CONNECTIONS 26(2): 70-81 http://www.insna.org/Connections-Web/Volume26-2/7.Cornwell.pdf
© 2005 INSNA

A Complement-Derived Centrality Index for Disconnected Graphs[1]

Benjamin Cornwell
University of Chicago, Department of Sociology

Freeman's (1979) measure of closeness centrality is valuable in network analysis, but its use is limited to connected networks. In this paper, I describe an approach for calculating actor closeness centrality that circumvents the problem of disconnectedness. I show how the complement, G_C, of a disconnected network, G, can be used to obtain weights that transform Freeman's measure, C'_C, into a universal measure, C'_{CW}, for actors in both connected and disconnected networks. In essence, this method incorporates information about how an actor is not proximate to all other actors in a network (captured by the structure of the complement network) to weight within-component closeness. C'_{CW} has several attractive properties. Aside from being universally applicable and ranging from 0 to 1, the value of C'_{CW} equals C'_C in connected networks. Furthermore, C'_{CW} cannot reach 1 for actors in disconnected networks.

INTRODUCTION

Centrality is a much-analyzed actor-level property of social networks. Measures of centrality attempt to identify the "most important" actors in a network using nomination degree, closeness, betweenness, or some comparable notion. Perhaps the most useful and popular of these measures, Freeman's (1979) closeness centrality index, relies on geodesic distances among actors. The measure is useful because it captures independence from the control of others in terms of accessing (the resources of) others in a network, but it can only be applied to connected networks. Though several attempts to calculate actor centrality measures for disconnected networks have been made, none seem to possess as much intuitive appeal or are as substantively interpretable as the original measure for connected networks.

I begin by describing some existing actor centrality measures and by summarizing the problem of disconnectedness. I follow with a discussion of the complement of a network (the network of ties that do not exist), and how it can be used to infer properties about actor closeness centrality across disconnected components. I show that one can extend Freeman's (1979) measure to actors in a disconnected network by considering the extent of disconnectedness in conjunction with a given

[1] Correspondence concerning this manuscript should be addressed to Benjamin Cornwell, Department of Sociology, University of Chicago, 1126 E. 59th St., Chicago, Illinois 60637. Email may be sent to: cornwell@uchicago.edu., and faxes may be sent to 614-292-6687. I wish to thank Jim Moody for his always indispensable counsel, Robert Poulin and Marie-Claude Boily for technical clarification, Erin York for her meticulous reading and thoughtful suggestions, and the Connections editor and reviewers for valuable advice.

actor's position in its complement. Finally, I compare the closeness measure that I derive from the complement (which I will call C'_{CW}) to another measure, C''_{CN}, developed in previous attempts to overcome the problem of disconnectedness.

C'_{CW} has attractive properties in that it ranges from 0 to 1 (it can only reach 1 in connected networks), and it is comparable across components and networks containing different numbers of actors and levels of disconnectedness. Freeman's (1979) closeness centrality is a special case of C'_{CW} that is found in fully connected networks, since the amount of disconnectedness experienced in these networks is constant for all actors (0). *C'_{CW} reflects independence from the limits on access posed by reachable others in the network and by disconnectedness in the overall network*. This operationalization results from conceiving of centrality as a dimension of an actor's local importance that is contextualized by a broader structure.

ACTOR CENTRALITY MEASURES

Early measures of actor centrality that focused on closeness were developed by Bavelas (1950), Harary (1959), and Beauchamp (1965; see Wasserman and Faust 1994). Measures of closeness centrality developed later are somewhat more useful because they convey the minimum number of steps separating actors from others (see Hakimi 1965; Sabidussi 1966). According to Freeman (1979), the simplest and most useful actor closeness measure was developed by Sabidussi (1966). The appeal of Sabidussi's measure is that it uses geodesic distances to measure actor closeness centrality. The measure is calculated as the inverse of the sum of the distances from a given actor, *i*, to all the other actors in a network (Wasserman and Faust 1994). The formula can be expressed as:

$$Cc(n_i) = \left[\sum_{j=1}^{g} d(n_i, n_j)\right]^{-1}$$

where *g* is the number of actors in the network, and *d* is the number of links in the shortest path from actor n_i to actor n_j.

Note that Sabidussi's (1966) measure depends on the number of actors in the network. *Ceteris paribus*, networks with fewer actors will have larger closeness centrality scores, since the sum of distances (which ends up in the denominator) is greater in larger groups. However, following Beauchamp (1965), we can set the ceiling of this index to 1 by dividing g-1 by the sum of the path distances, which is the same as multiplying the inverse of the sum of the path distances by g-1 (Freeman 1979). In essence, we are dividing by the maximum possible distance. This technique standardizes the measure, making it comparable across groups of different sizes. Thus, the revised closeness centrality index is:

$$C'c(n_i) = (g-1)\,Cc(n_i)$$

While Freeman's (1979) standardized closeness centrality index is useful, it has one major drawback: it can be calculated universally only for actors in *connected* networks. Connected networks are those in which each actor in the network can reach all of the others through direct or indirect ties. There are no isolated actors in connected networks. The reason is that the distance between disconnected actors (actors who are not connected—directly or indirectly) is infinite, or undefined. Therefore, "the distance sum for every actor is infinity, and the actor closeness indices are all 0" (Wasserman and Faust 1994:185). This limitation sometimes leads researchers to compute a localized index of closeness centrality *within* connected components of networks, thereby ignoring isolated actors or entire components that exist separately elsewhere in the network. The major drawback of this method is that one cannot adequately compare actor closeness centrality across networks when at least one of those networks is disconnected (see Donninger 1986; Stephenson and Zelen 1989; Altmann 1993). Another approach is to assign a distance to all unreachable nodes. One might use some arbitrary large number, such as 1,000, which simulates infinity in a more mathematically manageable way. Alternately, a researcher could use a value such as one plus the diameter (the largest observed distance) to imitate the distance between unreachable nodes (e.g., see Valente and Forman 1998). This approach is probably more realistic, particularly when the nodes are people, because it assumes that we are all connected to each other in a relatively modest number of steps (Milgram 1967; Watts 1999). These approaches are useful fixes but require imputing somewhat arbitrary values.

Attempts at Bypassing Graph Disconnectedness

Poulin et al. (2000) provide a detailed account of several options that exist for circumventing the problem of accounting for infinite distances among disconnected actors when calculating centrality. Stephenson and Zelen's (1989; see also Altmann 1993) S-Z index of centrality (C'_{Inf}) is one example. While such approaches often have reasonably good discriminant power (i.e., the ability to recognize obscure differences in closeness within components), they still suffer from some problems with comparability among actors in different components/networks. Poulin et al. (2000) argue that other limitations of these methods include the fact that they are computationally difficult (e.g., inversing large matrices), and can be time- and memory-consuming when analyzing large networks. Thus, they propose a centrality measure, C''_{CN}, for disconnected networks that overcomes these problems. Theirs is a mapping-based method that uses a *c*umulative *n*omination scheme to explore all of the possible paths between pairs of individuals in a network. Their idea is as follows:

> Initially . . ., every individual gets one nomination. Then, after the first round of nominations (stage 1), each individual gets additional nominations from their contacts, weighted by the number of nominations their contacts already have, which is 1 at this stage. Thus, a contact with many nominations will be considered more important than a contact with only few nominations. The process is repeated such that [an] individual cumulates nominations every new round . . .. After a while, individuals are ordered in the function of their cumulated number of nominations; more central individuals having cumulated more nominations. (Pp 199-200).

They normalize the measure (at this stage referred to as C_{CN}) by the level of nomination activity within the component of interest (a measure of the rate of growth of the cumulated nominations), yielding a new measure, C'_{CN}. Finally, they improve the discriminatory power of the score for each actor by multiplying it by its component size, yielding C''_{CN}.

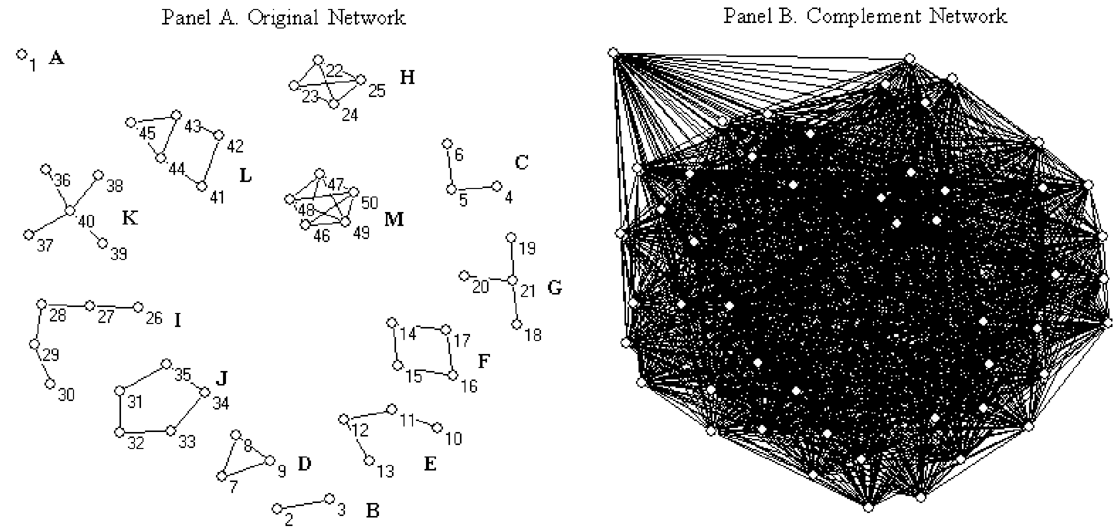
The application of Poulin et al.'s (2000) measure to a 50-person contact network is shown in Table 1, which is discussed in greater detail in the section on discriminatory power. The network is sparsely connected, and contains multiple autonomous subgroups of varying sizes, as shown in Panel A of Figure 1 (see Poulin et al. for a more visually organized layout of the network). The complement of this network, which is exceedingly busy and difficult to comprehend, is depicted in Panel B.

Despite the usefulness of the Poulin et al. (2000) measure for overcoming the problem of disconnectedness, it suffers from a lack of comparability and interpretability.[2] One problem is that the comparability of the measure in various networks is questionable. It is difficult to tell what is a "high" or "low" centrality score in this scheme. That is, the measure does not appear to be standardized within a certain range because its theoretical maximum has not been determined. [3] Relatedly, C''_{CN}

[2] See Poulin et al. (2000) for a detailed discussion of the limitations of the other measures.

[3] I thank Marie-Claude Boily and Robert Poulin for providing clarification of this point in a personal communication.

does not permit direct interpretation. The measure provides little language that can be used to discuss the relative centrality of one point versus another. Overall, while cumulative nomination mapping may be a useful and computationally convenient tool for examining disconnected networks it is neither comparable across disconnected and connected networks nor substantively interpretable.



Note: The network appearing in Panel A is from Poulin et al. (2000), which includes a clearer portrayal of the network and its components. In these panels, nodes are in fixed positions, determined by a spring embedding algorithm in NetDraw (Borgatti 2002). For example, the node in the upper left hand corner of Panel A (node 1) is the same as the node in the upper left hand corner of Panel B. Panel A provides both node labels(numbers) and names of the components in which the nodes appear (bolded etters,appearing to the right of the components). Node and component labels are not included inPanel B due to lack of space.

Figure 1. A Disconnected Network of 50 Actors, and its Complement

Below, I propose a method of calculating a closeness centrality measure which applies to actors in both connected and disconnected symmetric networks. This new measure is based on Freeman's (1979) original closeness centrality index, but it includes information about how an actor is *not* connected to others to infer its closeness centrality with respect to the entire network. In effect, this method generates a value representing the extent of disconnectedness in a network, which can be used to weight an actor's centrality within connected components. In the next section, I describe how information concerning an actor's disconnectedness can be generated and analyzed using standard network methods.

Complement Weighting

Examining the connections that *do not* exist in a network is potentially as interesting and useful as examining the connections that *do* exist. This idea, as applied to directed networks, is mentioned briefly by Wasserman and Faust (1994):

> . . . the complement of a digraph might be used to represent the absence of a tie, or as *not* the relation. For example, in the digraph representing the relation of friendship the arc $< n_i, n_j >$ means *i* "chooses" *j* as a friend. In the digraph representing the complement of the relation of friendship, the arc $< n_i, n_j >$ means *i* "does not choose" *j* as a friend. (P. 135.)

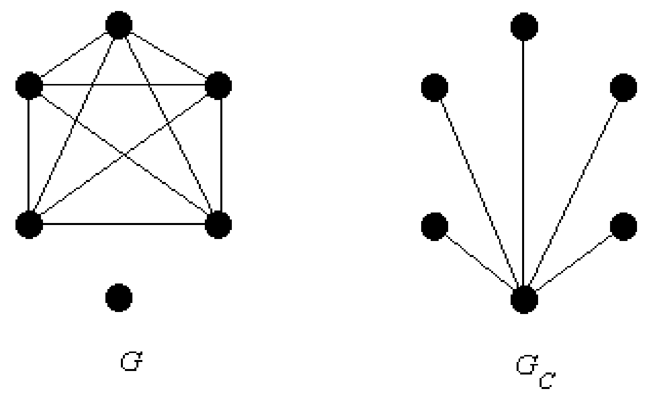


Figure 2. A Disconnected Graph, G, and its Complement, G_C.

In other words, the complement is a network of non-existent ties among actors. Thus, G_C has the same number of nodes as G, but has the exact opposite pattern of ties. If actor i is connected to actor j in an undirected network G, it is not connected to j in G_C. Unfortunately, the above passage from Wasserman and Faust (1994) is the most in-depth discussion of the complement that I could find in current social network research. Possibly because it is simply the "opposite" of a given network, the properties of the complement have not been fully explored. However, as we will see, this very fact makes the complement useful for overcoming limitations in G with respect to at least one unique network measure—actor closeness centrality. Figure 2 provides a clear example of the relationship between a given disconnected network, G, and its connected complement, G_C.

An important point is that, *if a symmetric network is disconnected, its complement will be connected*, because any isolate (or groups of disconnected actors) will have ties to *all* actors in other components, linking those actors together indirectly. Other isolates or actors in other components will have the same pattern of ties, linking them indirectly to each other. Applying this idea to the problem of calculating actor closeness centrality in disconnected networks, one can work backward from knowledge concerning which actors are and which actors are not central in the pattern of *non-existing ties* in a network to determine who is not and who is, respectively, central in the pattern of *existing* ties in a network. As I will show, one can calculate actor closeness centrality for nodes in all networks, including disconnected ones, by considering the extent to which actors contribute to a network's overall disconnectedness, and using that value to weight their within-component closeness centrality scores.

Obtaining the Complement Graph

Obtaining G_C is easily accomplished using matrix algebra. We simply solve:

$$G_C = 1 - G,$$

where

G_C = The $n \times n$ adjacency matrix of the complement of the disconnected graph;

1 = An $n \times n$ matrix containing all ones;

G = The adjacency matrix of the disconnected graph.

In a binary matrix, this operation effectively transforms 0s into 1s, and all 1s into 0s, since 1 - 0 = 1 and 1 – 1 = 0. Thus, this method works for undirected as well as directed networks. Any statistical program capable of solving matrix algebra expressions can generate the complement matrix with ease.

Complement-Weighted Centrality

The idea behind a complement-weighted centrality measure is to adjust observed closeness among connected actors by non-closeness among all actors. First, we calculate Freeman's (1979) centrality index for actors within connected components (see above), and set those values aside. Next, we evaluate the distance relationships among actors in the complement. The distance matrix for G_C can be obtained relatively easily using network analysis programs (e.g., Borgatti et al. 2002). The complement's distance matrix allows one to consider how close actors are to each other in the "opposite" reality. If an actor, *i*, is not reachable to an actor, *j*, in *G*, *i* is reachable to *j* in the connected complement, G_C. Distance in the complement represents what we might refer to as anti-distance, or distance to others given the exact counterfactual of the reality we observe in the original network. The structure of an anti-distance matrix, as derived from the complement, is unique and cannot be determined simply by reversing distances in the original networks of ties. It is this pattern of anti-distances that I use to derive weights for observed closeness among connected actors.

Recall that Freeman's (1979) closeness centrality measure is calculated as the inverse of the sum of the distances from actor *i* to all the other actors in the original network, *G*, normalized by *i*'s component size. Likewise, to obtain the complement-weighted measure, C_{CW}, I first calculate the inverse of the sum of the *anti*-distances for each actor, given in the complement's distance matrix. Those in small components in *G* will be more central (less anti-distant) in G_C, so I normalize the inverse of *i*'s anti-distances to others by multiplying it by the number of actors outside of *i*'s connected component in *G* (because distances to these persons will always equal 1 in the complement). Theoretically, an actor retains centrality status in *G* by being non-central in the complement, G_C. If *i* is central in this counterfactual reality, which presents the structure of non-connectedness, then *i*'s centrality in *G* should be adjusted downward to account for *i*'s actual disconnectedness. Because the weight should be a relatively low number for cases that are disconnected in G_C, I subtract the normalized complement centrality measure from 1 before using it to weight Freeman's measure. The final step is to multiply this complement-derived weight by Freeman's measure of centrality. Thus, the formula for the complement-derived closeness centrality index for a given actor, *i*, is:

$$C'_{CW}(n_i) = \left(1 - \left(\left[\sum_{j=1}^{g_C} d_{G_C}(n_i, n_j)\right]^{-1} (g_C - g(n_i))\right)\right) C'_C(n_i)$$

where

g_C = the number of nodes in G_C, the complement of *G*;

$d_{G_C}(n_i, n_j)$ = the distance between n_i and n_j in G_C;

$g(n_i)$ = the number of nodes in *i*'s component in *G*; and

$C'_C(n_i)$ = Freeman's within-component closeness centrality score for actor n_i.

This formula makes the complement-derived measure comparable to Freeman's (1979) measure for actors in connected networks. We can interpret a complement-weighted closeness centrality value for any given actor, *i*, as the closeness centrality of *i*, adjusting for the non-closeness of *i*. Freeman's centrality measure for actors in connected networks can be interpreted in the same way. Thus, C'_{CW} provides a way of comparing the centrality of actors in connected and disconnected networks against each other.

Table 1. Summary Statistics for Nodes in the Disconnected Network Pictured in Figure 1.

Node	Comp. Size	C'_C	Rank (C'_C)	C''_{CN}	Rank (C''_{CN})	C'_{CW}	Rank (C'_{CW})	Node	Comp. Size	C'_C	Rank (C'_C)	C''_{CN}	Rank (C''_{CN})	C'_{CW}	Rank (C'_{CW})
1	1	.000	9	.020	18	.000	12	26	5	.400	8	.183	13	.040	11
2	2	1.000	1	.080	17	.040	11	27	5	.571	6	.317	6	.067	7
3	2	1.000	1	.080	17	.040	11	28	5	.667	4	.366	3	.078	4
4	3	.667	4	.127	16	.040	10	29	5	.571	6	.317	6	.067	7
5	3	1.000	1	.180	14	.078	5	30	5	.400	8	.183	13	.040	11
6	3	.667	4	.127	16	.040	10	31	5	.667	4	.300	7	.078	4
7	3	1.000	1	.180	14	.078	5	32	5	.667	4	.300	7	.078	4
8	3	1.000	1	.180	14	.078	5	33	5	.667	4	.300	7	.078	4
9	3	1.000	1	.180	14	.078	5	34	5	.667	4	.300	7	.078	4
10	4	.500	7	.160	15	.040	11	35	5	.667	4	.300	7	.078	4
11	4	.750	3	.259	9	.074	6	36	5	.571	6	.250	10	.057	8
12	4	.750	3	.259	9	.074	6	37	5	.571	6	.250	10	.057	8
13	4	.500	7	.160	15	.040	11	38	5	.571	6	.250	10	.057	8
14	4	.750	3	.240	11	.074	6	39	5	.571	6	.250	10	.057	8
15	4	.750	3	.240	11	.074	6	40	5	1.000	1	.500	1	.151	1
16	4	.750	3	.240	11	.074	6	41	5	.667	4	.283	8	.078	4
17	4	.750	3	.240	11	.074	6	42	5	.667	4	.283	8	.078	4
18	4	.600	5	.185	12	.048	9	43	5	.800	2	.419	2	.108	3
19	4	.600	5	.185	12	.048	9	44	5	.800	2	.419	2	.108	3
20	4	.600	5	.185	12	.048	9	45	5	.667	4	.338	4	.078	4
21	4	1.000	1	.320	5	.115	2	46	5	1.000	1	.500	1	.151	1
22	4	1.000	1	.320	5	.115	2	47	5	1.000	1	.500	1	.151	1
23	4	1.000	1	.320	5	.115	2	48	5	1.000	1	.500	1	.151	1
24	4	1.000	1	.320	5	.115	2	49	5	1.000	1	.500	1	.151	1
25	4	1.000	1	.320	5	.115	2	50	5	1.000	1	.500	1	.151	1

Note: Measures for C''_{CN} and C'_C are from Poulin et al. (2000).

Properties of the Complement-Derived Centrality Measure

One of the most useful properties of C'_{CW} is that it can be calculated for any node in any symmetric network. Its use is limited to neither connected nor disconnected networks. When a network is connected, the complement-derived weight will always equal 1, because there are no outside-component nodes in the parenthetical expression. Thus, *when a network is connected, the complement-weighted closeness centrality measure will equal Freeman's (1979) within-component actor closeness centrality measure*. In this section, I describe some of the additional properties, including the range and discriminatory power of the complement-weighted centrality measure, C'_{CW}. I also discuss its interpretation and comparability across networks.

Range

C'_{CW} is attractive because it is standardized and interpretable. Isolates receive a centrality score score of 0, and the centrality score of nodes that are maximally connected within their components depends on the degree of disconnectedness elsewhere in the entire network, *G*. In the network shown in Figure 1, Panel A, the largest observed complement-weighted centrality value is .151. This centrality score is modest—despite the fact that the nodes that receive that score (nodes 40, 46 - 50) are maximally connected within their components—because of the extensive disconnectedness characterizing the rest of the network. As connectedness within an actor's component (in an otherwise disconnected network) increases, that actor's centrality also increases, but can never reach a value of 1. In fact, *the measure only equals 1 when there are no isolates or separate components in the network*. This makes sense, because an actor should not achieve a centrality score of 1 unless it is directly connected to *every* node in a network. Imagine that this network contains one complete component of size 999 and one isolate. In this case, the centrality of any of the actors in the connected component would be .999 (not 1, as Freeman's index would indicate), because some of the possible ties (999 of them, to be exact) are non-existent in this network.

Table 1 (above) shows how this measure applies to the disconnected graph given in Poulin et al. (2000), and Figure 3 plots them together on a scatter plot. As you can see, the measures are closely related (r = .97), but are slightly different in several respects. These discriminatory differences are described below.

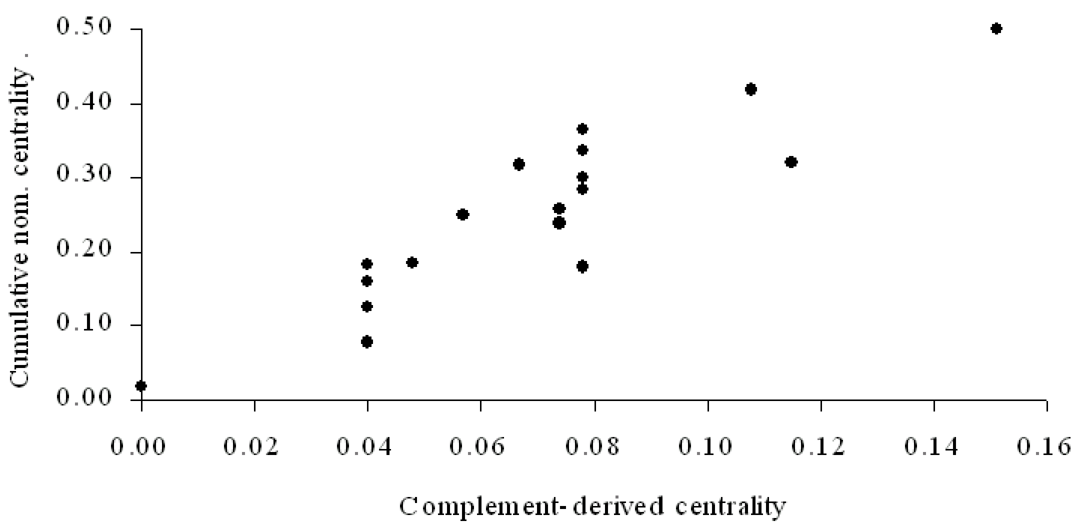


Figure 3. Scatterplot of Poulin et al.'s (2000) Centrality Measure against the Complement-Weighted Closeness Centrality Measure for a Disconnected 50-Actor Network

Discriminatory Power

Compare Freeman's (1979) C'_C, Poulin et al.'s (2000) C''_{CN}, and the complement-weighted centrality measures shown in Table 1. The differences in the node centrality ranks generated by each one are notable. For instance, while nodes 2 and 3 take on the maximum possible centrality values when using Freeman's measure within components, they have the 11th and 17th highest centrality values, respectively, when using the cumulative nomination mapping and complement-weighted methods.[4] This is a direct result of bringing the measure out of the within-component context.

One of the most striking (and I argue, useful) aspects of the complement-weighted measure is that the centrality values are small given the 0 - 1 range, and thus the differences among the raw complement-weighted values obtained for this network are small (ranging from .000 to .151). This is a direct result of the level of disconnectedness in the rest of the network as experienced by each node. After all, this is a 50-node network, containing only 51 lines, and we are dealing with components that are small relative to the total number of nodes in the overall network. This is an attractive property of C'_{CW} because it forces us to consider the overall properties of the network in determining an actor's centrality.

Because this method is based on geodesic distances, the relative centrality rank of a given node within its component is the same as indicated by Freeman's (1979) measure. Poulin et al.'s (2000) C''_{CN} does not always correspond to C'_C within components. For instance, C'_{CW} assigns the same centrality to nodes 41, 42, and 45, whereas 45 takes on a larger C''_{CN} value than that assigned to 41 and 42. Thus, while the complement-weighted method also considers disconnectedness in a network, it retains the emphasis on geodesic distance found in classic treatments of centrality.

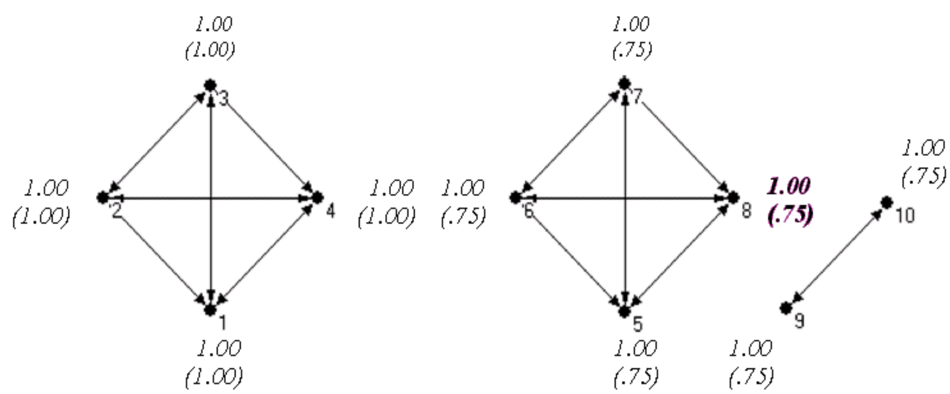
There are some other discriminatory differences between the complement-weighted centrality measure and C''_{CN}. Both measures allow nodes within larger components to take on values that are lower than some nodes in smaller components. For instance, both measures indicate that nodes 10 and 13 are less central than nodes 7 thru 9, despite the fact that the former two are in a larger component. However, the complement-weighted method allows this to happen more frequently. For instance, according to C'_{CW}, 5 is more central than 30, whereas it is less central according to C''_{CN}. This occurs because the complement-weighted method places greater emphasis on *closeness*. Node 30's lack of closeness to others within its component is weighted more heavily than its general connectedness to them, while node 5's closeness to the others in its component is weighted more heavily than its lack of connectedness to others in the network.

The final difference in the discriminatory powers of C'_{CW} and C''_{CN} is revealed in an examination of centrality of nodes in component *J* (31 thru 35) versus that of node 28. C''_{CN} gives more weight to node 28 (by .066). Freeman's (1979) measure does not discriminate between nodes 31 thru 35 and node 28, and neither does C'_{CW}. The average distance of node 28 to others in component *I* is the same as those in *J* with each other (avg. = 1.5) and at the same level of consistency (captured by the standard deviation, which is .578). The only difference is that node 28 binds the other nodes in *I* together, whereas those in *J* are similarly connected to each other. If we were to impute a direct link between nodes 26 and 30, components *I* and *J* would be identical. Thus, C''_{CN} gives weight to the structural role of node *28 relative to the other nodes within I* (resembling more of a betweenness centrality measure), which is somewhat irrelevant if what we are interested in is its geodesic closeness centrality.

[4] The reader should note that the rank scales are different for each measure.

When to Use the Complement-Weighted Centrality Measure

Versus C'_C Calculated within Components. As I will show, the choice between Freeman's (1979) closeness centrality measure, C'_C, and C'_{CW} hinges on whether one is interested in the local or universal conditions of nodes. C'_C is calculated as the inverse of the sum of the distances between *i* and the actors to whom *i* is connected divided by the number of other actors in *i*'s component. Because it is based on geodesic distances, the measure cannot be calculated for disconnected networks. C CW is appropriate when one wishes to calculate actor closeness centrality in any network, connected or disconnected. The complement-weighted method can be used for comparisons between actors across components, disconnected networks, connected networks, and even between a node in a connected network and a node in a disconnected network. The measure can be compared across components/networks of different sizes and levels of disconnectedness. It has the same range as Freeman's centrality index. In fact, C'_C is a special case of C'_{CW} , occurring when an actor is at least indirectly connected to all others in a network. C'_{CW} simply has the added benefit of allowing us to incorporate a node's lack of connectedness in determining centrality in disconnected networks.



Note: C_{CW} values are presented in parentheses below C_C values

Figure 4. C_C and C_{CW} Centrality Scores for a Connected and a Disconnected Network

Figure 4 compares two simple networks: one connected network made up of four actors and one disconnected network made up of six actors (split up into a maximally connected component containing four actors and one dyad). C'_C can be calculated for actors in both networks—though it must be calculated within components in the disconnected network—and is presented next to each node. C'_{CW} is calculated for actors in both, and is presented in parentheses under C'_C . As you can see, the complement-weighted measure conveys the centrality of a given actor, *i*, that is due to how close *i* is to actors to whom *i* is connected and relative to its level of disconnectedness to everyone else in the network.

This complement-weighted method allows us to improve our understanding of the extent to which actors can independently access all other actors in a network, which takes us away from mere local structural context. We can see that actors 1 and 5 have direct and indirect access to the resources of the same number of actors (four) at the same distance (one step away from the other three actors). To make the comparison across networks, one simply uses C'_{CW} for all actors. Using this measure, we see that actor 1 has more independent access to the resources held by others in his or her network than actor 5 because actor 5 also must deal with not having any access to the resources of actors 9 and 10.

Actors 9 and 10 have relatively little access independence, which is reflected in the relatively low centrality scores. With C'_C, we can only assess centrality relative to others in one's own component. It is not as useful when we are considering conditions across subgroups. Therefore, C'_C values calculated within components might be more appropriate for studying context-specific conditions, such as relative deprivation (Davis 1959; Gurr 1970; Runciman 1966), whereas C'_{CW} is better for understanding overall conditions.

Versus C''_{CN}. The centrality measure proposed by Poulin et al. (2000) involves a cumulative mapping technique, which weights the number of nominations an actor receives from contacts by the number of nominations those contacts receive, and subsequently weights those values by the amount of nomination activity in the component. Making a choice between the complement-derived C'_{CW} and C''_{CN} boils down to the relevance of the measures to one's substantive question. Both measures are useful because they overcome the problem of disconnectedness. C'_{CW} is more appropriate for those interested in the combined effects of within-component closeness and disconnectedness outside of components, but again with less emphasis placed on the local context.

C''_{CN} is perhaps most useful when one is interested in determining the effectiveness of a given node in distributing resources to others (though not necessarily in a shorter time frame). *C''_{CN} gives more weight to general connectedness, whereas C'_{CW} is more concerned with closeness*. This property also could have implications for diffusion. Closeness is important when the accuracy of a message is crucial, particularly when message accuracy breaks down with each successive relay. In such a case, the message is best left in the hands of an actor who can insure that the information is diffused using the *minimum number of steps* possible. In this case, the choice of node 36 over node 27 in Figure 1, for instance, is clear, since information beginning with node 27 takes three steps to reach node 3 (requiring both nodes 28 and 29 to "forward" the message along). Starting with node 36, though, the remaining nodes in component *K* can receive the message in two steps, at most. A logical implication is that C'_{CW} is perhaps better suited for analyzing a network where the quality of the information being relayed decays substantially with each step. This property of centrality might be relevant when studying diffusion among nodes that are inaccurate or prone to malfunction, or when the information conveyed is intimate in nature, thus translating best among intimate contacts. The use of C''_{CN} is advisable where we assume that all nodes are equally efficient and accurate, and where the researcher is concerned about modeling information relay to as many others as possible without constraints on time or rate. Relatedly, C'_{CW} is probably the better choice if we are interested in evaluating the influence status of a node on others, where independence from others' control over information flow in a given structure is relevant (see Freeman 1979).

Most importantly, it does not appear that the cumulative nomination mapping technique generates estimates of centrality that are comparable across networks. I have already shown that C''_{CN} is not comparable across networks because it assigns values based on the average centrality observed within a network. This average will vary from network to network. Thus, to say that a node has a higher C''_{CN} value in one network than a given node in another does not really convey to us which actor is better off overall, as they are evaluated relative to others in their respective networks.

CONCLUSION

The complement of a disconnected network can be employed to overcome problems with calculating measures related to closeness. Measures that rely on geodesic distances are limited to connected networks because the theoretic distance between two disconnected nodes is not defined. This paper describes how an existing and popular closeness centrality measure for connected networks—Freeman's (1979) measure of actor closeness centrality—can be extended to disconnected networks by adjusting for non-connectedness. The new measure proposed here, C'_{CW}, has several attractive properties. For one, it ranges from 0 to 1 and is comparable across connected and disconnected components and networks. C'_{CW} can be interpreted as the amount of closeness centrality that an actor, *i*, has in a network given the amount of disconnectedness in the network.

The method described here provides one way of circumventing the problem of disconnectedness while retaining the basic principles behind the measure of interest. The complement provides us with a snapshot of the counterfactual of an observed relational reality, which makes it attractive as a weighting tool. It is best suited to issues involving universal access to resources in a network other than local access within connected components.

REFERENCES

Altmann, M. 1993. Reinterpreting Network Measures for Models of Disease Transmission. *Social Networks,* 15:1-17.

Bavelas, A. 1950. Communication Patterns in Task-Oriented Groups. *Journal of the Acoustical Society of America,* 22:271-82.

Beauchamp, M.A. 1965. An Improved Index of Centrality. *Behavioral Science,* 10:161-63.

Bonacich, P. 1972. Factoring and Weighting Approaches to Status Scores and Clique Identification. *Journal of Mathematical Sociology,* 2:113-20.

Bonacich, P. 1987. Power and Centrality: A Family of Measures. *American Journal of Sociology,* 92:1170-82.

Borgatti, S., M. G. Everett, and L. C. Freeman. 1999. *UCINET V for Windows: Software for Social Network Analysis*. Vers. 5.2.0.1. Natick, MA: Analytic Technologies.

Donninger, C. 1986. The Distribution of Centrality in Social Networks. *Social Networks,* 8:191-203.

Freeman, L.C. 1979. Centrality in Social Network: 1. Conceptual Clarification. *Social Networks,* 1:215-39.

Hakimi, S.L. 1965. Optimum Locations of Switching Centers and the Absolute Centers and Medians of a Graph. *Operations Research,* 12:450-9.

Harary, F. 1959. Status and Contrastatus. *Sociometry,* 22:23-43.

Moody, James. 1999. SPAN: *SAS Programs for Analyzing Networks. User's Manual*. University of North Carolina at Chapel Hill (164 pp).

Poulin, R., M.-C. Boily, and B.R. Mâsse. 2000. Dynamical Systems to Define Centrality in Social Networks. *Social Networks,* 22:187-220.

Sabidussi, G. 1966. The Centrality Index of a Graph. *Psychometrika,* 31:581-603.

Stephenson, K., and M. Zelen. 1989. Rethinking Centrality: Methods and Examples. *Social Networks,* 11:1-37.

Wasserman, S., and K. Faust. 1994. *Social Network Analysis: Methods and Applications*. New York: Cambridge University Press.