



Induced, endogenous and exogenous centrality

Martin G. Everett^{a,*}, Stephen P. Borgatti^{b,1}

^a Mitchell Centre for Social Network Analysis, School of Social Sciences, University of Manchester, Arthur Lewis Building, Bridgeford Street, Manchester M13 9PL, UK

^b Department of Management, Gatton College of Business & Economics, University of Kentucky, Lexington, KY 40506, USA

ARTICLE INFO

Keywords:

Centrality
Degree
Closeness
Betweenness
Vitality measures

ABSTRACT

Centrality measures are based upon the structural position an actor has within the network. Induced centrality, sometimes called vitality measures, take graph invariants as an overall measure and derive vertex level measures by deleting individual nodes or edges and examining the overall change. By taking the sum of standard centrality measures as the graph invariant we can obtain measures which examine how much centrality an individual node contributes to the centrality of the other nodes in the network, we call this exogenous centrality. We look at exogenous measures of degree, closeness and betweenness.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Centrality has always been of primary interest to social network analysts. It is possibly the technique most commonly applied when analyzing network data. In addition, it has attracted the attention of those interested in devising new methods and there are a plethora of techniques available. The software package UCINET (Borgatti et al., 2002) for example currently implements 16 different centrality measures directly and it also produces a number of other centrality measures as direct products of other routines.

Freeman's (1979) classic paper brought together the important measures of the time and gave a general framework for understanding centrality measures. More recently Borgatti (2005) provided a framework and typology by considering network flow and Borgatti and Everett (2006) examined and categorised measures based on graph theoretic properties. The goal of the present paper is to continue this stream of research into the fundamental nature and interpretation of centrality in social networks. As such, it is not our primary goal to introduce new centrality measures, although it will be clear that our approach lends itself to creating new centrality measures appropriate for any specific research objective.

In this paper we examine a technique which has been in use – knowingly or not – for several decades. The basic strategy is to take any whole network property, such as density or maximum flow, and derive a centrality measure by deleting nodes and measuring the change in the network property. Measures derived in this way have been called “vitality measures” for a brief review see

Koschützki et al. (2005), although we prefer the more descriptive term “induced centrality measures”.

One of the advantages of this approach is that centrality measures can be constructed in a customized way for any research need, avoiding having to use an “off-the-shelf” measure which may not be wholly appropriate in a given setting. The technique is also easily extendable to the case of edge centralities by measuring the impact of removing an edge, and also to the case of group centralities (Everett and Borgatti, 1999; Borgatti, 2006), by removing sets of nodes or edges.

2. Induced centrality

A graph invariant is any quantitative graph property which only depends upon the graph structure and not on a representation or a labelling of a graph. For example, the density and the transitivity of graphs (proportion of transitive triples) are examples of invariants, as is the property of containing a cycle of length 5. The graph may be weighted or directed. In general, a graph invariant is usually a single number but the concept can be easily extended to sets or vectors. For example, the degree sequence of a graph is a graph invariant that is not a single number.

2.1. Definition

Let f be a graph invariant and G a graph with vertex (or edge) x . Then the induced centrality C_f of vertex (or edge) x is given by

$$C_f(x) = f(G) - f(G - \{x\}) \quad (1)$$

where $G - \{x\}$ is the graph G with vertex (or edge) x deleted, in an abuse of notation we shall often write this as $G - x$. It should be noted that we can induce centrality on nodes or edges for any graph invariant that is defined for all graphs. If the invariant is not

* Corresponding author. Tel.: +44 01612752515.

E-mail addresses: martin.everett@manchester.ac.uk (M.G. Everett), sborgatti@uky.edu (S.P. Borgatti).

¹ Tel.: +1 8592572257.

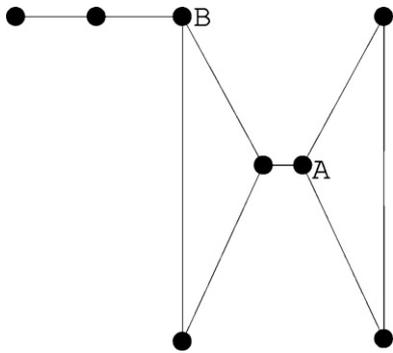


Fig. 1. A graph in which the deletion of vertex A or B results in the same subgraph.

a single number then we will need to define the difference of the sets or vectors appropriately. For convenience, we assume that the co-domain of f is a subset of the set of real numbers.

Several well known measures of centrality can be formulated as induced centralities. For example, if the graph invariant is the total number of edges in a graph, then the induced vertex centrality is simply degree centrality. If the graph invariant is the total maximum flow in the network then the induced vertex centrality is the max-flow betweenness (Freeman et al., 1991; Koschützki et al., 2005). These two examples exploit a simple and direct relation between the invariant chosen and the contribution that a node makes to that invariant.

Given that we can formulate degree and max-flow betweenness as induced centrality measures, it is reasonable to ask if all centrality measures are induced by some graph invariant. The answer to this question is no. Consider the graph in Fig. 1, then the deletion of either A or B results in the same residual graphs (i.e., $G-A$ is isomorphic to $G-B$). This means that for any graph invariant we choose, the induced centrality for nodes A and B must be the same value. Yet, if we measure the closeness centrality of each node, we find that A has a score 14 and B has a score of 13. Hence, closeness cannot be an induced centrality.

Borgatti and Everett (2006) show that betweenness can be formulated in a way similar to an induced centrality measure as follows:

$$C_x = \sum_{i,j} \frac{w_{i,j}^x - w_{i,j}^{*x}}{w_{i,j}^x} \quad (2)$$

In (2) w^x is the geodesic count matrix of the original graph with row and column x deleted and w^{*x} is the geodesic count matrix of the graph with vertex x deleted. The ij th entry of a geodesic count matrix is the number of geodesics connecting vertex i with vertex j . The reason this is not an induced measure is that we do not directly define a graph invariant but only consider measures that depend on identifying x .

2.2. Graph invariants

There are a large number of graph invariants in use today (and new ones being invented all the time) and in principle any of them could be used to define a new centrality measure. However, there are a number of properties of invariants that affect how useful they are in defining induced centrality measures.

First, the graph invariant should be defined on all graphs so that vertex (or edge) deletion does not result in a graph in which the invariant is not defined. For example, if we use graph diameter as our invariant, we will be unable to compute a centrality score for any node whose removal disconnects the graph. This is especially an issue with directed graphs. Graph invariants that suffer from this can usually be modified or extended to overcome this prob-

lem; we give a detailed example of this for closeness later in the paper.

Second, graph invariants should not be normalized for the number of nodes in a graph (or edges if computing an edge centrality) since the normalization may well remove the contribution made by the deleted vertex or edge. As an extreme example consider the case where we use density as the graph invariant. Suppose we have a complete network, so that density is 1.0. Removing a node leaves a residual network that is also complete and therefore also has density 1.0. As a result, each node would have an induced centrality score of zero, incorrectly indicating that no node contributes to the density of the graph. Other invariants that fall in this category are average path length and fragmentation (proportion of pairs of vertices that cannot reach each other, see Borgatti, 2006).

Third, invariants should be sufficiently sensitive to node removal so that they change significantly under vertex (or edge) deletion. For example, a graph invariant that is insensitive to node removal is the clique number of a graph. The clique number is the size of the largest clique. If the graph has two largest cliques, and they are disjoint, then vertex deletion will not change the invariant. Even if the graph has a unique largest clique then deletion of vertices not in this clique will not change the invariant while vertices in the largest clique will all have an induced centrality score of 1. Thus all values in the resulting centrality measure will be either 0 or 1.

Fourth, to be interesting, the change in an invariant due to the removal of a node should be dependent on the structure of the residual graph. This ensures that the measure is not just a simple count of a node's properties. For example, we could take as our invariant the total number of edges in a graph. Deleting a vertex removes all edges incident upon it and so the resultant induced measure is a simple count of the number of edges the node is involved in. Similarly, if we take as our invariant the number of transitive triples in a graph, the resulting induced measure is simply a count of the number of transitive triples containing the deleted vertex.

3. Endogenous, exogenous and total centrality

An interesting class of invariants to use for induced centrality is based on aggregating other centrality measures. That is, we can take as our invariant the sum of centrality scores (using any measure C we like) for all vertices in a network. Deleting a node x and recalculating the invariant gives us an induced centrality that we call the "total centrality" $C_T(x)$ of node x . This is shown in (3)

$$C_T(x) = \sum_{j \in V(G)} C(j) - \sum_{j \in V(G-x)} C'(j) \quad (3)$$

where $V(G)$ is the vertex set associated with the graph G and C' is calculated on the graph $G-x$. The idea is that a node's total centrality reflects not only the node's direct contribution to the overall centrality in the graph, but also the centrality that it gives to other nodes by being adjacent to them or being along paths that connect them. For example, consider a measure of centrality based on the shortness of paths from a node to all others, so that the shorter the paths, the greater the centrality. (We leave the details of such a measure for another section.) If we sum the centralities of all nodes, we obtain a network-level statistic—our graph invariant. Now let us consider the contribution of a node x to that invariant. Clearly, a piece of the invariant is the node's measured centrality, since it was one of the n quantities summed to get the invariant. This is the direct or endogenous contribution of the node to the invariant. But consider that if the node were to sever all ties with the rest of the network, it would not be just its own centrality that would be reduced. The node could have been a shortest path intermediary

such that without it the distances between certain pairs of nodes would be increased, reducing their centralities. This reduction in other node's centrality is an indirect or exogenous contribution of the node to the invariant.

To summarize, we are proposing the following. Given a suitable centrality metric to serve as an endogenous centrality measure, we can then construct an induced centrality by choosing as our graph invariant the sum of centralities across all nodes. We can then subtract the endogenous centrality from the induced centrality to obtain what we are terming exogenous centrality. Thus, we are decomposing an induced centrality based on a centrality sum into endogenous and exogenous contributions:

total (induced) centrality = endogenous centrality + exogenous centrality

We now consider several well-known kinds of centrality using this approach.

3.1. Degree

An illuminating example is provided by degree centrality in an undirected graph. We begin by calculating the simple degree of each vertex and setting this as our endogenous centrality measure. The graph invariant is the sum of all the degrees of all the vertices. If we delete a vertex this sum will be reduced by twice the degree centrality of that vertex and hence the induced centrality score for a vertex will be twice its degree. This total centrality therefore reflects the degree of the deleted vertex plus the degree given to all the vertices it is adjacent to.

Now consider the case of in-degree in a directed graph. Our graph invariant is now the sum of the in-degrees of all the vertices and the total centrality is the node's in-degree plus its out-degree. The in-degrees, of course, we were given and constitute the endogenous centrality of the individual vertex. The total centrality is the vertex's centrality plus the centrality it gives to other vertices, in this case simply its out-degree. Clearly we could just as well take the endogenous centrality as out-degree, in which case the exogenous centrality becomes in-degree.

3.2. Betweenness

We should first begin by noting that betweenness is in some respects an unusual centrality measure. As many have noted, the addition or deletion of edges not incident to a given vertex can in some cases increase that vertex's centrality and in some cases decrease it. This is in contrast with degree or closeness centrality where the addition of edges either increases the centrality or has no effect but can never decrease it.

As an example we look at the network of marriages among Renaissance Florentine families compiled by Padgett and Ansell (1993). Taking betweenness centrality as our endogenous centrality measure, Table 1 gives total, endogenous and exogenous scores (rounded to nearest integer).

There are a number of things to note about Table 1. First, the Medici family have the most endogenous centrality and are also quite high on exogenous centrality. As a result, they have the highest total centrality. The nodes with the highest exogenous scores are Salviati followed by Pazzi, both of which have very low endogenous scores (zero in the case of Pazzi). The main reason for their high exogenous scores is they are both on the one of the longer geodesics in the graph. Long geodesics contribute more to the total centrality since every actor on the path obtains a score for being an intermediary. This of course is a direct consequence of the way betweenness is constructed. Different betweenness type measures have been proposed that do not have this property, for example ego network betweenness (Everett and Borgatti, 2005).

Table 1
Induced centrality scores using betweenness for endogenous centrality.

	Total centrality	Endogenous centrality	Exogenous centrality
Medici	73	48	25
Guadagni	38	23	15
Albizzi	43	19	24
Salviati	57	13	44
Ridolfi	9	10	-1
Bischeri	11	10	1
Strozzi	14	9	5
Barbadori	14	9	5
Tornabuoni	13	8	5
Castellani	18	5	13
Peruzzi	24	2	22
Pazzi	35	0	35
Ginori	28	0	28
Acciaiuoli	24	0	24
Lamberteschi	29	0	29
Pucci	0	0	0

It is also interesting to note the negative exogenous score of Ridolfi showing that their presence takes centrality away from the other vertices. In fact it is possible to construct examples with large negative exogenous betweenness centrality scores so that the total centrality is also negative. As mentioned in Section 1, this is because deletion of edges can increase or decrease the betweenness of vertices in the network.

There is a fundamental difference in the way the total centrality and exogenous betweenness centrality scores are calculated compared to endogenous measures, and this needs to be taken into account when interpreting any results. If we take the approach suggested by Borgatti (2005) and think about centrality in terms of network flow then we can see that betweenness (that is the endogenous score for betweenness) only considers the role of intermediaries in the flow process. However, exogenous (and therefore by implication total centrality) also takes into account the endpoints of the geodesic path (provided it is not of length one) since they do send and receive flow. Once they are removed it is not possible for certain other vertices (e.g., the penultimate node in a geodesic) to have betweenness as they are now endpoints. As a result, nodes can contribute to the total betweenness centrality even though they do not contribute to the endogenous value. In the Padgett example we see that four actors with an endogenous score of zero have non-zero total centrality. Note that for an actor to have a zero total betweenness centrality score they would need to be in a component consisting of a complete subgraph. Whilst standard betweenness does not take account of the endpoints Brandes (2008) has proposed a number of variants of shortest path betweenness and some of these do take account of the endpoints. It should be noted that all of these variations can be used as a graph invariant in the same way as we have used standard betweenness here.

3.3. Reverse-closeness

Closeness centrality is defined by Freeman (1979) as the sum of geodesic distances to all other nodes. As a reminder that larger numbers indicate greater distance, we refer to this raw score as "farness" and reserve "closeness" for a normalization of farness obtained by dividing $n - 1$ by farness, where n is the number of nodes in the graph. Hence, a node that is distance 1 from all other nodes will receive a normalized closeness score of 1.0. Unfortunately, neither the sum of farness nor normalized closeness centrality scores satisfy our criteria for useful invariants since deleting a vertex or edge can disconnect the graph, yielding undefined distances between pairs of nodes in different components. Therefore to use Freeman's closeness to develop total centrality and exogenous centrality scores we would have to restrict ourselves to

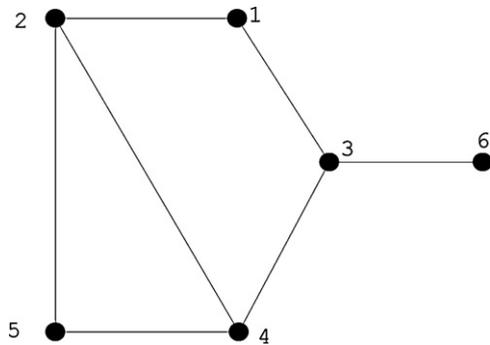


Fig. 2. Graph with a cutpoint used to demonstrate reverse-closeness calculation.

bi-connected graphs (or, in the case of edge centralities, bridgeless graphs), which is less than satisfactory.

Alternatively, we could adopt one of many ways to assign a value to the distance between two unreachable vertices. For example, one approach is to assign n to the distance between two unreachable, since this is one more than the longest possible path length in any graph with n nodes. Another approach is to redefine closeness in terms of reciprocal distances (proximities) and define the proximity of unreachable nodes as zero. Of course, this is a different measure of closeness centrality that will not necessarily give values linearly related to the Freeman measure even in connected graphs (since taking reciprocals is a non-linear transformation). Yet another approach is proposed by Valente and Foreman (1998) in which unreachable pairs are assigned a distance equal to the largest geodesic distance in the graph. For our purposes, however, this approach would be problematic since the largest distance could change across different node (or edge) deletions.

For our purposes, the first method of substituting n for unreachable distances works well, provided that instead of normalizing by dividing $n - 1$ by the sum of distances, we instead simply subtract the sum of distances from $n(n - 1)$, which is the maximum possible (this occurs when a node is an isolate, so that it is at a distance n from all $n - 1$ others). We refer to this measure of centrality as *reverse-closeness* centrality and define it as $n(n - 1)$ minus farness, where undefined distances are assigned a value of n . (Valente and Foreman, 1998 mention this but do not pursue it.) Note that in this measure isolates have a score of zero. The reverse-closeness scores for the graph in Fig. 2 which contains a cutpoint are given in Table 2.

We can now use the sum of reverse-closeness scores as our graph invariant. However there is one subtle point about the reversing process that requires us to make a small change when using it in this context. Consider a three-vertex graph consisting of a single undirected edge and an isolate ($P_2 \cup K_1$) (Fig. 3).

As shown in Table 3, the raw farness score for vertices a and b is 4, and the score for the isolate c is 6. Subtracting these from $n(n - 1) = 6$ gives the reverse-closeness scores of 0 for the isolate and 2 for the other two vertices. This gives a total reverse-closeness score of 4, which we now use as our graph invariant. If we now delete the isolate in order to calculate its total and exogenous contributions, we obtain a new graph with $n^* = n - 1$ nodes, and get a raw farness

Table 2
Farness and reverse-closeness for the graph in Fig. 2.

Vertex	Farness	Reverse-closeness
1	8	22
2	8	22
3	7	23
4	7	23
5	9	21
6	11	19

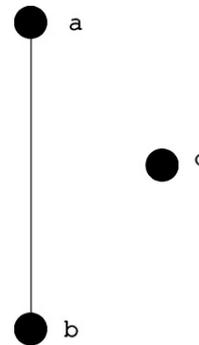


Fig. 3. The graph $P_2 \cup K_1$ which shows how an isolate can have a non-zero contribution to the non-adjusted total reverse-closeness score.

score of 1 for a and b . Subtracting these from $n^*(n^* - 1) = 2$ gives us the reverse-centrality score of 1 for each node, and summing these gives us a new graph invariant score of 2. The difference between 2 and the invariant for the graph as a whole is also 2, so the isolate contributes 2 to the invariant. This is counter intuitive and we would expect isolates to contribute nothing. This occurs because the maximum possible value for farness in the deleted graph is smaller than in the original. (This does not happen in the degree or the betweenness case as the minimum value is fixed at zero regardless of graph size.) This is easily remedied by using the size of the original graph in calculating the reverse-centrality for each node in the reduced graph, instead of using the size of the reduced graph. The result of this is shown in the last column of Table 3, labelled “adjusted reverse-centrality”.

Table 4 gives the total, endogenous and exogenous centrality scores for adjusted reverse-closeness for the graph in Fig. 2 together with the betweenness scores. It should be noted that if we rank the nodes in decreasing order of endogenous scores for reverse-closeness then the order is exactly the same as the ranking derived from the standard Freeman closeness. One of the problems with most closeness type measures is that they do not differentiate well between vertices. Clearly, adjacent vertices will have very similar scores and that is precisely what we see in this instance with nodes 3 and 4 having the same score as well as nodes 1 and 2. However the total centrality score and the exogenous score give a different value for each node. Note these are consistent with the endogenous scores and provides a way of separating the tied pairs. What is interesting is the high exogenous score (and hence total score) for vertex 3. This is because 3 is a cutpoint and hence gives a lot of centrality to the other nodes in the network.

The relationship between exogenous closeness and standard betweenness is striking. Clearly cutpoints give a lot of centrality to the rest of the network. Vertices with high betweenness scores are on many geodesics and the removal of such a vertex must generally move actors further apart and hence give high exogenous scores. We can therefore think of exogenous reverse-closeness as a betweenness type measure.

We now examine the total and exogenous reverse-closeness scores for Padgett data we looked at in the betweenness example in Table 1, these results are shown in Table 5. In the final column we have included the standard Freeman betweenness scores for this data.

Again we see that exogenous and standard betweenness result in a similar ranking pattern. The top four families in the total and the exogenous scores are all cutpoints in the graph. We also note that, as in the previous example, actors with a betweenness score of zero have equal endogenous and exogenous reverse-closeness scores. This makes intuitive sense as nodes with zero betweenness do not affect the distance between any other pairs of nodes. We state this result as a theorem below.

Table 3
Calculating exogenous adjusted reverse-closeness centrality for node *c* in Fig. 3.

Node	Graph G		Graph G-c		
	Farness	Reverse-closeness	Farness	Simple reverse-closeness	Adjusted reverse-closeness
<i>a</i>	4	2	1	1	2
<i>b</i>	4	2	1	1	2
<i>c</i>	6	0	–	–	–
Graph invariant		4	Graph invariant	2	4

Table 4
Reverse-closeness scores for Fig. 2.

Vertex	Total centrality	Endogenous reverse-closeness	Exogenous	Betweenness
1	44	22	22	1
2	46	22	24	1.5
3	74	23	51	4.5
4	50	23	27	3
5	42	21	21	0
6	38	19	19	0

3.4. Theorem

A vertex *x* with a betweenness score of zero has an endogenous reverse-closeness score equal to the exogenous reverse-closeness score.

3.5. Proof

Let *G* be a graph on *n* vertices with a vertex *x* which has a betweenness score of zero. Let *f(x)* be the farness score of *x*, *T_f* be the sum of all the farness scores in *G* and *T_f^x* the sum of all the farness scores in *G* with *x* deleted. It can easily be seen that the total adjusted reverse-closeness centrality *C_T(x)* of a directed graph with *n* vertices is

$$C_T(x) = 2n(n - 1) + T_f^x - T_f \tag{4}$$

Since *x* has a betweenness score of zero, deleting *x* will decrease the total farness score by *2f(x)*. This is because there will be no change in the distances between vertices not involving *x*, the farness score for *x* will be deleted and the value of *f(x)* will be removed in total from all the other vertices. We therefore have

$$T_f^x = T_f - 2f(x) \tag{5}$$

So that

$$C_T(x) = 2n(n - 1) - 2f(x) \tag{6}$$

Table 5
Closeness centrality scores for the Padgett Marriage Data.

	Total centrality	Endogenous centrality	Exogenous centrality	Freeman betweenness
Medici	1248	199	1049	47.5
Guadagni	702	194	508	23.2
Albizzi	692	195	497	19.3
Salviati	664	188	476	13.0
Ridolfi	402	196	206	10.3
Bischeri	398	189	209	9.5
Strozzi	392	192	200	9.3
Barbadori	396	192	204	8.5
Tornabuoni	394	195	199	8.3
Castellani	384	188	196	5.0
Peruzzi	372	186	186	2.0
Pazzi	350	175	175	0.0
Ginori	364	182	182	0.0
Acciaiuoli	372	186	186	0.0
Lamberteschi	362	181	181	0.0
Pucci	0	0	0	0.0

The endogenous centrality *e(x)* is simply adjusted reverse-closeness which by definition is given by

$$e(x) = n(n - 1) - f(x) \tag{7}$$

Since exogenous plus endogenous equals total centrality the result follows.

Note that the converse – i.e., that identical endogenous and exogenous reverse-centrality implies zero betweenness – is false, as can be seen by vertex 1 in Table 4.

3.6. Eigenvector centralities

We discussed the fact that standard closeness posed some problems and therefore used a different (but related) closeness implementation. This made sure the results had certain desirable properties (such as isolates having zero centrality across all the measures) and the measures were well defined. One class of measures for which this technique is not appropriate is eigenvector type measures. For this class of measures we do not have unique raw score since in any eigen type problem multiplying the answer by a constant is also a solution. Eigenvector type centralities are relative and allow us to compare actor scores within the network but we do not have absolute values. Typically in implementations the eigenvector is made to be of unit length for convenience but any normalization would serve the same purpose. When we look at the deleted network we cannot measure the contribution as this requires an absolute measure but this is really the only class for which this restriction is true.

3.7. Normalization of induced centralities and their constituents

We have not normalized our total centrality measure and this is of course possible. We do wish to preserve the idea that total centrality is the sum of the endogenous and exogenous and so each of these measures would need to be normalized using the same value. It is a simple matter to show that the reverse-closeness needs to be normalized by dividing through by $(n - 1)(3n - 2)$. For betweenness the situation is much simpler as the highest possible total centrality score occurs with the star and the total score is the same as the endogenous score and the exogenous score is zero. We can therefore use the standard betweenness normalization throughout.

3.8. Dyadic dependencies

As we have defined it, the total centrality of a node is given by the difference in the sum of centralities of all nodes when a given node is present or absent, as shown in Eq. (3). Equivalently, we can see total centrality as the sum of differences for each node. Each of these differences can be seen as a dyadic effect that each node has on every other. More specifically, we can define a node by node matrix *P* in which *p_{ij}* gives the difference between *j*'s centrality when *i* is present and *j*'s centrality when *i* is absent. Let *C_G(j)* be the centrality

Table 6
Betweenness dependency for Fig. 2.

	1	2	3	4	5	6	Sum	Diagonal	Off-diagonal sum
1	1.0	1.5	1.5	-1.0	0.0	0.0	3.0	1.0	2.0
2	1.0	1.5	-0.5	0.0	0.0	0.0	2.0	1.5	0.5
3	1.0	-0.5	4.5	3.0	0.0	0.0	8.0	4.5	3.5
4	-3.0	-1.5	1.5	3.0	0.0	0.0	0.0	3.0	-3.0
5	0.0	1.0	1.0	2.0	0.0	0.0	4.0	0.0	4.0
6	0.5	0.0	4.0	1.5	0.0	0.0	6.0	0.0	6.0
Sum	0.5	2.0	12.0	8.5	0.0	0.0	23.0	10.0	13.0

of vertex j in graph G then

$$p_{ij} = C_G(j) - C_{G-i}(j) \quad (8)$$

where $C_{G-i}(i)$ is defined as zero. Each p_{ij} indicates the contribution of node i to node j 's centrality. The row sums of P give the total centrality of node i , while the diagonal value, p_{ii} , gives the endogenous centrality of node i . The sum of all off-diagonal values in a row gives the row node's exogenous centrality. The matrix can be interpreted as a dependency matrix where each cell p_{ij} indicates the degree to which j is dependent on i for centrality.

As an example, consider the graph in Fig. 2. Table 6 gives the matrix of contributions P using betweenness as our base measure of endogenous centrality. We can see that the pendant node 6 is a big source (4 points) of betweenness centrality for node 3, since all paths from node 6 to all other nodes pass through 3. It is interesting to note that node 4's contribution to node 1's centrality is a negative -3, indicating 4's presence takes strongly away from 1's betweenness. Without node 4, nodes 3 and 6 need 4 to reach 2 and 5. With node 4, there is a shorter path to node 5 as well as an alternative (equally short) path to node 2.

4. Conclusion

In this paper we have presented a general method of constructing centrality measures using almost any graph invariant. The essential idea is to calculate a graph invariant on a graph, delete a given node and recompute the invariant. The difference between the two quantities gives the induced centrality for that node. Three key benefits are provided by this approach. First, the resulting centrality measure has a natural interpretation which is the contribution of a node to the network's overall score on the invariant. For example, if an invariant is chosen that represents network cohesion, the induced centrality based on that invariant can be interpreted as a node's contribution to the network's cohesion. The second key benefit is that we can construct a centrality measure from virtually any invariant, allowing us to tailor a measure to the specific research need at hand. For example, in the context of trying to disrupt a network by removing key nodes, Borgatti (2006) takes as his invariant a measure of network fragmentation. A measure of attractiveness for deletion is constructed by measuring the increase in fragmentation obtained by removing each node. The third advantage of induced centrality measures is that they provide a natural way to generalize node centrality to group centrality (Everett and

Borgatti, 1999), simply by calculating changes in invariants when removing sets of nodes instead of individuals.

In the second part of the paper we demonstrate that we can construct an induced centrality measure from virtually any other centrality measure by taking the sum of all centrality scores as the invariant for the induced centrality. We call this induced centrality, which measures the contribution of a node to the sum of centralities in the network, the total centrality of a node. We then decompose total centrality into an endogenous part, which is the base measure of centrality used to construct the invariant, and an exogenous part, which is the difference between the total centrality and the endogenous part. The exogenous centrality can be interpreted as the indirect contribution a node makes to the centrality of all the other nodes. We then show that we can construct a node by node contribution matrix that gives the amount of centrality contributed by each node to each other node's centrality, thereby indicating the latter node's dependency on the former node. Our examples have focused on centrality measures which are derived from shortest path type considerations, but as we have noted this technique is applicable to other variants such as maximum flow.

References

- Borgatti, S.P., 2005. Centrality and network flow. *Social Networks* 27, 55–71.
- Borgatti, S.P., 2006. Identifying sets of key players in a social network. *Computational & Mathematical Organization Theory* 12, 21–34.
- 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., Everett, M.G., 2006. A graph-theoretic framework for classifying centrality measures. *Social Networks* 28, 466–484.
- Brandes, U., 2008. On variants of shortest-path betweenness centrality and their generic computation. *Social Networks* 30, 136–145.
- Everett, M.G., Borgatti, S.P., 1999. The centrality of groups and classes. *Journal of Mathematical Sociology* 23, 181–201.
- Everett, M., Borgatti, S.P., 2005. Ego network betweenness. *Social Networks* 27, 31–38.
- Freeman, L.C., 1979. Centrality in social networks: 1. Conceptual clarification. *Social Networks* 1, 215–239.
- Freeman, L.C., Borgatti, S.P., White, D.R., 1991. Centrality in valued graphs: a measure of betweenness based on network flow. *Social Networks* 13, 141–154.
- Koschützki, D., Lehmann, K.A., Peeters, L., Richter, S., Tenfelde-Podehl, D., Zlotowski, O., 2005. Centrality indices. In: Brandes, U., Erlebach, T. (Eds.), *Network Analysis: Methodological Foundations*. Springer-Verlag, Heidelberg, Berlin, pp. 16–61.
- Padgett, J.F., Ansell, C.K., 1993. Robust action in the rise of the Medici, 1400–1434. *American Journal of Sociology* 98, 1259–1319.
- Valente, T.W., Foreman, R.K., 1998. Integration and radiality: measuring the extent of an individual's connectedness and reachability in a network. *Social Networks* 20, 89–109.