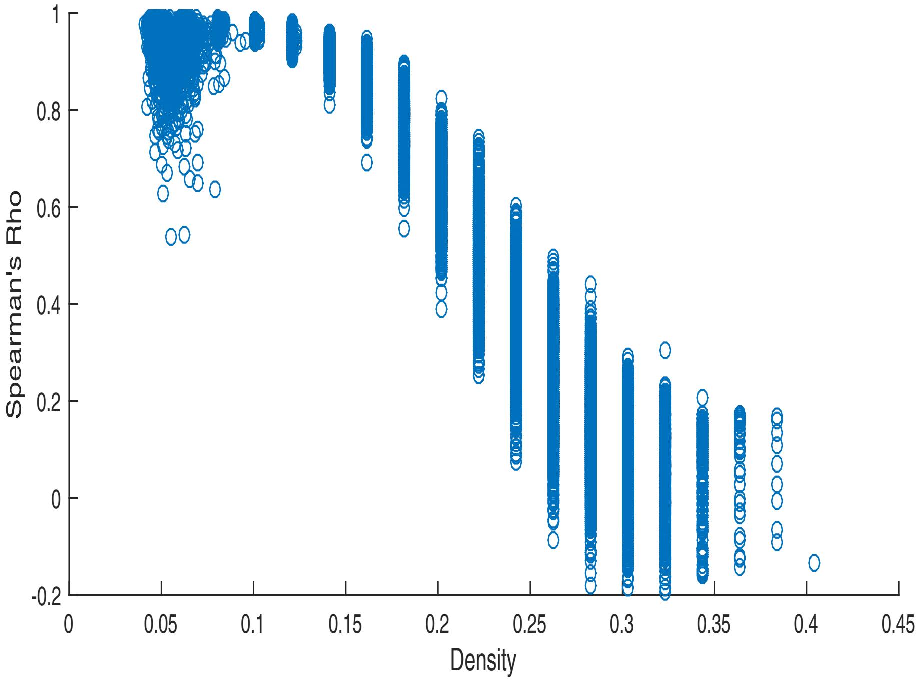
A specific analysis is then performed on simulations with 100 (or less) edges. In this case, the Spearman's Rho increases when both the standard deviation and the skewness of betweenness decreases, when skewness of eigenvector centrality decreases, and when the number of nodes increases.
This is not surprising, given the very small value of the network density. Under these circumstances, many nodes feature a small network centrality regardless of the measure considered. Similarly, few nodes are assigned to a high value by any network centrality measure. Consequently, the skewness of both betweenness and eigenvector centrality is higher compared to the value

---

[18] That said, one can still speculate on the reason for such differences in the rankings produced by the DDC and the BE rule when density is high. In these cases, the considerations made when discussing Example 7 and the higher order DDC might come in handy. In fact, one could speculate that the differences in the rankings are determined by the BE rule assigning more importance to nodes that are relevant for the mechanism considered by the higher order DDC.

found in all other simulations. When this is the case, the values produced by the *BE* rule tend to be very similar and close to zero, thus returning a ranking which is of little use to us. The reason is that both the betweenness and eigenvector centrality are rescaled between 0 and 1 when applying this rule, and when this is done while skewness is very high, most of the rescaled values approach zero. This confirms that the *BE* rule becomes less efficient in recovering the ranking of the DDC when the network is extremely sparse.

Figure 10: Spearman's Rho between DDC ranking and *BE* rule ranking related to network density.



Note: on the x-axis, we register the density of a simulated network. On the y-axis, we report the Spearman's Rho between DDC ranking and *BE* rule ranking.

4 Conclusions

This paper presents a simple measure to assess the importance of a node in decelerating diffusion processes in a network. Specifically, this measure quantifies the extent to which an agent is able to alter the average speed of diffusion in the network by putting him or herself in isolation. Consequently, it can be used to identify the agents to be targeted by a policy intending to reduce diffusion when the source of this process in unknown. The relevance of this topic in the literature has many different applications. For instance, this is the case when analyzing how efficiently real systems may be controlled by manipulating a single node (as in Klemm et al., 2012; Mobilia, 2003), or when investigating the impact of deleting a single node. This latter case is discussed in many different fields, when studying the failure of hierarchical networks such as nested split graphs (see among others König et al., 2012) and infrastructural networks (see for a recent discussion Currarini et al., 2016), an optimal change in the aggregate activity level of a network (e.g. Ballester et al., 2006; Lee et al., 2020), and strategic complementaries games on networks (Belhaj et al., 2004; Jackson and Yariv, 2005; Jackson and Zenou, 2015), where a shock propagating through a single central agent may end up affecting every other agent in the network.

We also show that the ranking of the DDC can be easily proxied by a different measure, that is the difference of the betweenness and the eigenvector centrality of the agents (the so called BE rule). This finding has the merit to facilitate the interpretation of the role played by high DDC nodes in the diffusion process. These are the nodes who control the access from the center to the periphery of the network. Once put into isolation, diffusion is confined within a certain area, or at least is significantly slowed down. This finding is also useful because it shows that one can easily reproduce the ranking of the DDC without actually computing DDC centrality, a task that might be often computationally heavy.

When examining the relation between BE rule and the DDC, we find that these metrics are most correlated when the network is not too sparse and it is not too dense: i.e. when considering non-degenerate networks, where diffusion is confined within certain areas or it rapidly spreads across the network because nodes are highly connected with one another. This is important because it shows that the BE rule can be trusted when considering most realistic and non-degenerate cases, which are the situations of most interest to us.

Future research should be dedicated to better understanding the relation between the BE rule and the DDC ranking in extremely sparse networks and/or with a small number of nodes, and in other canonical families of networks (e.g. scale-free networks). It would also be desirable to dedicate successive research to investigate the properties of the BE rule when networks are directed and/or weighted. Additional research should also be conducted on *higher order* delaying mechanisms, in order to develop an iterative version of the DDC allowing to design a strategy of isolation of multiple nodes at once. Finally, a promising venue of research would be to understand how the DDC can be applied to stochastic diffusion processes, modeling for instance the probability of contagion between agents through the weights of the network links, and to complex contagion processes (as in the work by Guilbeault and Centola, 2021), thus considering how peer reinforcement may act along the path of contagion to enable diffusion.

References

Anderson R.M, and May R.M. (1992). Infectious diseases of humans. *Oxford University Press.*

Ballester, C., Calvó-Armengol A., and Zenou Y. (2006). Who's who in networks, wanted: the key player. *Econometrica* 74, 1403–1417.

Banerjee A., Chandrasekhar A. G., Duflo E., and Jackson M. O. (2013). The Diffusion of Microfinance. *Science* 341, 6144.

Banerjee A., Chandrasekhar A. G., Duflo E., and Jackson M. O. (2014). Gossip: Identifying central individuals in a social network. *National Bureau of Economic Research (NBER)*, w20422.

Bardhan A. D., and Guhathakurta S. (2004). Global linkages of subnational regions: coastal exports and international networks. *Contemporary Economic Policy* 22(2), 225–236.

Bavelas A. (1950). Communication patterns in task?oriented groups. *The journal of the acoustical society of America* 6, 725–730.

Belhaj, M., Bramoullé, Y. and Deroïan, F. (2014). Network games under strategic complementarities. *Games and Economic Behavior* 88, 310–319.

Bloch F., Jackson M. O., and Tebaldi P. (2021). Centrality measures in networks. *arXiv*, 1608.05845.

Borgatti (2005). Centrality and network flow. *Social Networks* 27, 55–71.

Cai, J., De Janvry, A. and Sadoulet, E. (2015). Social networks and the decision to insure. *American Economic Journal: Applied Economics* 7(2), 81–108.

Centola D. (2013). Social media and the science of health behavior. *Circulation* 127(21), 2135–2144.

Centola D. (2018). How Behavior Spreads: The Science of Complex Contagions. *Princeton university press.*

Choi, H., Park, S. and Lee, J. (2011). Government-driven knowledge networks as precursors to emerging sectors: a case of the hydrogen energy sector in Korea. *Industrial and Corporate Change* 20(3), 751–787.

Colizza, V., Barrat, A., Barthélemy, M. and Vespignani, A. (2006). The role of the airline transportation network in the prediction and predictability of global epidemics. *Proceedings of the National Academy of Sciences* 103(7), 2015–2020.

Colizza, V., and Vespignani, A. (2006). Invasion threshold in heterogeneous metapopulation networks. *Physical review letters* 99(14), 148701.

Conley, T. G.and Udry, C. R. (2010). Learning about a new technology:

Pineapple in ghana. *American Economic Review* 100(1), 35–69.

Christakis N.A., and Fowler J.H. (2010). Social network sensors for early detection of contagious outbreaks. *PloS one* 5(9), e12948.

Currarini, S., Marchiori, C., and Tavoni, A. (2016). Network economics and the environment: insights and perspectives. *Environmental and Resource Economics* 65(1), 159–189.

Del Ferraro, G., Moreno, A., Min, B., Morone, F., Pérez-Ramírez, Ú., Pérez-Cervera, L., Parra, L.C., Holodny, A., Canals, S. and Makse, H.A. (2010). Finding influential nodes for integration in brain networks using optimal percolation theory. *Nature communications* 9(1), 1–12.

Dow, P. A., Adamic, L. A., and Friggeri, A. (2013). The anatomy of large facebook cascades. *AAAI Conference on Weblogs and Social Media (ICWSM)* 1, 12.

Erdős P. and Rènyi A. (1959). On Random graphs. *Publ. Math. Debrecen* 6, 290–297.

Erdős P. and Rènyi A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.* 5, 17–61.

Erdős P. and Rènyi A. (1961). On the strength of connectedness of a random graph. *Acta Mathematica Hungarica* 12(1-2), 261–267.

Fogli, A. and Veldkamp L. (2021). Germs, Social Networks, and Growth. *Review of Economic Studies* forthcoming.

Foster, A. D. and Rosenzweig M. R. (1995). Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of Political Economy* 103(6), 1176–1209.

Gould P. R. (1967). On the geographical interpretation of eigenvalues. *Transactions of the Institute of British Geographers*, 42, 53–86.

Guilbeault, D., Becker J., and Centola D. (2018). Complex contagions: A decade in review. Lehmann, S., Ahn, Y.-Y. (eds.) Complex spreading phenomena in social systems. *Springer*, 3–25.

Guilbeault, D., and Centola D. (2021). Topological Measures for Identifying and Predicting the Spread of Complex Contagions. *Nature Communications* 12(1), 4430.

Hüser A. C. (2015). Too interconnected to fail: A survey of the interbank networks literature.. *Journal of Network Theory in Finance*, 1(3), 1–50.

Jackson M. O. (2010). Social and Economic Networks. *Princeton university press.*

Jackson M. O. (2020). A Typology of Social Capital and Associated Network Measures. *Social Choice and Welfare*, 54, 311–336.

Jackson M. O., and López-Pintado D. (2013). Diffusion and contagion in networks with heterogeneous agents and homophily. *Network Science*, 1(1), 49–67.

Jackson M. O., Rogers B. W., and Zenou Y. (2017). The Economic Consequences of Social-Network Structure. *Journal of Economic Literature*, 55(1), 49–95.

Jackson M. O., and Yariv L. (2005). Diffusion on Social Networks. *Economie Publique*, 16(1), 69–82.

Jackson M. O., and Yariv L. (2007). Diffusion of Behavior and Equilibrium Properties in Network Games. *The American Economic Review (Papers and Proceedings)*, 97(2), 92–98.

Jackson M. O., and Yariv L. (2011). Diffusion, strategic interaction, and social structure. Benhabib, J., Bisin, A. and Jackson, M.O. (eds.) Handbook of social economics, Vol. (1). *North-Holland*, 645–678.

Jackson M. O., and Yariv L. (2011). Games on networks. Peyton Young H., Zamir S. (eds.) Handbook of game theory with economic applications, Vol. (4). *Elsevier*, 95–163.

Kang K. (2011). Overseas network of export promotion agency and export performance: the Korean case. *Contemporary economic policy*, 29(2), 274–283.

Karampourniotis, P. D., Sreenivasan, S., Szymanski, B. K., and Korniss, G. (2015). The impact of heterogeneous thresholds on social contagion with multiple initiators. *Plos One*, 10(11), e0143020.

Katz E., and Lazarsfeld P. (1955). Personal influence: The part played by people in the flow of mass communication. *Free press.*

Kempe D., Kleinberg J., and Tardos E. (2003). Maximizing the Spread of Influence through a Social Network. *Proceedings of the 9th International Conference on Knowledge Discovery and Data Mining* 137–146.

Kempe D., Kleinberg J., and Tardos E. (2005). Influential Nodes in a Diffusion Model for Social Networks. *Proceedings of the 32nd International Colloquium on Automata, Languages and Programming* 1127–1138.

Kim, D. A., Hwong, A. R., Stafford, D. D., Hughes, A. A., O'Malley, J. Fowler, J. H., and Christakis N. (2015). Social network targeting to maximise population behaviour change: a cluster randomised controlled trial. *The Lancet*, 36(9989), 145–153.

Klemm, K. M., Serrano, A., Eguíluz, V. M., and San Miguel M. (2012). A Measure of Individual Role in Collective Dynamics. *Scientific Reports*, 2(1), 1–8.

König, M.D., Tessone, C.J. and Zenou, Y. (2014). Nestedness in networks: A theoretical model and some applications. *Theoretical Economics*, 9(3), 695–752.

Lamberson P. J. (2016). Diffusion in networks. Bramoullé, Y., Galeotti, A. and Rogers, B. (eds.) The Oxford handbook of the economics of networks. *Oxford University Press*, 479–503.

Liu, J.-Gu., Lin J.-H., Guo Q., and Zhou T. (2016). Locating Influential Nodes via Dynamics-Sensitive Centrality. *Scientific Reports*, 6(1), 1–8.

Lee, L.F., Liu, X., Patacchini, E. and Zenou, Y. (2020). Who is the key player? A network analysis of juvenile delinquency. *Journal of Business & Economic Statistics*, forthcoming.

Loepfe, L., Cabrales, A., and Sánchez, A. (2013). Towards a Proper Assignment of Systemic Risk: The Combined Roles of Network Topology and Shock Characteristics. *Plos One*, 8(10).

López-Pintado D. (2006). Contagion and coordination in random networks. *International Journal of Game Theory*, 34, 371–381.

López-Pintado D. (2008). Diffusion in complex social networks. *Games and Economic Behavior*, 62, 573–590.

López-Pintado D. (2012). Influence Networks. *Games and Economic Behavior*, 75, 776–787.

Lü, L., Chen, D., Ren, X.-L., Zhang, Q.-M., Zhang, Y.-C., and Zhou T. (2016). Vital Nodes Identification in Complex Networks. *Physics Reports*, 650, 1–63.

Min, B., and San Miguel, M. (2018). Competing contagion processes: Complex contagion triggered by simple contagion. *Scientific reports*, 8(1), 1–8.

Mobilia, M. (2003). Does a single zealot affect an infinite group of voters?. *Physics Review Letter*, 91, 028701.

Munshi K. (2004). Social learning in a heterogeneous population: Technology diffusion in the Indian green revolution. *Journal of Development Economics*, 73(1), 185–213.

Neveu A. R. (2018). A survey of network-based analysis and systemic risk measurement. *Journal of Economic Interaction and Coordination*, 13(2), 241–281.

Newman, M.E.J. (2002). The spread of epidemic disease on networks. *Physical Review E*, 66, 016128.

Pastor-Satorras R., and Vespignani, A. (2000). Epidemic Spreading in Scale-Free Networks. *Physical Review Letters*, 86(14), 3200–3203.

Pastor-Satorras R., and Vespignani, A. (2001). Epidemic dynamics and endemic states in complex networks. *Physical Review E*, 63, 066117.

Pei, S., Teng X., Shaman J., Morone F., and Makse, H. A. (2017). Efficient collective influence maximization in cascading processes with first-order transitions. *Scientific reports* 7:45240, 1–13.

Pei, S., Morone F., and Makse H. A. (2018). Theories for influencer identification in complex networks. Lehmann, S., Ahn, Y.-Y. (eds.) Complex spreading phenomena in social systems. *Springer*, 125–148.

Piraveenan, M., Prokopenko, M., and Hossain, L. (2013). Percolation Centrality: Quantifying Graph-Theoretic Impact of Nodes during Percolation in Networks. *PLoS ONE*, 8(1), e53095.

Restrepo, J. G., Ott, E., and Hunt, B. R. (2006). Characterizing the dynamical importance of network nodes and links. *Physical Review Letter*, 97, 094102.

Rogers E. (1995). Diffusion of Innovations. *Free press.*

Saberi, M., Hamedmoghadam, H., Ashfaq, M., Hosseini, S.A., Gu, Z., Shafiei, S., Nair, D.J., Dixit, V., Gardner, L., Waller, S.T., and González, M.C. (2020). A simple contagion process describes spreading of traffic jams in urban networks. *Nature communications E*, 11(9), 1–9.

State, B., and Adamic, L. (2015). The diffusion of support in an online social movement: evidence from the adoption of equal-sign profile pictures. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 1741–1750.

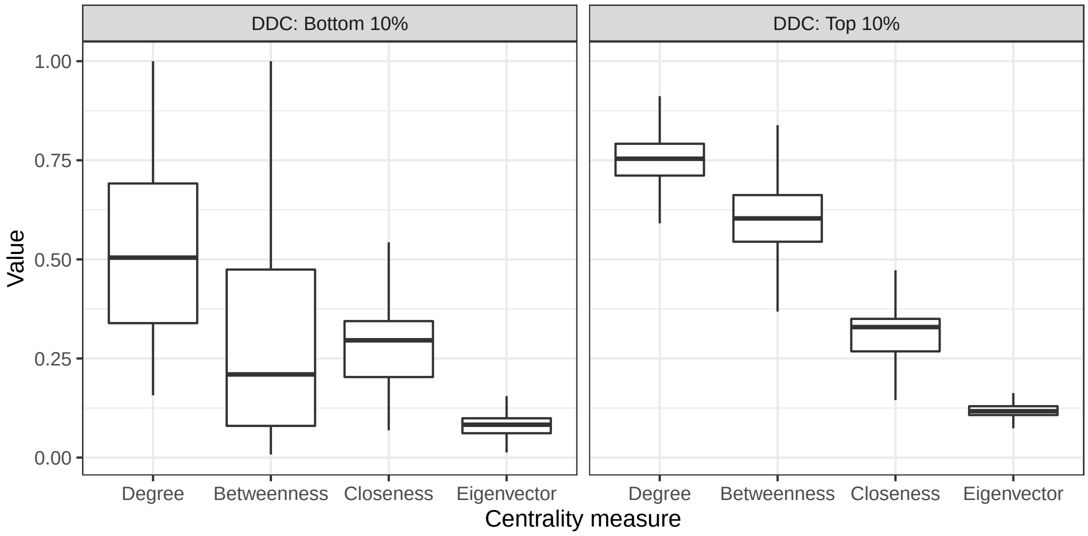
Teng, X., Pei, S., Morone, F., and Makse H. A. (2016). Collective influence of multiple spreaders evaluated by tracing real information flow in large-scale social networks. *Scientific reports* 6(1), 1–11.

Wasserman S, and Faust K. (1994). Social network analysis: Methods and applications. *Cambridge university press.*

A Topological features of low and high DDC nodes

In this section, we provide additional insights on the topological features of low and high DDC nodes. This exercise is conducted using the same simulated networks used in the Monte Carlo experiment presented in section 3.1 of the paper. Specifically, for each simulated network, we create two groups: i.e. the nodes featuring the top 10% and bottom 10% of DDC values, respectively. Then, we compute the average value of degree, betweenness, closeness, and eigenvector centrality for each group of nodes. Finally, we use a boxplot to observe the average and standard deviation of each of these metrics across all simulated networks. We report these values in Figure 11. For visualization purposes, each measure is rescaled between zero and one. A number of interesting insights can be gathered by observing this figure. First, high DDC nodes feature a high degree and betweenness centrality. Moreover, the small standard deviation of these two metrics, as indicated by the small scale of the boxplots, suggests that this is true for a large part of simulated networks. We thus infer that high DDC nodes tend to play the role of important hubs (i.e. high betweenness) with many connections (i.e. high degree) in all simulated networks. Second, low DDC nodes do not feature any recognizable pattern when looking at their degree, betweenness, closeness and eigenvector centrality values. These metrics greatly vary across networks, as suggested by the large scale of their associated boxplots, meaning that the information provided by the bottom ranking of the DDC is substantially different from that returned by other centrality measures. This is relevant because it confirms that DDC provides a new and non-redundant information on the role of nodes. Third, on average, high DDC nodes feature a higher degree and betweenness centrality value compared to low DDC. This further confirms that high DDC nodes are a crucial hub vehiculating diffusion across many nodes.

Figure 11: Topological features of low and high DDC nodes



Note: In the left panel, on the y-axis, we report the distribution of degree, betweenness, closeness, and eigenvector centrality, indicated on the x-axis, for the bottom 10% DDC nodes in the simulated networks. In the right panel, on the y-axis, we report the distribution of degree, betweenness, closeness, and eigenvector centrality, indicated on the x-axis, for the top 10% of DDC nodes in the simulated networks.