Social Networks 49 (2017) 12–26

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

Geodesic based centrality: Unifying the local and the global

Filip Agneessens[a,*], Stephen P. Borgatti[b], Martin G. Everett[c]

[a] Surrey Business School, University of Surrey, GU2 7XH, Guildford, UK
[b] LINKS Center for Social Network Analysis, Gatton College of Business and Economics, University of Kentucky, Lexington, KY 40508, USA
[c] Mitchell Centre for Social Network Analysis, School of Social Sciences, University of Manchester, Arthur Lewis Building, Bridgeford Street, M13 9PL, Manchester, UK

ARTICLE INFO

Article history:
Received 8 September 2015
Received in revised form 18 August 2016
Accepted 29 September 2016

Keywords:
Degree centrality
Closeness centrality
Betweenness centrality
Diversity
Resource-richness
Node-level measures

ABSTRACT

A variety of node-level centrality measures, including purely structural measures (such as degree and closeness centrality) and measures incorporating characteristics of actors (such as the Blau's measure of heterogeneity) have been developed to measure a person's access to resources held by others. Each of these node-level measures can be placed on a continuum depending on whether they focus only on ego's direct contacts (e.g. degree centrality and Blau's measure of heterogeneity), or whether they also incorporate connections to others at longer distances in the network (e.g. closeness centrality or betweenness centrality). In this paper we propose generalized measures, where a tuning parameter δ regulates the relative impact of resources held by more close versus more distant others. We first show how, when a specific δ is chosen degree-centrality and reciprocal closeness centrality are two specific instances of this more general measure. We then demonstrate how a similar approach can be applied to node-level measures that incorporate attributes. When more or less weight is given to other nodes at longer distances with specific characteristics, a generalized measure of resource-richness and a generalized measure for diversity among one's connections can be obtained (following Blau's measure of heterogeneity). Finally, we show how this approach can also be applied to betweenness centrality to focus on more local (ego) betweenness or global (Freeman) betweenness. The importance of the choice of δ is illustrated on some classic network datasets.

© 2016 Published by Elsevier B.V.

1. Introduction

A fundamental perspective in social network research is that, through social ties, individuals get access to information (Sparrowe et al., 2001), social support (Vaux, 1988; Marsden, 1987), and other resources. The more social access to resources, the greater the person's social capital. Thus measures of centrality such as degree and closeness (Freeman, 1979) speak directly to this form of social capital. Other types of centrality, such as betweenness (Freeman, 1977), speak to other forms of social capital. For example, Burt (1992) eloquently makes the case for the advantages that accrue to those in a position to broker between others. Accordingly, nodes high in betweenness centrality have the potential to control, filter or alter the flow of resources between parts of the network (Brass, 1984).

A wide variety of measures of centrality have been proposed (see Borgatti and Everett (2006) for reviews).[1] Some measures – such as degree centrality – focus on the local position in the network, while other measures – such as closeness centrality – also incorporate indirect connections. Hence, one important aspect to consider when choosing a measure for a specific research context (e.g. Marsden, 1990; Borgatti et al., 1998; Borgatti, 2005; Borgatti and Everett, 2006) relates to the question to what extent these flow processes are largely restricted to direct contacts or can travel through intermediaries to nodes further away. In the first case, the preferred measure should focus primarily or even exclusively on the flow of resources between people who are directly connected to ego. In the second case, the measure should (to some extent) also incorporate access to others that are indirectly reachable.

* Corresponding author.
E-mail addresses: f.agneessens@surrey.ac.uk (F. Agneessens), sborgatti@uky.edu (S.P. Borgatti), martin.everett@manchester.ac.uk (M.G. Everett).

http://dx.doi.org/10.1016/j.socnet.2016.09.005
0378-8733/© 2016 Published by Elsevier B.V.

[1] Moreover, centrality has been related with a person's ability to *influence* the views and behavior of others more easily (Brass and Burkhardt, 1993; Brass, 1984; Friedkin, 1991). However such measures often incorporate dependent effects (A influences B, and B subsequently influences A). We focus here purely on measures that focus on the geodesic distance, and thus will not focus on measures that involve *eigenvectors* (cf. Bonacich, 1972, 1987; Borgatti, 2005).

In this paper we propose a generalized approach, where a tuning parameter δ regulates the relative impact of resources held by more proximal versus more distal others. We show that two well-known centrality measures (*degree centrality* and *reciprocal closeness centrality*) are obtained as special cases of our measure by choosing δ appropriately. We also discuss how the choice of δ can be made *a priori* or determined empirically, based on maximizing the correlation with an outcome variable. We then show how our approach can be applied to other well-known measures of position. We illustrate this by focusing on three well known measures. First, focusing on betweenness centrality we yield a family of measures that vary along the local to global continuum.

In addition, following Everett and Borgatti (2012), we extend our approach to take into account the attributes of alters (cf. Agneessens and Koskinen, 2016). For example, numerous authors have constructed social capital measures by assessing the resource-richness of a node's direct contacts, as in looking at the total potential funding available to an entrepreneur through her rich contacts (e.g., Lin et al., 1981a,b; Hurlbert, 1991; cf. Lin, 2001). We generalize these types of measure by adding the δ parameter to take into account both direct and indirect access to resources. We similarly generalize measures of ego net heterogeneity by taking into account indirect contacts. The δ parameter allows us to tune the extent that contacts several links away impact the calculation of heterogeneity.

2. A generalization of measures of centrality

From a theoretical point of view, a large part of the social network literature takes for granted that indirect relations are relevant for obtaining resources (Borgatti and Halgin, 2011). For example, Granovetter has argued for the importance of weak ties in the transferring information about a potential new job among people at longer distances (Granovetter, 1973). Indirect relations are particularly applicable when considering easily transferable information and other types of resources or influence processes which require little energy from the persons involved (cf. Borgatti, 2005). In such cases, it makes sense to use centrality measures that incorporate both direct and indirect connections.

In other cases, however, contacts at longer distance might have less impact. For example, transferring complex, tacit procedural knowledge ("how to" knowledge) requires considerable effort and time (Hansen, 1999), and may not flow across long distances. Similarly, a considerable proportion of research on social support has focused almost exclusively on the importance of direct ties to others (e.g., Vaux, 1988; Thoits, 1982). Since most social support (and in particular emotional support) requires substantial investment from others and frequent contact, social support studies often use measures that exclusively focuses on the amount of support directly accessible to a person to predict well-being and buffer stress (e.g. Thoits, 1982; Lin, 1986; Marsden, 1987). This is largely because the contribution of persons indirectly reachable through third parties can be neglected and therefore the focus is on personal networks (e.g., Fischer, 1982; Campbell and Lee, 1991).

Hence, when deciding which measure to use to capture a person's access or influence in a specific context, one should question whether the flow-processes for the specific research is largely restricted to directly connected others, or whether it can also involve others who are indirectly reachable. The centrality measure used should reflect this decision.

2.1. Closeness centrality

In general, closeness centrality is based on the distance of a person to all others in the network. Closeness centrality is for example used to consider the time it would take for a person to access resources which are distributed over the actors in a network (Brass, 1984), or alternatively the total number of steps needed for a person to access everyone in a network (cf. the key player problem discussed in Borgatti, 2006). Distance is generally based on the *shortest distance* between ego and all its others (i.e., the geodesic distance, g_{ij}).

In the classic version of closeness centrality, shorter distances between two actors (high closeness) are preferred over longer distances as a shorter distance might imply "fewer message transmissions, shorter times and lower costs" (Freeman, 1979: 225). Hence, for closeness centrality, others close by are weighted as more important (i.e., more easy to access or more easy to influence) than those at longer distance, because they involve less intermediary third parties. While people at shorter distances are more important, longer distances may still matter.

One version of closeness centrality is reciprocal closeness (cf. Borgatti, 2006) which uses reciprocal distance (Latora and Marchiori, 2001; Newman, 2003). Reciprocal closeness centrality is based on the sum of the closeness from an actor i to all other actors j, where the closeness between two actors i and j is obtained by taking the geodesic distance from actor i to all others j (g_{ij}) and transforming this distance into a closeness score by taking its reciprocal ($1/g_{ij}$).[2] As a result, the closeness between i and j ($1/g_{ij}$) is 1 if the geodesic distance (g_{ij}) is equal to 1 (a geodesic distance of 1 being the minimal distance between two nodes), and the value for $1/g_{ij}$ becomes smaller as the shortest path from i to j is longer. In the limit, the closeness-score goes to 0 as distance goes to infinity. Conveniently, we can use this to deal with disconnected graphs by declaring the distance between nodes in different components to be infinite, thus yielding a closeness score of zero (cf. Latora and Marchiori, 2001; Newman, 2003). The resulting normalized version of reciprocal closeness centrality is:

$$C'_C(i) = \frac{\sum_{j \neq i} \left(g_{ij}^{-1}\right)}{N-1} \quad (1)$$

This normalization involves dividing by the total number of others (*N-1*, the total number of nodes in the network besides ego), as the maximum closeness a person can reach is obtained when he or she is directly connected to all (*N-1*) other actors (i.e. a geodesic distance of 1, so that $1/g_{ij}$ equals 1), and the sum over all others is then *(N-1)*.[3] The normalized measure is also known as average reciprocal distance or ARD (Borgatti et al., 2002).

2.2. Generalized measure of centrality based on closeness

The measure described by Eq. (1) assumes that a node at distance 4 from ego contributes only 1/4 while a node at distance 2 contributes 1/2. As discussed before, the importance of connections at longer distances might be high or low, depending on the type of process, and therefore different weight should be allowed to actors at different geodesic distances depending on the relevance of different distances. Following this logic, we propose a generalization of Eq. (1) that weights nodes at different distances depending on the

[2] While we focus in this paper on undirected networks, one can easily extent this to either outgoing ties from actor i to all other actors, or incoming ties from all others to actor i.

[3] An alternative approach, *Freeman's closeness* centrality (Freeman, 1979), takes the sum of the number of steps needed to reach all other actors. It subsequently takes 1 over this farness and normalizes the measure by multiplying by N-1: $C'_C(i) = \frac{N-1}{\sum_{j \neq i} g_{ij}}$. However, whenever at least one actor is not reachable from another actor in the network, farness is more difficult to define, and as Freeman (1979: 225) states the measure above generally only makes sense for a connected graph – i.e. a network which consists of a single component.

value of a gradient parameter δ. More specifically, g_{ij}^{-1} is replaced by $g_{ij}^{-\delta}$, as follows:

$$C'^{\delta}_{C}(i) = \frac{\sum_{j \neq i} g_{ij}^{-\delta}}{N-1}, \text{ where } \quad \delta \geq 0 \tag{2}$$

Note that both degree and reciprocal closeness are special cases of this measure, as they can be obtained by setting δ to infinity or 1, respectively. In addition, we can choose other values of δ. For example, we can choose values between infinity and 1 to construct measures intermediate between degree and closeness.[4] The greater the δ, the more the measure favors short distances, and therefore the greater the similarity to degree. On the other hand, we can choose values below 1 (and close to 0) to reflect a situation where nodes at all distances are (almost) weighted equally. It should be noted that for all positive δs, the maximum value of the numerator cannot be greater than N-1.

The choice of the gradient δ could follow from a theoretical idea about the relative impact of the longer distances, compared to shorter distances. Alternatively one might be interested in the centrality of a person when different δs are considered and empirically try to find the δ that optimizes a specific outcome and hence provide an empirical answer to this question. In what follows we will use an example to consider the impact of different values for δ. In Section 2.4 we then provide an example where we search for the optimal value for the gradient δ. We will also provide theoretical criteria that might guide us when we want to make an *a priori* specific choice of δ.

2.3. Example

To illustrate the impact of different choices of δ, we use the longitudinal friendship network among 32 students collected by van de Bunt (Van de Bunt et al., 1999). We focus on the fourth wave and we consider two students to be friends if either (or both) of them named the other as a friend or best friend. A graphical representation of the resulting friendship network can be found in Fig. 1. Table 1 shows the centrality scores for each student, given different values for δ.

δ equal to 1. When δ is 1, the connection between two actors is given by $1/g_{ij}{}^{1}$, and therefore the results are exactly the same as those for the reciprocal closeness (Equation 1). In that case student 27 is the most central (with a value of 0.470), and student 3 is a close second (with a value of 0.469), while student 24 follows as 3rd (0.430), and student 10 as 4th (0.415). As can be seen more clearly in Fig. 2, student 27 has a high number of direct contacts (8) as well as a good proportion of people at distance 2 and 3. Student 3 has a good amount of direct contacts (5) together with a high level of close-by indirect contacts (13 at distance 2).

High positive δ values. When we increase the value for δ to 2, 3 or higher, people at a longer geodesic distance become less important, since their value for $1/g_{ij}{}^{\delta}$ becomes relatively small. For example, when δ is set to 2, a geodesic distance of 3 generates a value $3^2 = 9$, and therefore the contribution of the person at distance 3 is now 1/9, compared to a contribution of 1/3 when δ is 1 for the same geodesic distance. In the limit as δ increases further, any node who is indirectly connected to ego will generate a value for $g_{ij}{}^{\delta}$ that tends to infinity, so that $1/g_{ij}{}^{\delta}$ approximates 0. As a result only direct relations (with geodesic 1) still have an impact on ego. Hence, degree

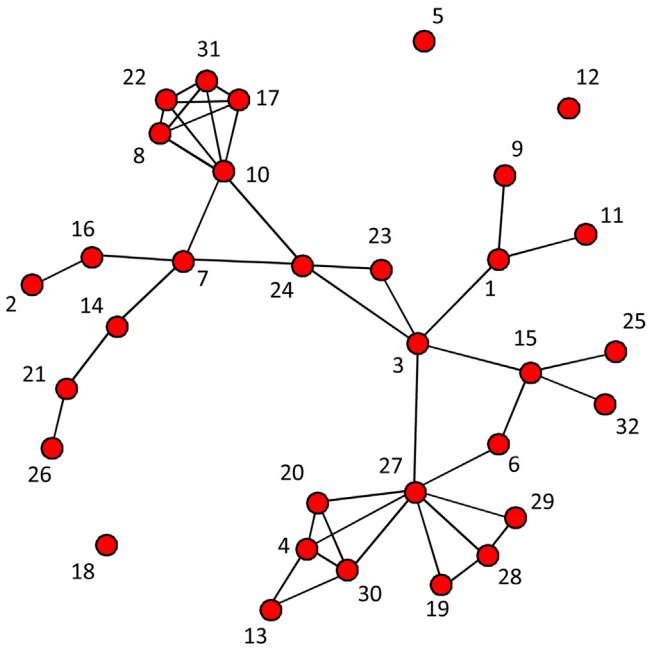


Fig. 1. Symmetrized friendship network among 32 Dutch students (Van de Bunt, 1999).

centrality[5] can be seen as a special case of the general formula with a high δ (e.g. 100), resulting in Eq. (3):

$$C'_{D}(i) = \frac{\sum_{j \neq i} x_{ij}}{N-1} \tag{3}$$

In the example dataset, the centrality of student 10 increases from 4th to 3rd place from the moment that δ is 2 (0.269), overtaking student 24 (with a value of 0.252). Student 27 remains most central (0.335) followed by student 3 (0.296). Student 10 even gets to a 2nd place when δ is 5 or higher, leaving student 3 at place 3 (0.199 compared to 0.175 for student 3). Clearly, compared to students 24 and 3, student 10 benefits more from its direct connections and has less (close) indirect connections, while students 3 and 24 have a lower number of direct connections, but have a high level of people two and three steps away (see their respective degree distributions in Fig. 2). In all these cases student 27 remains the most central, benefiting especially from a high number of direct ties and people at distance 2. When δ goes to infinity student 27 gets a value 0.258 (8/31), student 10 gets a value 0.194 (6/31), student 3 gets a value 0.161 (5/31) and student 24 gets a value 0.129 (4/31), similar to a number of other nodes in the network. This ordering matches the ordering based on node degree (see Fig. 2).

Small positive δ values. On the other hand, as the value for δ approaches 0, the importance of the differences in geodesic distances to other nodes become far less important, provided the person can at least be reached indirectly. When a δ close to 0 is used, the contribution of all directly and indirectly connected others will be relatively similar, irrespective of their geodesic distance from ego, as long as the node can be reached at all. For δ close to 0, the value $g_{ij}{}^{\delta}$ will be close to 1, irrespective of the distance, and therefore $1/g_{ij}{}^{\delta}$ will also be close to 1. The only exception is a geodesic distance equal to infinity, which by definition we assign a value of 0.

[4] Note that many other transformations are possible. However, we focus on this transformation because it generates degree and reciprocal closeness as two outcomes, which are established measures.

[5] Degree centrality can be defined as the number of other nodes a person is directly connected with, i.e. the number of nodes reached through a single step and without the need of any intermediary.

Table 1
Results of generalized closeness centrality with different δ's (van de Bunt friendship data) (normalized values reported).

	δ													
	-10	-2	-1	-0.5	-0.1	0	0.1	0.25	0.5	1	2	3	5	100
1	0.000000032	0.030	0.177	0.413	0.778	0.903	0.814	0.698	0.548	0.354	0.186	0.130	0.103	0.097
2	0.000000363	0.066	0.268	0.511	0.813	0.903	0.779	0.625	0.439	0.227	0.081	0.046	0.034	0.032
3	0.000000004	0.017	0.131	0.355	0.755	0.903	0.838	0.753	0.635	0.469	0.296	0.223	0.175	0.161
4	0.000000171	0.041	0.200	0.433	0.784	0.903	0.809	0.690	0.539	0.355	0.206	0.159	0.135	0.129
5	0.000000000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
6	0.000000170	0.039	0.196	0.430	0.784	0.903	0.809	0.688	0.534	0.339	0.170	0.110	0.074	0.065
7	0.000000008	0.026	0.163	0.395	0.771	0.903	0.822	0.716	0.578	0.396	0.231	0.172	0.138	0.129
8	0.000000067	0.045	0.214	0.451	0.791	0.903	0.802	0.675	0.516	0.327	0.186	0.148	0.132	0.129
9	0.000000176	0.050	0.232	0.476	0.801	0.903	0.790	0.647	0.469	0.257	0.096	0.053	0.035	0.032
10	0.000000008	0.027	0.165	0.396	0.770	0.903	0.823	0.720	0.586	0.415	0.269	0.222	0.199	0.194
11	0.000000176	0.050	0.232	0.476	0.801	0.903	0.790	0.647	0.469	0.257	0.096	0.053	0.035	0.032
12	0.000000000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
13	0.000000753	0.062	0.252	0.491	0.805	0.903	0.787	0.645	0.469	0.267	0.122	0.083	0.067	0.065
14	0.000000063	0.042	0.210	0.449	0.791	0.903	0.801	0.671	0.506	0.305	0.143	0.094	0.070	0.065
15	0.000000032	0.030	0.173	0.407	0.775	0.903	0.817	0.706	0.561	0.376	0.215	0.161	0.135	0.129
16	0.000000063	0.043	0.214	0.454	0.793	0.903	0.798	0.667	0.500	0.297	0.137	0.090	0.069	0.065
17	0.000000067	0.045	0.214	0.451	0.791	0.903	0.802	0.675	0.516	0.327	0.186	0.148	0.132	0.129
18	0.000000000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
19	0.000000171	0.042	0.206	0.443	0.789	0.903	0.803	0.677	0.516	0.317	0.153	0.100	0.072	0.065
20	0.000000171	0.041	0.202	0.437	0.786	0.903	0.807	0.685	0.530	0.339	0.182	0.131	0.104	0.097
21	0.000000359	0.064	0.260	0.501	0.809	0.903	0.784	0.636	0.456	0.253	0.112	0.078	0.066	0.065
22	0.000000067	0.045	0.214	0.451	0.791	0.903	0.802	0.675	0.516	0.327	0.186	0.148	0.132	0.129
23	0.000000005	0.024	0.161	0.397	0.773	0.903	0.818	0.707	0.558	0.359	0.173	0.107	0.072	0.065
24	0.000000001	0.018	0.139	0.368	0.761	0.903	0.832	0.737	0.609	0.430	0.252	0.182	0.140	0.129
25	0.000000176	0.048	0.228	0.471	0.800	0.903	0.792	0.652	0.476	0.265	0.103	0.056	0.036	0.032
26	0.000001617	0.091	0.315	0.553	0.826	0.903	0.767	0.603	0.408	0.199	0.069	0.042	0.034	0.032
27	0.000000032	0.025	0.153	0.377	0.762	0.903	0.832	0.741	0.623	0.470	0.335	0.289	0.264	0.258
28	0.000000171	0.042	0.204	0.439	0.787	0.903	0.805	0.682	0.526	0.334	0.177	0.128	0.103	0.097
29	0.000000171	0.042	0.206	0.443	0.789	0.903	0.803	0.677	0.516	0.317	0.153	0.100	0.072	0.065
30	0.000000171	0.041	0.200	0.433	0.784	0.903	0.809	0.690	0.539	0.355	0.206	0.159	0.135	0.129
31	0.000000067	0.045	0.214	0.451	0.791	0.903	0.802	0.675	0.516	0.327	0.186	0.148	0.132	0.129
32	0.000000176	0.048	0.228	0.471	0.800	0.903	0.792	0.652	0.476	0.265	0.103	0.056	0.036	0.032

In our data, as more importance is given to a longer distances (i.e., the δ is positive, but close to 0), the centrality of student 3 overtakes that of student 27. For example when δ is set to 0.25, student 3 obtains a value of 0.753 compared to only 0.741 for student 27 (with student 24 approaching very closely in third position with 0.737). Hence, when more (equal) importance is given to both close and far away connections, student 3 seems to reach more others rather than 27. A close look at the picture reveals that student 27 occupies a more peripheral position than 3, and benefits from a larger number of close connections (cf. Fig. 2). When δ is set to 0.25, the 5 direct connections for student 3, compared to 8 direct connections for student 27, is outweighed by a higher number of close yet indirect connections (especially the 13 nodes that are 2 steps away from student 3).

δ set to 0. In the extreme case – when δ is set to 0–the importance of other nodes will be equal as long as the person can be reached by some path, as $1/g_{ij}^{\delta}$ will be 1 whatever the distance. When a node j is not reachable, the geodesic distance g_{ij} can be declared infinite, and in this case $1/g_{ij}^{\delta}$ is assigned the value 0. Hence the sum reflects the number of nodes that ego can reach by some path, no matter how long it is. Equivalently, the sum equals the number of other nodes in ego's component. For a strongly connected graph, the maximum value is *N–1*.

In the example, when δ further approaches 0, all actors in the largest component become equally well positioned, while the three isolated students (students 5, 12 and 18), who are not part of the main component, are in a worse situation since they have no access to resources from the main component.

Negative δ. It is also interesting to note that, when a negative δ is chosen, connections at a longer distance receive a higher value than others at shorter distances and hence further away (but still connected) nodes are preferred over shorter distance others.[6] Such a choice of δ could make sense when it is important to be (indirectly) connected to as many people as possible, but at the same time to keep them as much as possible at a long distance. For example, in the case of a criminal network, to avoid being uncovered, it might be important for the "leader" of a criminal organisation not to be directly associated with other important criminals, but to nevertheless be connected at some distance (e.g., Duijn et al., 2014). Similarly, when direct connections require a lot of social investment (e.g., with respect to time or returning of favors), it may be advantageous to have just a few direct connections, but to nodes that ultimately connect ego indirectly to a large number of distant connections.

Note that when δ is negative, the maximum value occurs when distances from ego are as large as possible. This occurs when the entire network is arranged as a path, with ego at one endpoint. As a result, the maximum score for δ = –1 is *N(N-1)/2*. As the negative δ increases in magnitude, the maximum score continues to increase. In general, for any negative δ, the maximum is:

$$\text{Max} = \sum_{k} k^{-\delta} \text{where } k \text{ goes from 1 to } N-1 \quad (4)$$

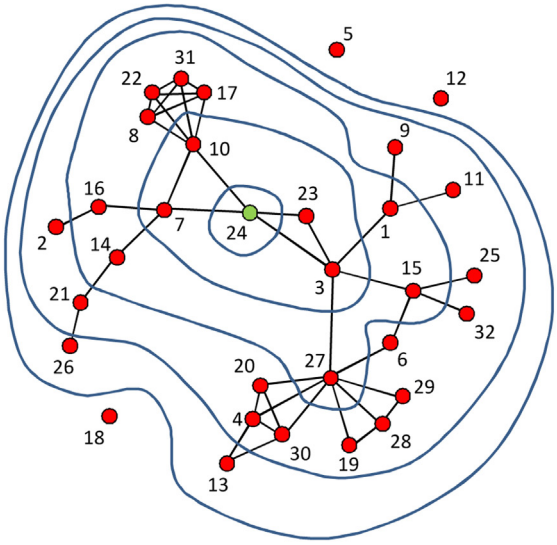
We can use this maximum to normalize the measure for negative δs:

$$C'^{\delta}_{C}(i) = \frac{\sum_{j \neq i} g_{ij}^{-\delta}}{\sum_{k} k^{-\delta}} \text{where} \quad \delta < 0 \quad (5)$$

In the example, student 26 is clearly the furthest away from all others (but still connected), hence generating the highest centrality

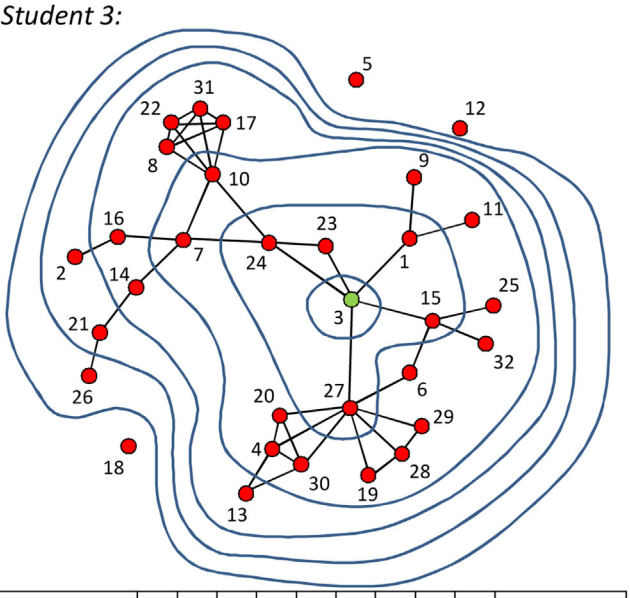
[6] Again we have to assume that the contribution of an actor is 0 if the distance to that actor is infinity.

Student 24:



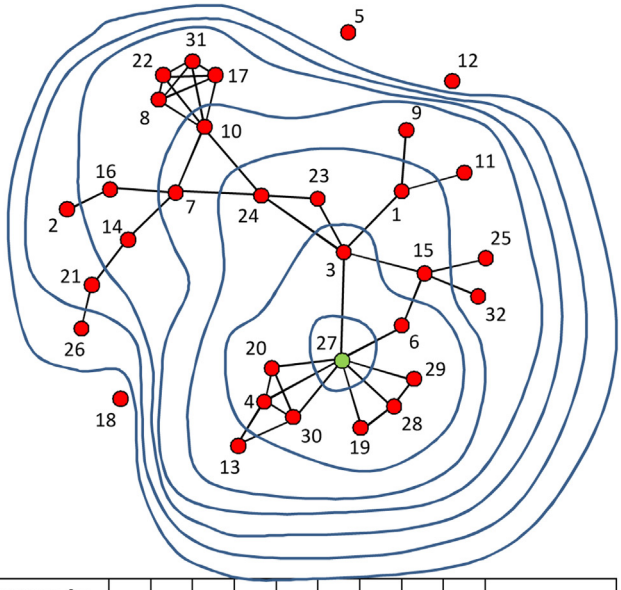
Distance to student 24	1	2	3	4	5	6	7	8	9	Infinity
Number of nodes at distance	4	9	13	2	0	0	0	0	0	3

Student 3:



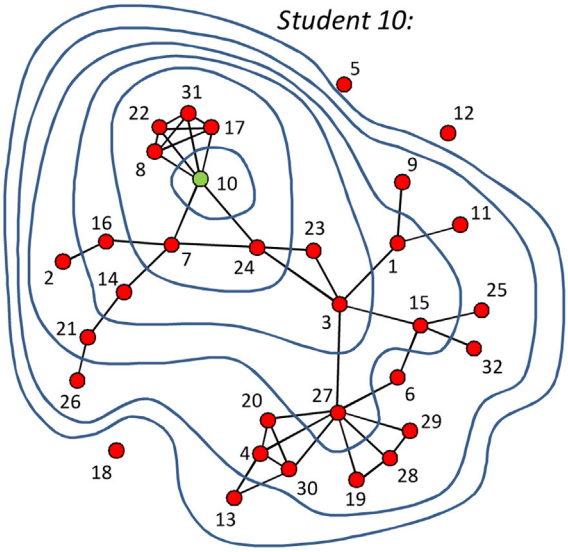
Distance to student 3	1	2	3	4	5	6	7	8	9	Infinity
Number of nodes at distance	5	13	7	2	1	0	0	0	0	3

Student 27:



Distance to student 27	1	2	3	4	5	6	7	8	9	Infinity
Number of nodes at distance	8	5	6	6	2	1	0	0	0	3

Student 10:

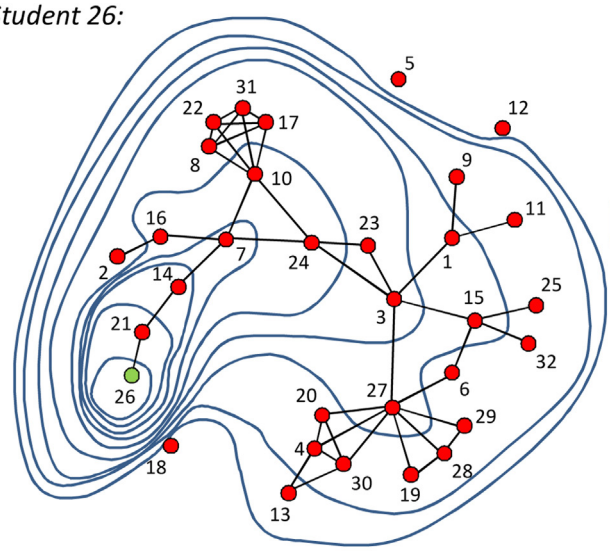


Distance to student 10	1	2	3	4	5	6	7	8	9	Infinity
Number of nodes at distance	6	4	5	12	1	0	0	0	0	3

Fig. 2. Degree distributions of the 4 most central students when δ is positive.

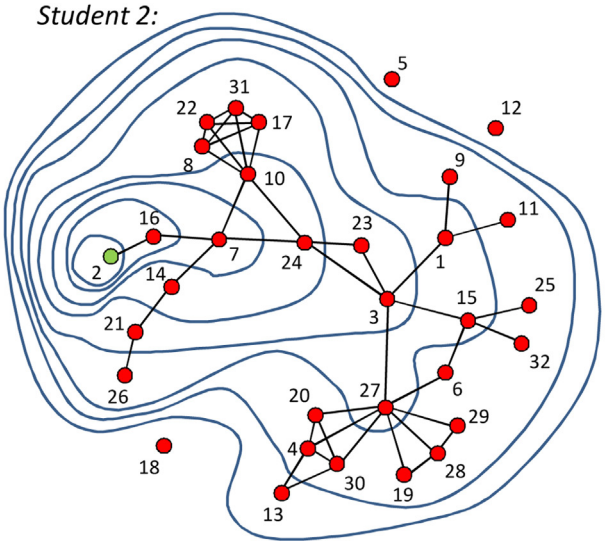
when δ is negative (see Fig. 3). However the second most important node changes depending on the value of δ. The positions of student 2 and 21 are preferred over student 13 when there is a moderate preference for nodes at longer distance (0.268 and 0.260 when δ is -1). However, when the importance of really long distances is given more importance, student 13 is more important than 2 or 21

Student 26:



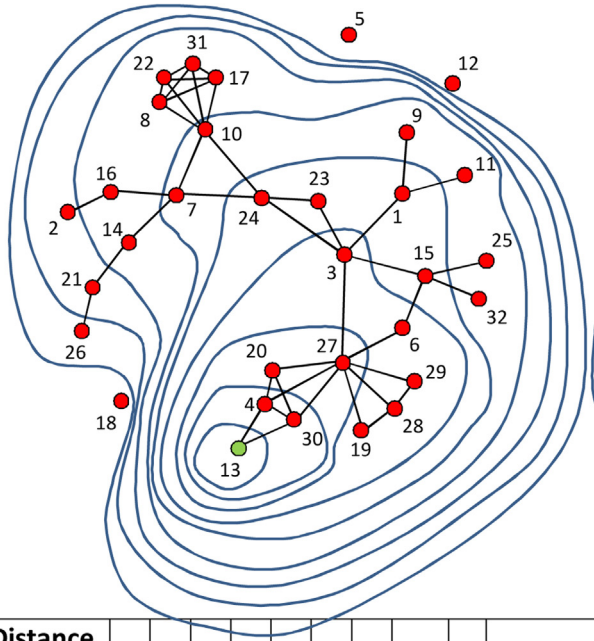
Distance to student 26	1	2	3	4	5	6	7	8	9	Infinity
Number of nodes at distance	1	1	1	1	3	7	3	11	1	3

Student 2:



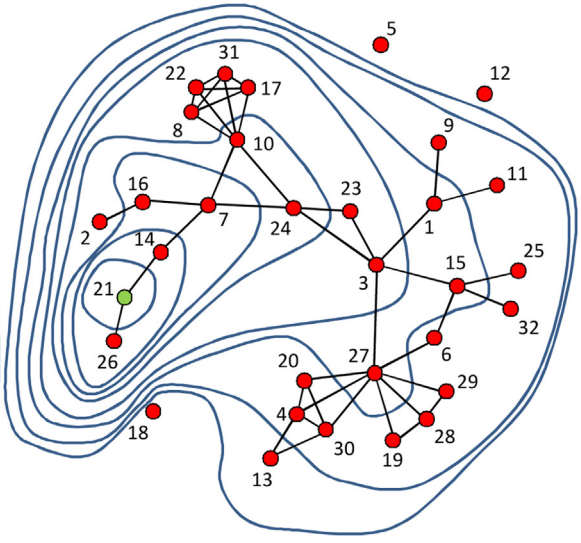
Distance to student 2	1	2	3	4	5	6	7	8	9	Infinity
Number of nodes at distance	1	1	3	7	4	11	1	0	0	3

Student 13:



Distance to student 13	1	2	3	4	5	6	7	8	9	Infinity
Number of nodes at distance	2	2	5	4	6	6	2	1	0	3

Student 21:



Distance to student 21	1	2	3	4	5	6	7	8	9	Infinity
Number of nodes at distance	2	2	3	6	3	11	1	0	0	3

Fig. 3. Degree distributions of the 4 most central students when δ is negative.

(.7.53 * 10^{-7} over 3.63 * 10^{-7} or 3.59 * 10^{-7} when δ is set to −10). This is largely due to two nodes at distance 7 (rather than 1) and 1 person at distance 8 for student 13. Fig. 3 gives an overview of these nodes degree distribution.

Centrality scores for different δs. To explore this further, we ordered the nodes according to their centrality score when δ is set to different values (see Fig. 4). The figure reiterates the point made before: the relative centrality of quite a number of nodes can change considerably in the network when a different choice for δ is made.

We also correlated the centrality scores for different δs (Table 2). Not surprisingly the centrality scores for nodes are highly correlated when similar δs are used. However, for rather different δs the correlation is considerably lower. For example, the correlations between the centrality scores when δ is set to 0.1 and when it is set to 5 is only 0.54. These results are in line with earlier research (e.g., Valente et al., 2008) that found that in many real networks different centrality measures (such as degree and closeness centrality) "are distinct, yet conceptually related" (p. 24). Moreover, when δ is set to a relatively high positive value (e.g., $\delta = 2$) versus a relatively high negative value (e.g., $\delta = -2$) we even obtain a negative correlation (−0.05). Both the correlations and the rank order results reiterate the importance of the choice of δ.

2.4. The choice of δ: empirical and theoretical approach

The previous section illustrates the impact of the choice of δ. We now focus on addressing the reasoning for choosing a specific δ. Two approaches can be taken: one is based on an *a priori* choice guided by theory, while the second is based on an empirical search for the δ that maximizes the correlation between the centrality measure and an outcome variable.

Using an *a priori* approach the focus is on identification of the questions that help guide us in the decision which δ to use. We identify at least three relevant criteria that help guide this theoretical choice: (1) the willingness of the intermediary actors to transmit the resource between two nodes, (2) the ability to forward the resource, and (3) the usefulness of the resources themselves for nodes at longer distance.

One first crucial criterion relates to the willingness (and motivational strength) of intermediary actors to transmit the resource further. As we discussed briefly before, indirect relations might be particularly relevant when considering resources and information with a low costs to transmit these on to others (such as where to get a new job (Granovetter, 1973)), as this would make it likely that the intermediary actor is willing to take on this intermediary role. For more complex (tacit) information and other resources which require considerable investment and costs to transmit to the next person (Hansen, 1999) the willingness of an intermediary is likely to be low and therefor the contribution of the indirect ties is likely to be limited. Hence, one of the crucial reasons guiding the choice of δ might be the energy that it requires from the person(s) in between the sender and the receiver.

However, there might be a second reason why indirect ties might have limited relevance. When considering complex resources, such as learning a new skill, it might simply be very difficult for an intermediary to transfer this complex information. Even when intermediary actors might be willing to make time and energy available to play the role of broker, the resource might be too difficult to be passed on through indirect relations. Hence, they might be (largely) limited to direct interaction, whereas for easily transferable knowledge the ability to forward is likely to be a lot higher. In a somewhat analogue way, a virus might be prevented from reaching longer distant contacts because the people in-between are immune to the virus (or at least less likely to develop it). Note that we require both willingness and ability from people for them to take on the role of intermediary.

Another important aspect is the relevance of a resource as it flows through the network and reaches people at longer distances. Gossip for example is generally not very useful for people at longer distances, since gossip is more interesting if it is about people one knows (e.g. Ellwardt et al., 2012). In a similar way, one can argue that emotional support is not transmittable to other people because it is very person-specific and therefore not relevant to those third parties. Emotional support is generally directly focused on the target and it is hard to provide emotional support through an intermediary (e.g., Vaux, 1988; Thoits, 1982). Resources are also potentially vaporous as they move through a network, especially when they are to be split over different others (cf. parallel duplication; Borgatti, 2005). On the other hand, innovation and fashion is relevant to longer distances and – provided the intermediary people are willing and able to transmit – will be impacting all its contacts at longer distances (serial duplication, Borgatti, 2005). Similarly, viruses can easily remain relevant at very long distances.

Hence, when the choice of the gradient δ is based on an *a priori* choice, δ should reflect answers to such questions as:

- How *willing* are intermediary actors to transmit resources (taking into account their limited energy and time)?
- How *able* are intermediary actors to transmit resources (taking into account their limited ability and the "complexity" of the resource)?
- How *relevant* is the resource *to longer distance others* in the network (taking into account the potential local "nature"/relevance of the resource)?

Alternatively, when there is no clear *a priori* choice, one can empirically "search" for the optimal choice for gradient δ based on how the centrality for specific values for δ "maximize" a specific objective. To illustrate this we consider two networks.

The first example is the classic marriage network among Florentine families (Padgett and Ansell, 1993, see Fig. 5), with the attribute wealth as a measure of the richness of a family. Suppose we are interested in knowing whether the centrality in this network is related to the wealth of those families.

Fig. 6a provides the explanatory power (R squared) of the variable wealth based on reciprocal closeness centrality with different values for the gradient δ. We removed the isolate from the analysis. Clearly in this case the maximal explanatory power is obtained when δ is high (R^2 = 25% when $\delta > 5$). This indicates that the relation between the position of the family in the marriage network and their wealth is largely based on the direct number of marriage ties (degree), and that the indirect contacts can be neglected.

In Fig. 6b we consider the EIES friendship network (Freeman and Freeman, 1979). The network data consist of friends or good friends, symmetrized, among early network researchers. In addition, we have the number of citations amassed by each researcher. We used the natural log transformation (log (x + 1)) due to the skewness of the citations variable. Again we removed any isolates. There seems to be a clear relationship between the (log) number of citations and centrality when using a gradient value for δ that is approximately 0.65. This is much smaller than in the Padgett dataset, where the optimal δ was above 5. In contrast to the Padgett case, the results here indicate that the outcome variable is best predicted by a measure that places considerable weight on distant contacts (even more than the reciprocal closeness, when $\delta = 1$). Comparing the results in the two datasets, we note that searching for the optimal δ has a nice benefit of telling us something about the processes being studied in each. That is, we can use the empirically optimal δ as an indicator of the relative importance of long paths for a given process. For example, we previously speculated that gossip was something that

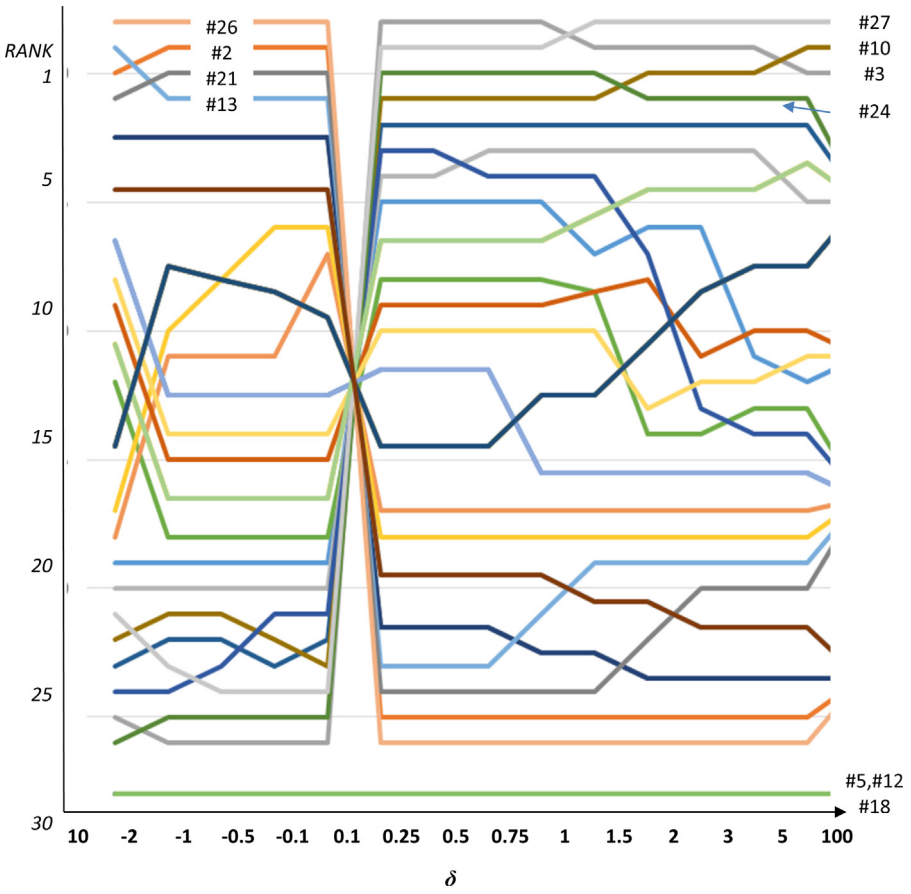


Fig. 4. Rank order for generalized closeness for 32 students (based on results in Table 1).

Table 2
Correlations between generalized closeness centrality with different δ's for main component (in Table 1).

	-10	-2	-1	-0.5	-0.1	0.1	0.25	0.5	1	2	3	5	100
-10	-	0.74	0.55	0.39	0.23	0.14	0.08	-0.02	-0.17	-0.28	-0.29	-0.27	-0.27
-2	0.74	-	0.95	0.84	0.70	0.60	0.53	0.40	0.17	-0.05	-0.10	-0.10	-0.10
-1	0.55	0.95	-	0.97	0.88	0.82	0.76	0.65	0.44	0.20	0.12	0.10	0.10
-0.5	0.39	0.84	0.97	-	0.97	0.93	0.89	0.81	0.64	0.41	0.32	0.28	0.27
-0.1	0.23	0.70	0.88	0.97	-	0.99	0.97	0.93	0.80	0.60	0.50	0.45	0.44
0.1	0.14	0.60	0.82	0.93	0.99	-	1.00	0.97	0.87	0.69	0.59	0.54	0.52
0.25	0.08	0.53	0.76	0.89	0.97	1.00	-	0.99	0.91	0.75	0.66	0.60	0.59
0.5	-0.02	0.40	0.65	0.81	0.93	0.97	0.99	-	0.97	0.84	0.76	0.70	0.68
1	-0.17	0.17	0.44	0.64	0.80	0.87	0.91	0.97	-	0.95	0.89	0.83	0.81
2	-0.28	-0.05	0.20	0.41	0.60	0.69	0.75	0.84	0.95	-	0.99	0.96	0.94
3	-0.29	-0.10	0.12	0.32	0.50	0.59	0.66	0.76	0.89	0.99	-	0.99	0.98
5	-0.27	-0.10	0.10	0.28	0.45	0.54	0.60	0.70	0.83	0.96	0.99	-	1.00
100	-0.27	-0.10	0.10	0.27	0.44	0.52	0.59	0.68	0.81	0.94	0.98	1.00	-

-1.00 - .00 **.00 - .29** **.30 - .59** **60 - .79** **. 80 - .94** **.95 - 1.00**

Note: When $\delta = 0$, all nodes in the main component have the same value and therefore was excluded from the table above.

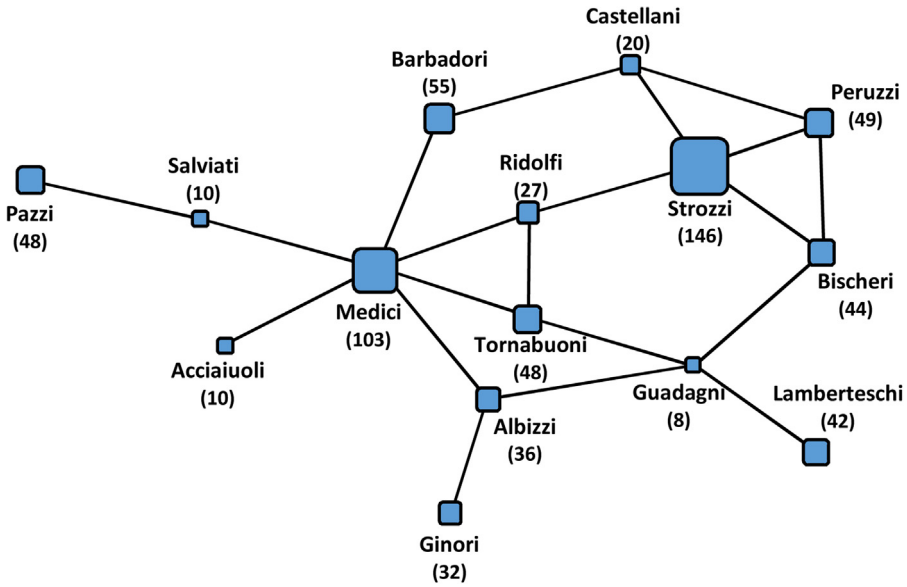


Fig. 5. Marriage network among 16 Florentine families (values reflect wealth).

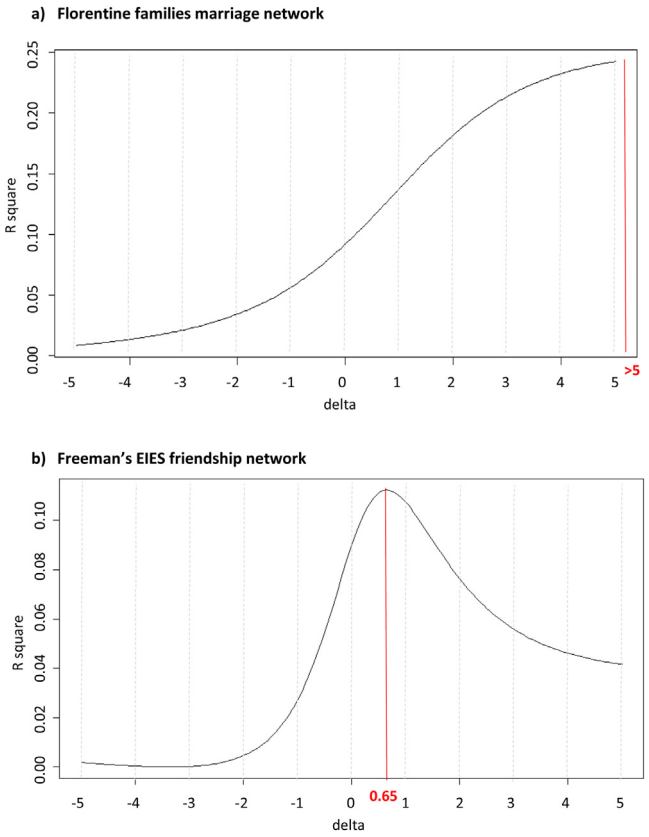


Fig. 6. Predictive power under different choice of δs.
a) Florentine families' centrality (marriage network) on wealth (excluding isolates)
b) Centrality in Freeman's EIES network (time 1 friendship symmetrized) on natural log of number of publications plus 1 (excluding isolates)

did not travel far. Hence, across multiple gossip datasets, we should find that, on average, it is the higher δs that will be associated with outcomes that depend on gossip.

3. Extending the generalization to measures with attributes

Besides purely structural measures – such as degree and closeness centrality – other node-level measures (cf. Agneessens and Koskinen, 2016) have incorporated characteristics of ego's connections in order to reflect the idea that some of these connections might be more relevant for ego than others (even when they are at a similar distance from ego).

In this part we show how the approach used before can easily be extended to measures that incorporate attributes. While this can be applied to a range of measures, to illustrate how this works, we will focus on two major types of measures that incorporate characteristics of the other nodes in the network: *resource-richness* and *diversity*. Generally the measures developed to capture the resource-richness (e.g., Lin et al., 1981a,b; Hurlbert, 1991) and diversity of ego's surrounding (Burt, 1983; Marsden, 1987; Reagans and McEvily, 2003; Campbell et al., 1986) focus exclusively on direct contacts. However, following the logic proposed before, one can easily think of more global variants of such measures, where the measure not only takes into account the direct contacts, but also those at longer distance, and where these contacts are weighted in one way or another by the amount of steps needed to reach a person (the distance) (cf. Everett and Borgatti, 2012). We will first focus on measures that explicitly focus on the *resource-richness* of the other nodes, by incorporating attributes that reflect differences in resources, power or status of these nodes (Lin, 1999). We will then discuss measures that incorporate attributes to reflect the *diversity* or range in characteristics of the nodes around ego, as this diversity might indicate access to (more) unique properties or resources that are different from the resources of the other existing connections of ego (cf. Granovetter, 1973).

3.1. Resource-richness

A central argument in Social Resource Theory is that connections are of little relevance if these others do not possess resources that are useful for ego (Lin, 1999). In other words, being connected to people with more money, more power or higher status is better than being connected with people with less money, less power or low status. This approach focuses on incorporating the

resource-richness of the others explicitly by measuring the extent to which others around an actor possess more or less *useful* characteristics, resources, power or status (e.g., Lin et al., 1981a,b; Brass, 1984; Hurlbert, 1991; cf. Lin, 2001; van der Gaag and Snijders, 2005; O'Connor, 2013). For example, Hurlbert (1991) investigated whether having direct access to highly educated persons makes people more satisfied with their life, than being linked to lower educated others. Similarly, Lin et al. (1981a,b) focused on the occupational prestige of one's connections to explain a person obtaining a high status job.

Most studies that explicitly incorporate the resource-richness of others (using a specific "resource-richness"-scale) focus on a local variant (Hurlbert, 1991; Lin et al., 1981a,b), i.e. where ego is directly connected to these other nodes. In the *resourceful degree* measure the impact of the other nodes (j) is weighted by the specific resources (r_j) that these nodes possess:

$$C'_R(i) = \frac{\sum_{j \neq i}(x_{ij} r_j)}{\sum_{j \neq i} r_j} \tag{6}$$

This measure is then normalized by dividing by the total amount of resources that all the others in the network have available, and hence could be reached if ego would be directly connected to all the other nodes.[7] However, since indirect relations might be important, resources from indirect connections should potentially also be taken into account. To incorporate the resources of actors at longer geodesic distance, but weighting by their distance, we can use the following formula:

$$C'^{\delta}_R(i) = \frac{\sum_{j \neq i}({g_{ij}}^{-\delta} \cdot r_j)}{\sum_{j \neq i} r_j} where \quad \delta \geq 0 \tag{7}$$

In parallel with the generalization proposed before, a node's (j) resources is given more or less weight based on the geodesic distance to ego i ($1/g_{ij}$) and depending on the value δ. As before, the measure captures how much the level of 'indirect connectedness' between ego and the other node (i.e., the number of steps) has an impact on ego. In this generalized equation the same δ is used as before to regulate the importance of distance. Similar to the measure before, a high δ would assign more weight to the resources of other nodes that are very close by, while a δ close to 0 would focus on the total resources available in a component. A negative value would indicate that it is important to have resourceful people at a large distance, but still reachable.

In a similar way as before, when δ is positive, standardization is possible by dividing by the total resources available from other nodes. However, when δ is negative the best situation is obtained when nodes with the highest resources are at the longest distance from ego, i.e. where the resource is still available but at a long distance from ego. In other words, in order to normalize the measure the value in the denominator consists of a path, with ego at one end and in which the other node with the least resources is put closest to ego, the node with the second least resources is put as second and so forth, until the last (most resourceful other) who is put last on the path. In the equation below this is represented by $C'^{\delta}_R(i)$. The sum of these resources (r^+ being the resources ordered in ascending order excluding the resource for i) is therefore weighted by their respective distances (and the choice of δ):

$$C'^{\delta}_R(i) = \frac{\sum_{j \neq i}({g_{ij}}^{-\delta} \cdot r_j)}{\sum_{k \neq i}(\mathrm{k}^{-\delta} \cdot r_k^+)} where \quad \delta < 0 \tag{8}$$

Returning to the example of the classic marriage network among Florentine families (Padgett and Ansell, 1993, see Fig. 5), with the attribute wealth now being used to measure the resource-richness of a person, we see (Table 3) that while the Peruzzi clearly have a lot of resources in their direct environment (a non-normalized value of 20+146+44=210, when δ =100), the Guadagni become more resource-rich when indirect connected are given importance (a non-normalized value of 501.29 when δ =0.5, compared to a value of only 453.32 for the Peruzzi). The reason being that the Peruzzi are directly connected to the wealthy Strozzi, while the Guadagni are two steps away from both the wealthy Strozzi and the wealthy Medici. Hence when longer distances are also given importance (i.e. a lower, but still positive δ is chosen) the indirect resource-richness of these two families increases. Hence, while the Peruzzi might be better at marrying with wealthy others, this will be beneficial if wealth is only impacting the direct neighbours. However, if there are reasons to believe that the impact of wealth might spread beyond those directly connected to others in the network (with a certain decay function), then we might find that the Guadagni are better connected. If connection to resource-rich nodes at long distances are preferred over short distances, then negative values for δ should be chosen. This might indicate an opportunity to acquire resources from a longer distance without having the disadvantages of having to be under their direct influence of local contacts. Hence the Ginori and the Lamberteschi are still able to benefit from the resources in their component through their indirect contacts, while remaining (relatively) far away from either wealthy families (see Fig. 5). Note that the Pucci are completely unconnected to the main component, and therefore are unable to benefit from the resources of the main component, while also not being under their pressure.

More generally, if only the resources of direct contacts matter, then it would suffice to use the resourceful degree measure (or in other words use a δ set to 100). However, if there is any reason to assume that resources at a longer distance from ego matter, and in other words resources are transmittable through intermediary nodes in a network, the general measure proposed in this paper is appropriate with a lower or higher (positive) value for δ. Moreover, if being part of a resourceful component is important, but the precise distance is not, then a δ equal to 0 might be more appropriate.

3.2. Diversity

Another approach to incorporate attributes has been to capture the diversity or heterogeneity in resources of nodes in ego's surrounding (Marsden, 1987; Burt, 1983). The diversity in characteristics of alters (Burt, 1983) might indicate one's access to more diverse or unique sets of information or other resources, as the attributes used to capture diversity can indicate access to different social circles (cf. Granovetter, 1973; Reagans and McEvily, 2003). Measures of diversity, such as Blau's measure of heterogeneity (Blau, 1977), generally consider the diversity among ego's direct contacts. Hence, such a measure does not focus on the amount of access to a specific characteristic (like the resourceful approach), but rather focuses on the diversity of the other nodes in ego's surrounding with respect to some characteristics which represent specific social circles or categories of people with unique benefits (Marsden, 1987; Campbell et al., 1986). Often these measures focus on characteristics of nodes which could (implicitly) indicate the type of resources and other benefits that they might provide for

[7] Alternatively the sum could be divided by the number of other nodes reached. In that case the mean resourcefulness of the other nodes is captured. However, in most cases we are interested in the total amount of resources that are accessible (i.e., the richness) rather than the average amount of resources as these resources from different sources all contribute (i.e. accumulate) and therefore we focus here on the sum of resources.

Table 3
Results of (non-normalized) generalized resource-richness measure with different δ's (ordered from most to least resource-richness).

δ = 0.1		0.5		1.0		1.5		2		100	
GUADAGNI	630.00	RIDOLFI	531.49	RIDOLFI	453.67	RIDOLFI	402.50	RIDOLFI	368.56	RIDOLFI	297
RIDOLFI	622.77	GUADAGNI	501.29	CASTELLANI	404.83	CASTELLANI	348.32	CASTELLANI	313.57	CASTELLANI	250
CASTELLANI	618.90	CASTELLANI	498.70	GUADAGNI	392.17	GUADAGNI	320.50	GUADAGNI	272.81	PERUZZI	210
SALVIATI	617.31	TORNABUONI	466.56	BISCHERI	367.43	BISCHERI	307.52	BISCHERI	270.53	BISCHERI	203
ACCIAIUOLI	612.31	BISCHERI	466.49	TORNABUONI	359.17	MEDICI	301.78	MEDICI	264.50	MEDICI	186
ALBIZZI	598.83	SALVIATI	461.88	MEDICI	358.00	PERUZZI	296.68	PERUZZI	263.11	GUADAGNI	170
BISCHERI	592.92	ALBIZZI	461.72	PERUZZI	353.93	TORNABUONI	287.94	TORNABUONI	240.31	SALVIATI	151
TORNABUONI	591.53	PERUZZI	453.32	ALBIZZI	348.67	ALBIZZI	277.05	ALBIZZI	231.22	ALBIZZI	143
PERUZZI	585.28	BARBADORI	450.73	SALVIATI	341.42	SALVIATI	269.75	SALVIATI	226.33	STROZZI	140
BARBADORI	582.00	MEDICI	443.55	BARBADORI	340.50	BARBADORI	268.91	BARBADORI	221.79	TORNABUONI	138
GINORI	579.97	ACCIAIUOLI	441.60	ACCIAIUOLI	309.42	STROZZI	244.52	STROZZI	209.39	BARBADORI	123
LAMBERTESCHI	570.85	STROZZI	387.90	STROZZI	299.67	ACCIAIUOLI	230.98	ACCIAIUOLI	183.66	ACCIAIUOLI	103
PAZZI	558.06	GINORI	382.24	GINORI	235.92	GINORI	153.54	GINORI	106.41	GINORI	36
MEDICI	543.94	LAMBERTESCHI	373.59	LAMBERTESCHI	224.35	LAMBERTESCHI	138.13	LAMBERTESCHI	87.52	PAZZI	10
STROZZI	497.12	PAZZI	347.82	PAZZI	198.67	PAZZI	118.11	PAZZI	73.58	LAMBERTESCHI	8
PUCCI	0.00	PUCCI	0.00	PUCCI	0.00	PUCCI	0.00	PUCCI	0.00	PUCCI	0

ego, for example by focusing on the (type of) job they do. For example, in a recent study, Shipilov et al. (2014) have argued how being connected to different units in the organisation might be important because of the boundary spanning possibilities, such as translating ideas from other units to their own unit (Seibert et al., 2001), gluing the organisation together (Reagans and McEvily, 2003), and be more innovative (Burt, 2004).

At a local level (i.e., where ego is directly connected with other nodes) the heterogeneity has been captured by Blau's measure of heterogeneity (Blau, 1977):

$$C_V(i) = 1 - \sum_v \left(\frac{\sum_{j \neq i} (x_{ij} \cdot [V_j = v])}{\sum_{j \neq i} x_{ij}} \right)^2 \quad (9)$$

With v the number of distinct categories for the attribute and V_j the attribute-category for actor j. Taking into account the heterogeneity among the indirect reachable nodes, a generalized version of heterogeneity can be built (cf. Everett and Borgatti, 2012), so that a person might be more heterogeneous if the person is also connected to more diverse others at distance 2, 3 or more. Hence this measure focuses on being connected directly or indirectly with diverse people, where the diversity is weighted according to the geodesic distance between ego and other nodes and with δ indicating the impact that one wants to give to different distances:

$$C_V^{\delta}(i) = 1 - \sum_v \left(\frac{\sum_{j \neq i} (g_{ij}^{-\delta} [V_j = v])}{\sum_{j \neq i} g_{ij}^{-\delta}} \right)^2 \text{ where } \delta \geq 0 \quad (10)$$

Considering the example of the friendship among students again, if we include their grade (3 levels), we can see in Fig. 7 that both node 4 and node 8 are primarily directly connected with blue (circle) others (3 out of 4), but if we consider indirect contacts we can see that node 4 has quite some diversity at distance 2 and 3 (with a high proportion of red (square) and green (triangle) contacts at distance 2), while node 8 remains quite homogeneous in its contacts (as most contacts are blue (circle)).

Like the measures before, when δ is positive, but close to 0, diversity among nodes at a longer distances from ego are also incorporated in the measure. At zero, the heterogeneity score for an ego is equal to the heterogeneity in ego's component. For connected graphs, this is not particularly useful, since every node will receive approximately the same score. In contrast, for higher values of δ, the impact of heterogeneity among more distant nodes becomes less important relative to the impact of diversity among those close by.

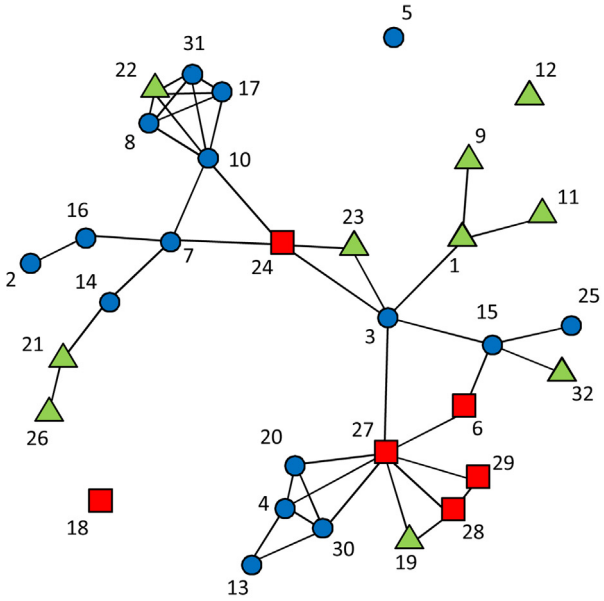
In the case of Fig. 7 when δ is 0, nodes from the same component will have approximately the same value. Finally, a negative value for δ focuses on the diversity among the more distant nodes *within a component*, while (largely) ignoring any variation among nodes who are more close to ego.

Although we have focused on two widely used measures that incorporate attributes, we do want to emphasize that the approach is more general and can be used to many of the attribute-based measures, such as the Gini coefficient and even the standard deviation (for continuous alter attributes).

4. Ego-betweenness, Freeman's betweenness and generalized betweenness centrality

The approach above can also be applied to many other measure that make use of the geodesic distance or only incorporate direct relations, by replacing these with ${g_{jk}}^{-1}$. An interesting further extension incorporates Freeman's betweenness and local or ego betweenness (Everett and Borgatti, 2005).

Betweenness centrality focuses less on access to information for a focal actor, but rather concentrates on the power resulting from being on the shortest path among others. Having a high betweenness centrality can generate power and influence because one is a broker between others in the network (Brass, 1984; Burt, 1992). Hence, this involves three actors, with the focus on actor i being on the shortest path between actors j and k. Let t_{jk} denote the total number of shortest paths connecting j to k and t_{jik} be the num-



Distance to node 4	1	2	3	4	5	6	7	Infinity
Blue (circle)	3	1	1	3	5	1	0	1
Red (square)	1	3	1	0	0	0	0	1
Green (triangle)	0	1	2	3	1	1	1	0
Total	***4***	***5***	***4***	***6***	***6***	***2***	***1***	***2***

Distance to node 8	1	2	3	4	5	6	7	Infinity
Blue (circle)	3	1	3	2	4	1	0	1
Red (square)	0	1	0	1	3	0	0	1
Green (triangle)	1	0	1	2	5	0	0	0
Total	***4***	***2***	***3***	***5***	***12***	***1***	***0***	***2***

Fig. 7. Symmetrized friendship network among 32 Dutch students (Van de Bunt, 1999) with attribute grade (focus on students 4 and 8).

ber of shortest paths connecting j to k that pass through i then the betweenness of i, $C_B(i)$ is given by:

$$C_B(i) = \sum_{j<k} \left(\frac{t_{jik}}{t_{jk}} \right) \tag{11}$$

We can then make the contribution of people (j and k) at specific distances from i, dependent on the tuning factor δ, so that the most local version is obtained when j is directly connected to i and i directly connected to k. Denote by g_{jk} the geodesic distance from j to k, then our generalized measure is given by:

$$C_B^{\delta}(i) = \sum_{j<k} \left(\frac{t_{jik}}{t_{jk}} \right) (g_{jk} - 1)^{-\delta} \tag{12}$$

When δ is high this reduces to a form of ego betweenness (Everett and Borgatti, 2005), as only direct relations between i and k, and i and j are now counted (i.e., when g_{jk} equals 2).[8] On the other hand when δ is set to 0, then this reduces to the standard Freeman betweenness as all j and k dyads now count irrespective of distances (as long as they can be reached eventually). When δ equals 1 then we obtain length scaled betweenness first proposed by Borgatti (2005) and available in the UCINET program (Borgatti et al., 2002). It should be noted that Brandes (2008) coined the term length scaled betweenness and gave an algorithm for its computation. His implementation differs from the one proposed here as he uses g_{jk} in place of our g_{jk}-1. We suggest that our formulation better reflects the original concept as it means that in graphs of diameter 2 length scaled betweenness is exactly the same as ordinary betweenness. In the Brandes formulation geodesics of minimum length that contribute to betweenness must have length 2 and so contribute 0.5 to the betweenness score. The original Freeman betweenness formulation gives a contribution of 1 and we suggest that in order to be comparable length scale betweenness should equally give a score of 1 for geodesics of length 2.

Consider the friendship network among 32 students again (Fig. 1), we find that different nodes become important depending on whether we consider more local or global betweenness (Table 4). Consider the betweenness (Fig. 8) we find that student 27 has the highest betweenness when focusing on local (ego) betweenness (δ=100), while 3 becomes more central if we also consider pairs of dyads at longer distance (δ=1 or lower). The high value for student 27 at a local level is not surprising given the amount of local contacts. Similarly the global betweenness of student 3 is not surprising given the closeness at longer distance.

5. Discussion and conclusion

A central theme in network theorizing is the notion that certain individuals are better positioned than others to access resources possessed by others or that are flowing through the network. A variety of node-level measures capture relative quality of a node's location in the network. Some measures focus solely on direct ties,

[8] In this case we include paths of length two, connecting two alters that do not go through ego. Everett and Borgatti (2005) do not include these paths.

Table 4
Results of generalized betweenness centrality with different δ's (van de Bunt friendship data).

	δ									
	0	0.1	0.25	0.5	1	1.5	2	3	5	100
1	53.00	47.44	40.30	30.97	18.97	12.31	8.54	5.08	3.36	3.00
2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	243.00	216.08	181.73	137.34	81.26	50.74	33.69	18.14	10.27	8.50
4	13.00	11.54	9.70	7.36	4.51	3.01	2.20	1.47	1.09	1.00
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	12.00	10.93	9.51	7.58	4.92	3.29	2.29	1.26	0.66	0.50
7	121.00	106.65	88.58	65.69	37.80	23.33	15.60	8.89	5.70	5.00
8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	96.00	86.04	73.27	56.65	35.47	23.87	17.38	11.53	8.63	8.00
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14	52.00	45.17	36.68	26.14	13.78	7.71	4.65	2.20	1.18	1.00
15	61.50	55.22	47.17	36.68	23.25	15.84	11.65	7.82	5.91	5.50
16	27.00	23.76	19.67	14.49	8.20	4.96	3.25	1.79	1.13	1.00
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
21	27.00	23.29	18.72	13.15	6.83	3.88	2.47	1.43	1.06	1.00
22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
24	187.00	164.46	135.96	99.68	55.18	32.05	19.76	9.38	4.83	4.00
25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
26	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
27	168.17	151.54	130.30	102.83	68.03	48.99	38.29	28.43	23.35	22.17
28	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34
29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
30	13.00	11.54	9.70	7.36	4.51	3.01	2.20	1.47	1.09	1.00
31	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
32	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 5
A typology of node-level measures of access to resources.

		Utilize Node Attributes	
		No	Yes
Utilize indirect connections	No	A. Degree	B. Resource richness & heterogeneity
	Yes	C. Closeness, Betweenness	D. E-I/G-F Centrality

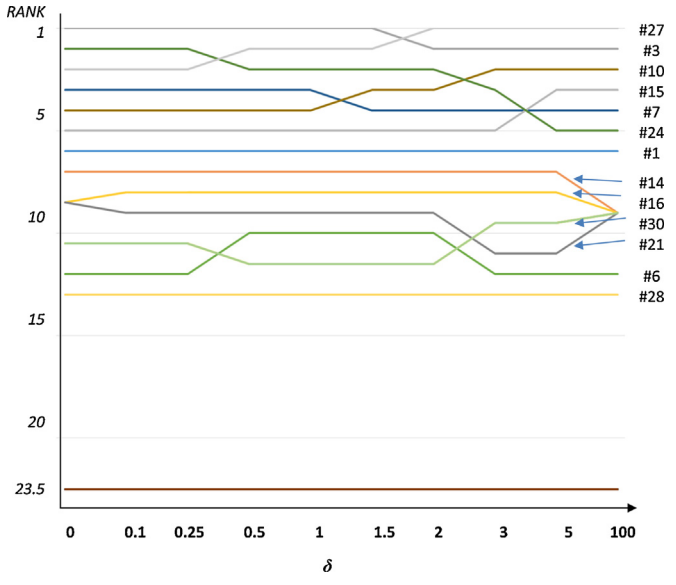


Fig. 8. Rank order for generalized betweenness for 32 students (based on results in Table 4).

while other take account of indirect connections to others. For example, degree centrality counts only direct ties, whereas closeness and betweenness centrality include longer paths. At the same time, some measures are purely topological in the sense that they do not make use of any information about the actors represented by the nodes, whereas others take node attributes into account. Degree, closeness and betweenness, as normally formulated, are purely topological. In contrast, measures of ego network composition, such as resource richness or heterogeneity, are explicitly based on node attributes such as wealth or power. Table 5 organizes these measures in a two-by-two table.

In this paper we accomplish two things. First, we convert the binary distinction between utilizing direct ties versus utilizing indirect connections into a continuous gradient. We do this by constructing a parameterized measure where the parameter δ controls the relative weighting of near and far connections. The measure yields both degree and closeness as special cases generated by different choices of δ. We also show that when a node outcome is available, we can empirically fit δ by choosing the value that maximizes the correlation between the centrality measure and the node outcome. The fitted value then tells us about the flow processes in our research setting because large values of δ indicate rapid decay of flows such that only short paths matter, whereas small (includ-

ing negative) values of δ indicate that what is flowing continues to give benefit even when it travels via long paths.

As a further extension, we showed how this approach can also be applied to measures that focus on a node being a broker between other nodes, and more specifically provide an example based on betweenness centrality. The extent to which a pair of nodes with a specific geodesic distance will contribute to the node's betweenness value will depend on the choice of δ. We showed how local (ego) betweenness and global (Freeman) betweenness centrality can be considered specific outcomes of a more general betweenness measure, when a specific δ is chosen. Again, while we focus on betweenness, one could also apply these principles in order to, for example, generalize Burt's constraint index (Burt, 1992) or Gould and Fernandez' brokerage roles (Gould and Fernandez, 1989).

The second contribution of this paper is to similarly generalize node-level measures of ego net composition, including resource richness and heterogeneity. Thus, instead of measuring resource richness by summing the wealth only of ego's direct contacts, we allow for the possibility of benefiting from resources two steps away, three steps away, and so on. The parameter δ again controls the relative weight we place on near or far resources. The generalization of resource heterogeneity is accomplished in the same way. Substantively, the result is that our measure of resource heterogeneity is not limited to just those controlled by a node's direct contacts, but also those controlled by friends of friends, and so on. And, of course, our method is perfectly general: we can generalize any measure of ego net composition, including measures of homophily. It is not difficult to imagine research settings where we would want to take account of the *depth* of homophily, in the sense that not only are the friends similar to ego, but also the friends of friends.

While we have focused on shortest paths, the principle proposed here is not restricted to the geodesic distance. We could consider measures that incorporate paths of different length and even consider random walks or trails (Borgatti, 2005; Newman, 2005; Bonacich, 1972, 1987). In this way, the general approach here can be extended further. We can also extend the measures to deal with valued or weighted networks (e.g, Opsahl et al., 2010) or measures that incorporate positive and negative relations (Marineau et al., 2016). All these are potential avenues for further research. One of the fruitful benefits of this approach is that it fosters discussion about the importance of shorter and longer distances for specific cases as well as allow more empirical research on exploring how direct and more indirect distances contribute to an outcome in specific situations, leading to a volume of studies which help understand when distances are important and when direct relations are sufficient.

Acknowledgement

This work was partly funded by Leverhulme Trust Research Project Grant number R116318, PI Martin Everett.

References

Agneessens, F., Koskinen, J., 2016. Modelling individual outcomes using a Multilevel Social Influence (MSI) model Individual versus team effects of trust on job satisfaction in an organisational context. In: Lazega, E., Snijders, T.A.B. (Eds.), Multilevel Network Analysis for the Social Sciences. Theory, Methods and Applications. (Methodos Series). Springer, New York.
Blau, P.M., 1977. Inequality and Heterogeneity. The Free Press, New York.
Bonacich, P., 1972. Factoring and weighting approaches to status scores and clique identification. J. Math. Sociol. 2, 113–120.
Bonacich, P., 1987. Power and centrality: a family of measures. Am. J. Sociol. 92, 1170–1182.
Borgatti, S.P., Everett, M.G., 2006. A graph-theoretic framework for classifying centrality measures. Soc. Netw. 28, 466–484.
Borgatti, S.P., Halgin, D.S., 2011. On network theory. Organiz. Sci. 22, 1168–1181.
Borgatti, S.P., Jones, C., Everett, M.G., 1998. Network measures of social capital. Connections 21, 36.
Borgatti, S.P., Everett, M.G., Freeman, L.C., 2002. Ucinet for Windows: Software for Social Network Analysis. Analytic Technologies, Harvard, MA.
Borgatti, S.P., 2005. Centrality and network flow. Soc. Netw. 27, 55–71.
Borgatti, S.P., 2006. Identifying sets of key players in a network. Computational. Math. Organiz. Theory 12, 21–34.
Brandes, U., 2008. On variants of shortest-path betweenness centrality and their generic computation. Soc. Netw. 30, 136–145.
Brass, D.J., Burkhardt, M.E., 1993. Potential power and power use: an investigation of structure and behavior. Acad. Manage. J. 36, 44–470.
Brass, D., 1984. Being in the right place: a structural analysis of individual influence in an organization. Adm. Sci. Q. 29, 518–539.
Burt, R.S. 1983. Range. Pp. 176–194 in R.S. Burt and M.J. Minor (Eds.) Applied Network Analysis. Beverly Hills: Sage Publications.
Burt, R.S., 1992. Structural Holes: The Social Structure of Competition. University Press, Harvard.
Burt, R.S., 2004. Structural holes and good ideas. Am. J. Sociol. 110, 349–399.
Campbell, K.E., Lee, B.A., 1991. Name generators in surveys of personal networks. Soc. Netw. 13, 203–221.
Campbell, K.E., Marsden, P.V., Hulbert, J.S., 1986. Social resources and socioeconomic status. Soc. Netw. 8, 97–117.
Duijn, P.A.C., Kashirin, V., Sloot, P.M.A., 2014. The relative ineffectiveness of criminal network disruption. Nat. Sci. Rep. 4 (Article number: 4238).
Ellwardt, L., Labianca, G., Wittek, R., 2012. Who are the objects of positive and negative gossip at work? A social network perspective on workplace gossip. Soc. Netw. 34, 193–205.
Everett, M.G., Borgatti, S.P., 2005. Ego network betweenness. Soc. Netw. 27, 31–38.
Everett, M.G., Borgatti, S.P., 2012. Categorical attribute based centrality: E-I and G-F centrality. Soc. Netw. 34, 562–569.
Fischer, C.S., 1982. To Dwell Among Friends: Personal Networks in Town and City. University of Chicago Press, Chicago.
Freeman, S.C., Freeman, L.C., 1979. The Networkers Network: A Study of the Impact of a New Communications Medium on Sociometric Structure Social Science Research Reports No 46. University of California, Irvine CA.
Freeman, L., 1977. A set of measures of centrality based on betweenness. Sociometry 40, 35–41.
Freeman, L., 1979. Centrality in social networks: conceptual clarification. Soc. Netw. 1, 215–239.
Gould, R.V., Fernandez, R.M., 1989. Structures of mediation: a formal approach to brokerage transaction networks. Sociol. Methodol. 19, 89–126.
Granovetter, M.S., 1973. The strength of weak ties. Am. J. Sociol. 78, 1360–1380.
Hansen, M.T., 1999. The search-transfer problem: the role of weak ties in sharing knowledge across organization subunits. Adm. Sci. Q. 44, 82–111.
Hurlbert, J.S., 1991. Social networks, social circles, and job satisfaction. Work Occup. 18, 415–429.
Latora, V., Marchiori, M., 2001. Efficient behavior of small-world networks. Phys. Rev. Lett. 87 (19), 198701.
Lin, N., Ensel, W.M., Vaughn, J.C., 1981a. Social resources and the strength of weak ties: structural factors in occupational status attainment. Am. Sociol. Rev. 46, 393–405.
Lin, N., Vaughn, J.C., Ensel, W.M., 1981b. Social resources and occupational status attainment. Soc. Forces 59, 1163–1181.
Lin, N., 1986. Conceptualizing social support. In: Lin, N., Dean, A., Ensel, W. (Eds.), Social Support, Life Events and Depression. Academic Press, New York.
Lin, N., 1999. Building a network theory of social capital. Connections 22, 28–51.
Lin, N., 2001. Social Capital: A Theory of Social Structure and Action. University Press, NY: Cambridge.
Marineau, J.E., Labianca, G., Kane, G.C., 2016. Direct and indirect negative ties and individual performance. Soc. Netw. 44, 238–252.
Marsden, P.V., 1987. Core discussion networks of Americans. Am. Sociol. Rev. 52, 122–131.
Marsden, P.V., 1990. Network data and measurement. Ann. Rev. Sociol. 16, 435–463.
Newman, M.E.J., 2003. The structure and function of complex networks. SIAM Rev. 45, 167–256.
Newman, M.E.J., 2005. A measure of betweenness centrality based on random walks. Soc. Netw. 27, 39–54.
O'Connor, L.T., 2013. Ask and you shall receive: social network contacts' provision of helpduring the job search. Soc. Netw. 35, 593–603.
Opsahl, T., Agneessens, F., Skvoretz, J., 2010. Node centrality in weighted networks: generalizing degree and shortest path. Soc. Netw. 32, 245–251.
Padgett, J.F., Ansell, C.K., 1993. Robust action and the rise of the Medici, 1400–1434. Am. J. Sociol. 98, 1259–1319.
Reagans, R., McEvily, B., 2003. Network structure and knowledge transfer: the effects of cohesion and range. Adm. Sci. Q. 48, 240–267.
Seibert, S.E., Kraimer, M.L., Liden, R.C., 2001. A social capital theory of career success. Acad. Manage. J. 44, 219–247.
Shipilov, A., Labianca, G., Kalnysh, V., Kalnysh, Y., 2014. Network-building behavioral tendencies, range, and promotion speed. Soc. Netw. 39, 71–83.
Sparrowe, R.T., Liden, R.C., Wayne, S.J., Kraimer, M.L., 2001. Social networks and the performance of individuals and groups. Acad. Manage. J. 44, 316–325.
Thoits, P.A., 1982. Conceptual, methodological, and theoretical problems in studying social support as a buffer against life stress. J. Health Soc. Behav. 23, 145–159.

Valente, T., Coronges, K., Lakon, C., Costenbader, E., 2008. How correlated are network centrality measures? Connections 28, 16–26.
Van de Bunt, G.G., van Duijn, M.A.J., Snijders, T.A.B., 1999. Friendship networks through time: an actor-oriented statistical network model. Comput. Math. Organiz. Theory 5, 167–192.
Van de Bunt, G.G., 1999. Friends by Choice. An Actor-oriented Statistical Network Model for Friendship Networks Through Time. Thesis Publishers, Amsterdam.
Vaux, A., 1988. Social Support: Theory, Research, and Intervention. Praeger, New York, pp. 1988.
van der Gaag, M.P.J., Snijders, T.A.B., 2005. The Resource Generator: measurement of individual social capital with concrete items. Soc. Netw. 27, 1–29.