Social Networks 32 (2010) 212–220

Contents lists available at ScienceDirect

Social Networks

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

# Bridging: Locating critical connectors in a network

Thomas W. Valente*, Kayo Fujimoto

*Institute for Prevention Research, Department of Preventive Medicine, Keck School of Medicine, University of Southern California, 1000 Fremont Ave, Bldg A Room 5110, Alhambra, CA 91803, United States*

ARTICLE INFO

*Keywords:*
Bridges
Bridging
Strength of weak ties
Disease transmission
Behavior change
Opinion leaders

ABSTRACT

This paper proposes several measures for bridging in networks derived from Granovetter's (1973) insight that links which reduce distances in a network are important structural bridges. Bridging is calculated by systematically deleting links and calculating the resultant changes in network cohesion (measured as the inverse average path length). The average change for each node's links provides an individual level measure of bridging. We also present a normalized version which controls for network size and a network-level bridging index. Bridging properties are demonstrated on hypothetical networks, empirical networks, and a set of 100 randomly generated networks to show how the bridging measure correlates with existing network measures such as degree, personal network density, constraint, closeness centrality, betweenness centrality, and vitality. Bridging and the accompanying methodology provide a family of new network measures useful for studying network structure, network dynamics, and network effects on substantive behavioral phenomenon.



The network analysis field has devoted considerable energy developing methods for identifying central nodes in a network which are important to diffusion and other actions that occur on networks (Borgatti and Everett, 2006). In contrast, Granovetter (1973) introduced the concept of bridging which emphasized the importance of structural bridges for diffusion. According to Granovetter (1973, 1982), bridges reduce the overall distance between individuals in a network, enabling information to spread more rapidly throughout a network. The over-emphasis on identifying central nodes has led to the creation of many centrality measures with comparatively less attention given to measures for bridging. Further, most (perhaps all) of the centrality measures developed to date are strongly correlated with a node's degree, its number of direct links.

As Fig. 1 illustrates, there are many centrality measures created to identify important nodes in a network. Degree is a local measure calculated by counting the number of links for each node. Betweenness and closeness (as well as other measures) are global measures calculated using information from the entire network. There is one measure for bridging, constraint, which is calculated using local information only. There are no measures of bridging calculated using complete network information.

In this paper we propose measures for bridging using complete network data that are independent of degree. There are at least three reasons these measures may be useful. First, bridging individuals with few links might act as more efficient diffusion agents than individuals of high degree because they have fewer relationships over which to persuade others (Holme and Ghoshal, 2008). People who are in contact with many others may have less capacity to persuade any one individual because they must spread their persuasive energies across many people, thus diminishing their capacity to persuade any one person. A critical node with few others to persuade may devote more energy to persuading those others and hence be a more effective change agent.

Second, bridging individuals may be more receptive to behavior change and more likely to be persuaded by targeted communications. Individuals with high degree occupy prominent and visible positions in the network. This prominence can inhibit behavior change because prominent individuals need to support the status quo in order to maintain their positions of prominence (Becker, 1970; Cancian, 1979). Bridging individuals, in contrast, have fewer direct contacts and therefore less direct pressure to support prevailing norms and behaviors, and hence perhaps more susceptible to change.

Finally, it may be that occupying a bridging position is indicative of attitudinal and behavioral dispositions such as being open to new ideas and practices. Many studies have identified associations between degree and attitudes and behaviors. Degree is often equated with opinion leadership and many studies conducted to determine correlates of opinion leadership (Rogers, 2003). Similarly, it is reasonable to expect that there may be attitudinal and behavioral correlates of bridging. Burt (1992) has shown that spanning structural holes accrues advantages to managers. We suspect that the graph theoretic measures of bridging proposed here will

* Corresponding author. Tel.: +1 626 457 4139; fax: +1 626 457 6699.
*E-mail address:* tvalente@usc.edu (T.W. Valente).

0378-8733/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.socnet.2010.03.003

| | | Measure | |
|---|---|---|---|
| | | Centrality | Bridging |
| Network Information | Local | Degree | Constraint |
| | Global | Closeness Betweenness | Bridging based on link deletions |

**Fig. 1.** Existing nodal measures of structural position can be classified by whether they measure centrality or bridging and whether they use local or complete information.

also be associated with such advantages, or with other individual characteristics.

In sum, bridging individuals may be more effective at changing others, more open to change themselves, and intrinsically interesting to identify. The efficacy of and susceptibility to behavior change of bridging individuals may be a function of the innovation's attributes such as its cultural or normative compatibility. Innovations that are radically new, less compatible with cultural norms, or have the potential to change power dynamics within a community or organization may be more readily embraced by bridging individuals than leaders because leaders have a vested interest in maintaining the status quo.

To measure bridging, Burt developed the concept of structural holes and argued that individuals who spann structural holes form bridges in the network (Burt, 1992). This spanning function was measured by constraint which is the degree a person's links (ego network) are to people not connected to one another. Constraint calculates bridging using network data from the individual's local or personal network rather than considering the structure of the complete network. Given the importance of bridging behavior to interpretation of network structure and diffusion, it seems warranted to develop measures of bridging based on complete network information.

Doreian and Fujimoto (2004) proposed three methods for identifying linking-pin organizations of (1) blockmodeling, (2) centrality/centralization, and (3) cut-points/sets of the graph. By using empirical data, they found that the necessary (but not sufficient) condition for a node to be a linking-pin organization is that it be a singleton in a position of a blockmodel image network, and further if it is a cut-vertex in the image network, it is a strong linking-pin organization.

In graph theory, two concepts have been used to describe bridging. A cut-point is a node whose removal disconnects a network and a bridge is a link whose removal disconnects the network (Harary et al., 1965). These measures (cut-point and bridge) have only been used to identify one or few nodes and links in a network and do not provide individual measures that can be used in subsequent behavioral analyses. Further, these two measures are very limited definitions of the much broader concept of bridging. The bridging measures described in this paper calculate the change in average path length of the network when each link is removed. These values are then summarized for each node. Before presenting the mathematical derivation of the measures, we provide some background on the use of node and link deletion in networks.

*Link and node deletion*. Many researchers have used link deletion for blockmodeling and subgroup identification (Everett, 1983; Schwartz and Sprinzen, 1984; Borgatti et al., 1990). Another use of link deletion and addition has been in the analysis and measurement of small worlds (Watts, 1999). Bridges make networks small world networks by reducing the overall path length between nodes in a network. Another example of link deletion is the Girvan and Newman (2002) procedure for removing links and recalculating network properties to define community structure. Motter et al. (2002) removed selected links from various prototypical networks to demonstrate network vulnerability. White and Harary (2001) and Moody and White (2003) proposed deleting links to assess the overall cohesiveness of a network. These link deletion analyses remove links based on some criterion and then calculate a network-level property, typically until an optimal or desired level of some network-level outcome is reached.

In addition to link deletion, some researchers have deleted nodes to assess their importance. Koschützki et al. (2005) proposed a vitality measure calculated by removing nodes and calculating change in closeness centrality (Koschützki et al., 2005, p. 36). Similarly, flow betweenness is calculated by identifying the flow through a node divided by total flows in the network with that node removed (Freeman et al., 1991). Borgatti's (2006) Key Player concepts and algorithms also use node removal to find sets of nodes that optimally span the network. Node deletion measures (vitality, flow betweenness, key players) differ from the bridging measures proposed in this paper since they use node removal not link removal. Link removal is quite distinct and more versatile than node removal.

The present approach differs from these prior techniques in at least two ways: (1) all links in the network are systematically removed, and (2) the resultant change in a network-level measure is used to characterize nodes rather than the link or network. Conceptually the proposed measure is similar to Borgatti's (2006) Key Player analysis with the difference being that Key Player uses node deletion while the bridging measures proposed here use link deletion and then aggregates the changes to the nodes. The main contribution of the present approach is that it provides an individual measure of the strategic function of a node's links. The calculation involves taking the average of each node's link changes, not the sum, and this average captures the importance of each person's position; and unlike other positional measures is independent of the node's degree.

*The measure*. As in other studies (Borgatti, 2006), the network property of interest is the overall cohesion in the network defined as (Freeman, 1979):

$$C = \frac{\sum 1/d_{ij}}{N(N-1)} \quad (i \neq j) \tag{1}$$

where $d_{ij}$ is the geodesic distance between dyads. Disconnected dyads are given a value of infinity, and therefore, $1/d_{ij}$ is equal to $1/\infty$ which reduces to zero. For the purposes of this derivation, it is useful to rewrite the common matrix representation of cohesion into its row/column constituent elements:

$$C = \frac{\sum_i^N \sum_j^{N-1} 1/d_{ij}}{N(N-1)} = \frac{1}{N}\sum_i^N \frac{1}{N-1}\sum_j^{N-1} \frac{1}{d_{ij}} \tag{2}$$

The value of the reciprocal of the geodesic distance ranges from 0 to 1, assuming unreachable nodes are infinitely separated. Thus, the maximum value of the inner sum is unity (the $N-1$ reflects the fact that the node cannot have a link to itself). Likewise, the outer (leftmost) sum has a maximum possible value of unity. To calculate change from link deletion we have:

$$\begin{aligned} \Delta C_{ij} &= C - C_{ij'} = \left( \frac{1}{N}\sum_i^N \frac{1}{N-1}\sum_j^{N-1} \frac{1}{d_{ij}} \right) - \left( \frac{1}{N}\sum_i^N \frac{1}{N-1}\sum_j^{N-1} \frac{1}{d_{ij'}} \right) \\ &= \frac{1}{N}\sum_i^N \frac{1}{N-1}\sum_j^{N-1} \left( \frac{1}{d_{ij}} - \frac{1}{d_{ij'}} \right) \end{aligned} \tag{3}$$

where $C$ is an overall cohesion, $C_{ij'}$ is the cohesion of the network without a link from $i$ to $j$, and $\Delta C_{ij}$ the difference in cohesion when the link from $i$ to $j$ is removed. Given that the removal of an existing link increases (or leaves unchanged) the geodesic distance to the remaining connected nodes, this implies that the result of $\Delta C_{ij}$ is zero or a manifestly positive number. Furthermore, the upper limit of $\Delta C_{ij}$ is unity, since the limit of $C$ is also unity.

Borgatti (2006) also introduced a node-based index, whereby a node's importance is based on the sensitivity of the network to the loss of its links. Borgatti (2006) suggested several aggregation methods such as taking the maximum or the average of the links. To measure bridging we propose using the average and so the formula for bridging ($B_i$) for each actor $i$ is:

$$B_i = \frac{\sum_j^{N-1} \Delta C_{ij}}{D_i} \tag{4}$$

where $\Delta C_{ij}$ is defined above and $D_i$ is the degree of actor $i$. $B_i$ is simply the average of the differences ($\Delta C_{ij}$) for a node's links. In the case of a directed network, the measure can be calculated on both outgoing and incoming links providing a directional measure. It is also possible to simply take the maximum $\Delta C_{ij}$ of a given node's link changes,

$$B_i = \max(\Delta C_{ij}; \forall j) \tag{5}$$

and we do this for comparison purposes later during the simulation analysis. In addition to the node level measure, we wish to create a measure normalized by network size so calculations between networks of different sizes can be compared.

*Normalized bridging*. To normalize the bridging measure, we use a star network. In addition to being the most centralized network, the star network has the very desirable property when it comes to link deletions. In a star network, all link deletions change overall cohesion (or average path length) the same. This is not true for other prototypical network structures, such as a chain. Consequently, we can use the star network to determine how size affects bridging calculations irrespective of network size.

We first consider the directed network case and then show the derivation for an undirected network using an ideal star network configuration. This increases generality of this derivation and covers both directed and non-directed networks. When a link is removed from a directed network, a single dyad is affected. In a star network, by definition, this dyad must consist of the central node and one of the peripheral nodes. Thus, there are two possible cases: the link from the peripheral node to the central node is removed, or the link from the central node to the peripheral node is removed. We address both cases in the following paragraphs and show how link removal affects distances from the affected peripheral node, the central node, and the other peripheral nodes (the $N-2$ other nodes not connected to the removed link).

For the link from the peripheral to the central node, the central node's distances to other nodes are not affected and only the affected peripheral node experiences changes in distances. Fig. 2(a) uses a block arrow to indicate the "reference node" being the affected peripheral node. The numbers within each node are the geodesic distance from the node indicated by the block arrow. Geodesic distances and changes are only calculated on this node since the other distances in the network are unchanged. Once the link is removed all nodes become unreachable from the affected peripheral node and are at infinite distance (i.e., $1/d_{ij} = 1/\infty$ is defined as 0). It is worth noting that the equation groups the peripheral nodes into two types: The affected peripheral node whose link is removed ($B_{perip-}$), and the rest of the peripheral nodes ($B_{perip+}$). For the peripheral nodes, the observed $\Delta C_{ij}$ is:

$$\begin{aligned}\Delta C &= (C - C_{ij'})|_{i=perip-} = \frac{1}{N}\sum_i^N \frac{1}{N-1}\sum_j^{N-1}\left(\frac{1}{d_{ij}} - \frac{1}{d_{ij'}}\right)\Bigg|_{i=perip-} \\ &= \frac{(1) + (1/2)(N-2)}{N(N-1)} = \frac{1}{2(N-1)}\end{aligned} \tag{6}$$

For the derivation of this formula, refer to Appendix A.

For links from the central to the peripheral nodes, distances between nodes are again grouped into two subsets: distance from the central node to the affected node and distances to the other peripheral nodes (the ones not part of the deleted link). Thus, we have two reference points (Fig. 2(b) and (c)) that will need to be combined to obtain the total change in reachability and so $\Delta C_{ij}$ is the sum of these two reference points. Again, the numbers within each node are the geodesic distance from the node indicated by the block arrow. Unreachable nodes are at an infinite distance (i.e., $1/d_{ij} = 1/\infty$ is defined as 0). The figures show that changes in reachability due to link removal only affect the geodesic distances of the node whose path is removed ($B_{perip-} = 0$ in Fig. 2(b)). For the star network, the contribution to $\Delta C_{ij}$ for the central node is:

$$\begin{aligned}\Delta C_{central} &= (C - C_{ij'})|_{i=central} = \frac{1}{N}\sum_i^N \frac{1}{N-1}\sum_j^{N-1}\left(\frac{1}{d_{ij}} - \frac{1}{d_{ij'}}\right)\Bigg|_{i=central} \\ &= \frac{1}{N(N-1)}\end{aligned} \tag{7}$$

For the derivation of this formula, refer to Appendix A.

Considering the star network in Fig. 2(c), the contribution to $\Delta C_{ij}$ for peripheral nodes is:

$$\begin{aligned}\Delta C_{perip+} &= (C - C_{ij'})|_{i=perip+} = \frac{1}{N}\sum_i^N \frac{1}{N-1}\sum_j^{N-1}\left(\frac{1}{d_{ij}} - \frac{1}{d_{ij'}}\right)\Bigg|_{i=perip+}, \\ \forall perip+ &= \frac{(1/2)(N-2)}{N(N-1)}\end{aligned} \tag{8}$$

For the derivation of this formula, refer to Appendix A.

Combining the results for Eqs. (7) and (8) results in a $\Delta C_{ij}$ of:

$$\Delta C = (\Delta C_{central} + \Delta C_{perip+}) = \frac{1 + ((1/2)(N-2))}{N(N-1)} = \frac{1}{2(N-1)} \tag{9}$$

which is the same result obtained by Eq. (6) when considering link deletions from the peripheral nodes. This implies that regardless of which of the two links in the dyad are removed, it does not change the resulting $\Delta C_{ij}$. Thus, regardless of a particular link's location in the star structure, the effect of its removal is always:

$$\Delta C_{dyad} = \Delta C_{\text{Eq. (6)}} = \Delta C_{\text{Eq. (9)}} = \frac{1 + ((1/2)(N-2))}{N(N-1)} = \frac{1}{2(N-1)} \tag{10}$$

Consequently, the $B_i$ for any node in a star network is:

$$\begin{aligned}B_{i_{star}} &= \frac{\sum_j^{N-1} \Delta C_{ij}}{D_i} = \frac{\sum_j^{N-1} \Delta C_{dyad}}{D_i} = \frac{\sum_j^{D_i} \Delta C_{dyad}}{D_i} \\ &= \frac{D_i(\Delta C_{dyad})}{D_i} = \Delta C_{dyad} = \frac{1}{2(N-1)}\end{aligned} \tag{11}$$

where $D_i$ is the degree of $B_i$.

It is apparent that for undirected networks $\Delta C_{dyad}$ is:

$$\Delta C_{dyad} = \Delta C_{\text{Eq. (6)}} + \Delta C_{\text{Eq. (9)}} = 2\left(\frac{1 + ((1/2)(N-2))}{N(N-1)}\right) = \frac{1}{N-1} \tag{12}$$

This would introduce a factor of two into $B_i$, but then this factor of two cancels out when placed in the fraction $BN_i$, thus making for a more robust normalized measure. Additionally, because of the property of invariance of $B_i$ to anything but the parameter $N$, it is appropriate that the $B_i$ from an ideal star network can be used as a basis for comparing nodes between networks of different size.

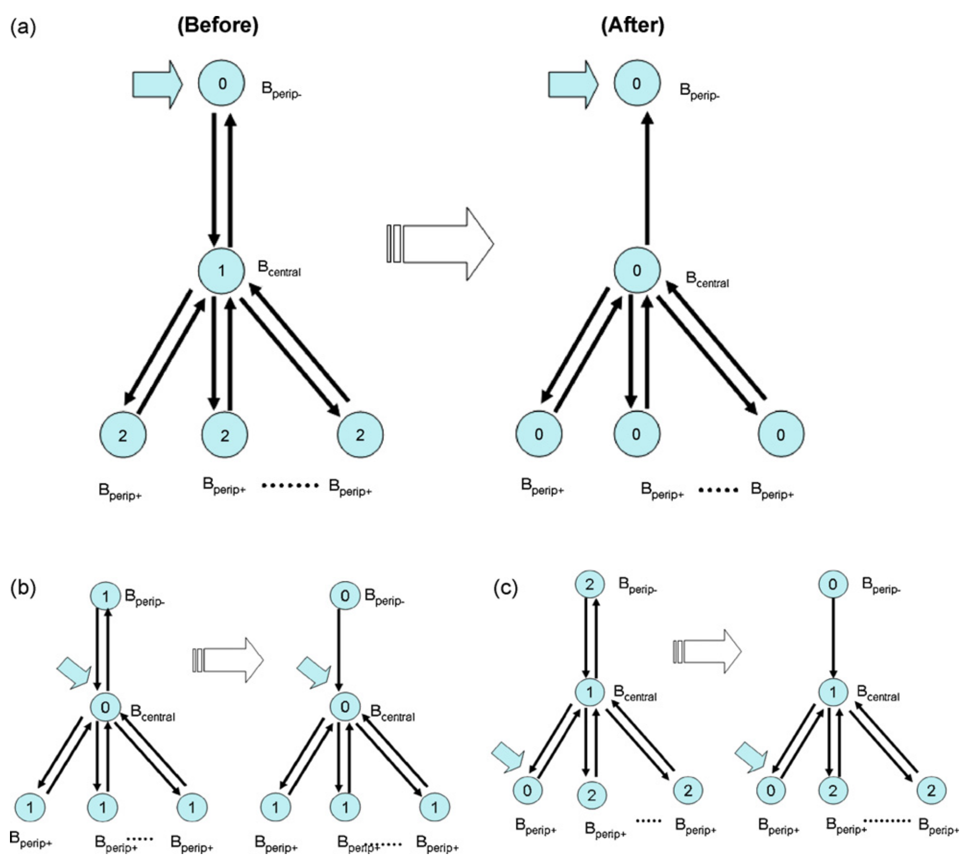


**Fig. 2.** (a) Diagram before and after tie removal from a peripheral node to the central node. The numbers inside the nodes indicate the geodesic distance from the peripheral node denoted by the block arrow. The white double-arrow indicates the transition of the network from the tie being present (before) to the tie being absent (after). The sum of the distances from this one peripheral node to all other nodes is: $\Delta C_{ij} = C - C_{ij} = [0+1+2(N-2)] - [0+0+0(N-2)] = 2N-3$. (b and c) Diagram after removal of tie from the central node to a peripheral one. The numbers inside the nodes indicate the geodesic distance from the central node to the peripheral ones. In (b) the distance changes for the central node are from $\Delta C_{ij} = [1+0+1(N-2)] - [0+0+1(N-2)] = 1$; in (c) the distance changes for a peripheral node indicated by the block arrow are $\Delta C_{ij} = [2+1+2(N-3)] - [0+1+2(N-3)] = 2$.

We propose that since all links in a star network are equally sensitive/important (as has been shown above), we can use this as the basis for a normalized bridge measure:

$$NB_i = \frac{B_i}{B_{i_{star}}} = \frac{B_i}{\Delta C_{dyad}} = \frac{B_i}{1/(2(N-1))} \tag{13}$$

where $B_i$ is bridging for an actor $i$ defined in formula (4), $B_{i_{star}}$ is bridging for actor $i$ in the star structure (see Eq. (11)). $NB_i$ ranges from 0 to 1 because $B_{i(star)}$ is the maximum possible value for given $N$.

Finally, to form a network-level measurement ($BR$) that has the property of ranging from 0 to 1 in a manner similar to centralization measures, the individual $B_i$, are summed relative to the star network ($BR_{star}$):

$$\begin{aligned} BR_{star} &= \sum_{i}^{N} B_i = B_1 + \sum_{i=2}^{N} B_i = \frac{N(2+(1/2(N-2))}{N(N-1)} \\ &= \frac{(2+(1/2)(N-2))}{(N-1)} = \frac{N+2}{2(N-1)} \\ BR &= \frac{\sum_i^N B_i}{BR_{star}} = \frac{\sum_i^N B_i}{(N+2)/(2(N-1))} = \sum_i^N NB_i \end{aligned} \tag{14}$$

By normalizing against the symmetric star network ($BR_{star}$), $BR$ is always in the range of 0–1, which is a useful property and controls for the size of the network.

Conceptually, bridging nodes cannot be nodes linked only to one other node. The one link these nodes have does not influence distances between other nodes. So bridging ($B_i$) is set to zero for nodes connected only to one other node (pendants).

## 1. Hypothetical examples

The hypothetical networks in Fig. 3 (drawn with Netdraw, Borgatti, 2005) illustrate the bridging measure. Most discussions of bridging usually begin by presenting an archetypical network such as the one in Fig. 3(a). In this network there are two stars connected by a link connecting the central nodes (1 and 6). In this network bridging is well measured by betweenness centrality and the bridge is obvious. Betweenness and bridging scores are the largest and equivalent for both nodes 1 and node 6 (betweenness = 72.2 and bridging = 11.1). Additionally, the network-level bridge score was 91.5. In Fig. 3(b) the two stars are connected by a bridging node, number 11. In this network, bridging for node 11 was the highest (16.8) and nearly twice as high as the next highest score for nodes 1 and 6 (9.1), with network-level bridge score of 92.3. On the other hand, betweenness is highest for nodes 1 and 6 (66.7) and next highest for node 11 (55.6). This simple example in Fig. 3(b) illustrates the distinction between bridging and betweenness by showing that node 11 occupies a critical position relative to nodes 1 and 6.

See Table 1 for the scores of three bridging measures, raw ($B_i$), normalized ($NB_i$), and network-level ($NBR$) for the hypothetical and empirical networks discussed below.

Fig. 3(c) illustrates what happens when the network is composed of groups of unequal size. Node 6 has the highest

**Table 1**
Raw ($B_i$), normalized ($NB_i$), and network-level bridging measures for the hypothetical examples.

| Network Bridging (%) ID | Fig. 3a | | Fig. 3b | | Fig. 3c | | Fig. 4 Granovetter | | Fig. 5 Kirke | | Fig. 6 HIV | | ID |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 91.5 | | 92.3 | | 83.8 | | 21.7 | | 44.4 | | 62.9 | | |
| | B | NB (%) | B | NB (%) | B | NB (%) | B | NB (%) | B | NB (%) | B | NB (%) | |
| 1 | 0.061 | 11.067 | 0.050 | 9.091 | 0.025 | 4.687 | 0.015 | 2.951 | 0.009 | 0* | 0.013 | 2.593 | O |
| 2 | 0.048 | 0* | 0.039 | 0* | 0.017 | 0* | 0.001 | 0.275 | 0.014 | 2.614 | 0.011 | 2.224 | FL1 |
| 3 | 0.048 | 0* | 0.039 | 0* | 0.017 | 0* | 0.002 | 0.289 | 0.011 | 0* | 0.008 | 1.553 | GA1 |
| 4 | 0.048 | 0* | 0.039 | 0* | 0.017 | 0* | 0.001 | 0.278 | 0.011 | 0* | 0.012 | 2.341 | LA2 |
| 5 | 0.048 | 0* | 0.039 | 0* | 0.017 | 0* | 0.001 | 0.272 | 0.020 | 3.868 | 0.030 | 5.785 | LA3 |
| 6 | 0.061 | 11.067 | 0.050 | 9.091 | 0.048 | 9.045 | 0.006 | 1.144 | 0.014 | 2.614 | 0.008 | 1.488 | LA4 |
| 7 | 0.048 | 0* | 0.039 | 0* | 0.026 | 4.962 | 0.026 | 4.993 | 0.009 | 0* | 0.005 | 0.970 | LA6 |
| 8 | 0.048 | 0* | 0.039 | 0* | 0.026 | 4.962 | 0.009 | 1.688 | 0.016 | 3.047 | 0.008 | 1.575 | NJ1 |
| 9 | 0.048 | 0* | 0.039 | 0* | 0.026 | 4.962 | 0.001 | 0.276 | 0.002 | 0.331 | 0.006 | 1.177 | NY11 |
| 10 | 0.048 | 0* | 0.039 | 0* | 0.026 | 4.962 | 0.001 | 0.255 | 0.004 | 0.808 | 0.008 | 1.556 | NY13 |
| 11 | | | 0.092 | 16.804 | 0.058 | 11.025 | 0.001 | 0.259 | 0.012 | 0* | 0.020 | 3.810 | NY14 |
| 12 | | | | | 0.017 | 0* | 0.001 | 0.178 | 0.001 | 0.242 | 0.014 | 2.791 | NY17 |
| 13 | | | | | 0.017 | 0* | 0.002 | 0.390 | 0.009 | 1.807 | 0.009 | 1.804 | NY18 |
| 14 | | | | | 0.017 | 0* | 0.001 | 0.218 | 0.001 | 0.278 | 0.002 | 0.459 | NY19 |
| 15 | | | | | 0.017 | 0* | 0.007 | 1.270 | 0.001 | 0.204 | 0.001 | 0.148 | NY2 |
| 16 | | | | | 0.017 | 0* | 0.002 | 0.366 | 0.001 | 0.204 | 0.010 | 1.880 | NY24 |
| 17 | | | | | 0.017 | 0* | 0.006 | 1.192 | 0.051 | 9.777 | 0.007 | 1.365 | NY5 |
| 18 | | | | | 0.017 | 0* | 0.009 | 1.655 | 0.010 | 1.947 | 0.010 | 1.888 | NY7 |
| 19 | | | | | 0.017 | 0* | 0.002 | 0.360 | 0.001 | 0.282 | 0.012 | 2.277 | NY9 |
| 20 | | | | | | | 0.003 | 0.620 | 0.001 | 0.221 | 0.002 | 0.429 | SF1 |
| 21 | | | | | | | 0.002 | 0.360 | 0.001 | 0.242 | | | |
| 22 | | | | | | | 0.001 | 0.249 | 0.001 | 0.259 | | | |
| 23 | | | | | | | 0.001 | 0.216 | 0.001 | 0.262 | | | |
| 24 | | | | | | | 0.007 | 1.388 | 0.006 | 1.135 | | | |
| 25 | | | | | | | 0.003 | 0.561 | 0.013 | 2.514 | | | |
| 26 | | | | | | | | | 0.009 | 0* | | | |

* Pendants set to zero (see text).

betweenness score (82.4), followed by node 11 (42.5) and node 1 (40.5). On the other hand, node 11 has the highest bridging score (11.02), followed by node 6 (9.0) and node 1 (4.7). This illustrates that in an unbalanced network the critical bridging function of node 11 is more accurately measured by bridging than betweenness. The network-level bridge score was 83.8.

As a second example, consider the hypothetical network described by Granovetter (1973) in the strength of weak ties article (Fig. 4). Node 7 has the highest bridging score (5.0), followed by nodes 1 (3.0), 8 (1.7), 18 (1.7), and 24 (1.4) with a network-level bridge score of 21.7. In contrast, node 1 has highest betweenness score (36.4), followed by nodes 24 (35.8), 6 (33.9), and 8 (33.4). Granovetter (1973) focused on the link between nodes 17 and 18 as the one which reduced overall path lengths the most. Several other links are also critical including those between nodes 1 and 23; 13 and 15; 6 and 7; and 7 and 8. Node 7 is unique, however, because it has only two links, both of which are critical to the network.

To gain a better understanding of how bridging correlates with other network measures, we created 100 randomly generated networks of 100 nodes and mean degree 10. Table 2 reports average Spearman's rank correlation coefficients between selected network measures as well as the univariate statistics for the measures. For measures we included degree, personal network density (the proportion of links among 1st degree alters), closeness centrality, betweenness centrality, constraint, vitality (node removal), bridging based on the maximum of each node's link changes (bridging maximum), and bridging based on the average of each node's link changes (bridging average).

Degree, closeness, and betweenness are all strongly correlated with one another as expected since they are centrality measures. Specifically, the rank correlations are degree and closeness, 0.883

**Table 2**
Average Spearman's rank correlation coefficients between selected network measures from 100 randomly generated symmetric networks of size 100 and degree 10.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1. Degree | 1 | | | | | | | |
| 2. Personal network density | 0.017 | 1 | | | | | | |
| 3. Closeness centrality | 0.883 | −0.077 | 1 | | | | | |
| 4. Betweenness centrality | 0.929 | −0.160 | 0.851 | 1 | | | | |
| 5. Constraint | −0.485 | 0.789 | −0.518 | −0.612 | 1 | | | |
| 6. Vitality (node removal) | −0.970 | −0.080 | −0.786 | −0.851 | 0.416 | 1 | | |
| 7. Bridging (maximum) | 0.120 | −0.163 | 0.211 | 0.251 | −0.192 | −0.01 | 1 | |
| 8. Bridging (average) | −0.115 | −0.345 | 0.051 | 0.118 | −0.205 | 0.279 | 0.613 | 1 |

| | Mean | S.D. | Range |
|---|---|---|---|
| Degree | 10 | 2.036 | 3.75–15.21 |
| Personal network density | 0.092 | 0.047 | 0.000–0.273 |
| Closeness centrality | 0.450 | 2.173 | 0.377–0.501 |
| Betweenness centrality | 0.013 | 0.515 | 0.001–0.029 |
| Constraint | 0.195 | 0.053 | 0.107–0.453 |
| Vitality (node removal) | 0.216 | 0.043 | 0.157–0.450 |
| Bridging (maximum) | 0.049 | 0.008 | 0.034–0.079 |
| Bridging (average) | 0.035 | 0.004 | 0.026–0.051 |

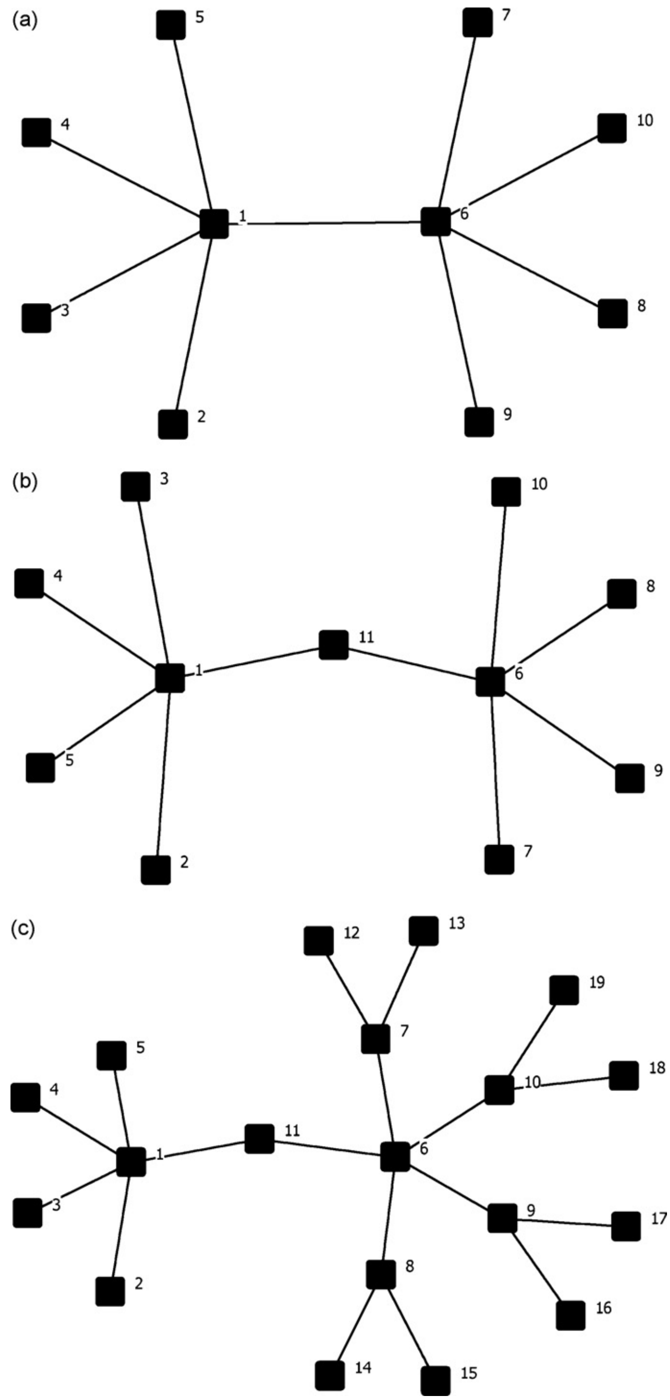


**Fig. 3.** (a) Illustrates a bridge link, the connection between nodes 1 and 6. The network in (b) illustrates a bridging node, number 11 that connects two stars. Nodes 1 and 6 have higher betweenness than node 11 which has the highest bridging score, but also the highest closeness score. In (c) the network is unbalanced and node 11 has the highest bridging score with node 6 having the highest betweenness and closeness scores (figures made with Netdraw, Borgatti, 2005).

($p < 0.001$); degree and betweenness, 0.929 ($p < 0.001$); and closeness and betweeenness, 0.851 ($p < 0.001$). Constraint is negatively correlated with the centrality measures: constraint and degree, −0.485 ($p < 0.001$); constraint and closeness, −0.518 ($p < 0.001$); constraint and betweenness, −0.612 ($p < 0.001$). Bridging (average) is weakly negatively correlated with personal network density, −0.345 ($p < 0.001$) and constraint, −0.205 ($p < 0.05$); and not correlated with degree, closeness or betweenness. Additionally, the rank

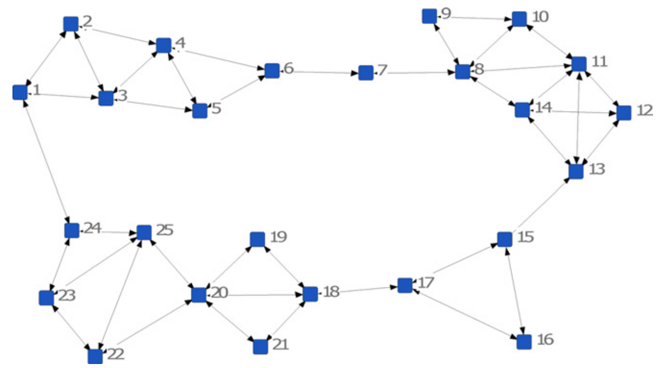

**Fig. 4.** Network from Granovetter's (1973) strength of weak tie article (Fig. 2(b) on p. 1365). Node 7 has the highest bridging score (figure made with Netdraw, Borgatti, 2005).

correlation between bridging (average) and bridging (maximum) is not so high, 0.613 ($p < 0.001$), indicating that researchers may use either the average or maximum score for different applications. It is worth noting that the rank correlation between bridging as the average of the link changes and vitality is low, 0.279 ($p < 0.01$), and there is no correlation between bridging as the maximum of the link changes and vitality. This indicates that removing a node is distinctly different from the average or maximum of its link deletions. These correlations demonstrate that bridging provides a measure of critical nodes in a network that are not central nodes, not the inverse of centrality, and not correlated with degree.

## 2. Empirical examples

Kirke (2004, 2005) collected friendship data from Dublin adolescents to investigate social network influences on substance use behavior. Fig. 5 shows one network from the study (from Fig. 1 in Kirke, 2004 and Fig. 8.1 Kirke, 2005). The graph clearly indicates that node 17 functions as a bridge between two subgroups. Our analysis also showed that node 17 had the highest bridging score (9.8), followed by node 5 (3.9) and node 8 (3.0). The network-level bridge score was 44.4. As for the betweenness scores, node 13 has the highest betweenness scores, 58.8, followed by 17 (48.0) and 18 (46.8). In addition to the highest betweenness score, Node 13 had the highest closeness score yet the eighth highest bridging score (1.81). Nodes 13 and 18 are critical access points to the two subgroups, yet both are members of each of those subgroups whereas node 17 is independent of them. Node 17 is a bridging node, a person whose links connects him/her to separate subgroups.

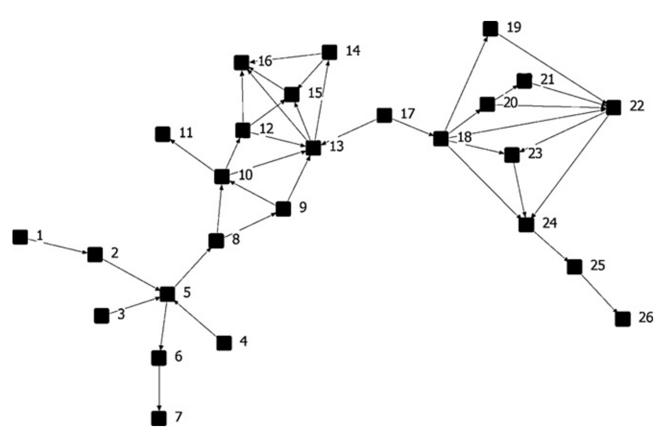

**Fig. 5.** Friendship links among adolescents in one Dublin community. Node 17 has the highest bridging score (9.8) and links two otherwise disconnected subgroups. Data are from Kirke (2004, 2005).

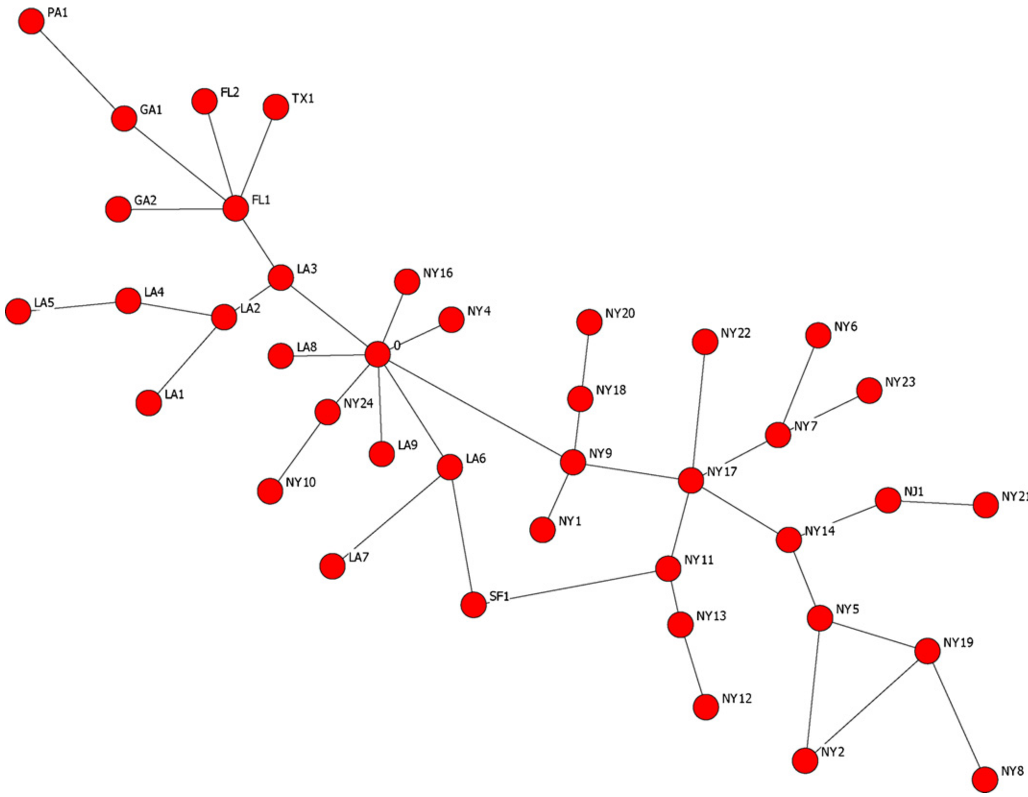


**Fig. 6.** First 40 diagnosed HIV cases. LA3 has the highest bridging score (5.8). Data are from Auberbach, and others 1984 and Klovdahl (1985).

A second example comes from the well-known data collected among the first 40 HIV cases diagnosed in the US (Auerbach et al., 1984; Klovdahl, 1985). Fig. 6 shows the network with node labels indicating the geographic location of the patients (with patient zero labeled as zero[1]). Bridging analysis returned the largest bridging score for patient LA3 (5.8), followed by NY14 (3.8) and NY17 (2.8). The seven nodes comprising the spine of the network (FL1, LA3, 0, NY9, NY17 and NY14) were among the top seven highest bridging scores (LA2 has the fifth highest score), yet LA3 and NY14 were among the highest ones because they were the only patients with three sexual contacts, all others have four or more and their average change scores are higher than the others. Interestingly, LA3 was the first patient interviewed in this study (Darrow, personal communication). The network-level bridge score was 62.9.

As most analyses focus on patient zero as a key node in the transmission network, betweenness analysis also identified patient zero as having the highest betweenness score (64.2), followed by NY17 (51.3) and NY9 (49.3). The nodes along the spine of the network (FL1, LA3, 0, NY9, NY17 and NY14) were all among the top six highest betweenness scores. Patient zero may have contributed disproportionately to the epidemic by having many relationships, but this also means that disrupting transmission from this person is more difficult because there are so many relationships to sever. In contrast, LA3 has fewer relationships and so they may be more critical to transmission dynamics.

## 3. Utility for understanding network structure

In addition to the individual bridging measures, network-level indicators can be derived from the individual distributions of the bridge scores by calculating the mean and variances of bridging for any network. For example, network centralization can be measured as the variance of the degree distribution (Wasserman and Faust, 1994). High degree variance indicates a centralized network. Likewise with the bridging measures, high variance in bridging scores indicates a vulnerable or potentially fragmented network whereas low variance indicates a cohesive network.

In this paper, bridging was calculated by computing the change in cohesion for each link deletion. Although not discussed here, potential bridging can be calculated by computing the change in cohesion for each link addition. The change values are stored in a change matrix that represents a potentially useful piece of information for understanding network structure. The change matrix can also be created for other network indicators such as change in centralization, transitivity and so on. In sum, the methodology proposed here can be expanded to create directional measures, potential bridging measures, and measures based on other network indicators such as centralization, clustering, and transitivity.

## 4. Discussion

In this paper we have introduced a measure of bridging designed to measure the degree a node in a network occupies a strategic position such that changes in links to or from this node have maximal impact on the overall structure of the network, in this case by changing network cohesion. We demonstrated bridging calculations on hypothetical networks, created simulated networks to correlate it with other well-known network measures, and applied the measure to empirical networks.

Bridging may be used specifically for network interventions (Valente, 2010). A manager or policymaker wishing to create a more cohesive or more fragmented network can target changes to those links and/or those individuals that would provide the max-

[1] Originally this patient was labelled "O" for "outside California" (Darrow, personal communication).

imum benefit for changing network structure. Researchers have primarily recommended targeting central nodes for this purpose, but this paper suggests that targeting bridging nodes may be a more optimal way to achieve such changes. For example, scholars recommend using opinion leaders to accelerate diffusion of innovation but such leaders may already be overburdened given their status as leaders (Borgatti, 2006; Valente and Davis, 1999). Consequently, finding key linkages and individuals that increase cohesion may be an alternative and more efficient way to use network information to accelerate change or otherwise improve performance. In addition, behavior changes that specifically disrupt existing norms may be more likely among bridging than central ones.

The network analysis field has developed many measures of centrality, but emphasis on centrality may have detracted from developing measures for other structural positions in a network. Centrality measures, particularly betweenness, do an excellent job of finding bridges as long as the links between disparate groups emanate from the center of the network. When bridging nodes are not central, however, existing measures of centrality are not well suited for identifying these critical connectors. The bridging measures proposed here seem to capture those structural positions.

A general method for calculating network positions by deleting a link, recalculating a network property, and taking the difference has been introduced. Here we calculated the difference in cohesion (inverse of the average path length distances), but the approach could be extended to the calculation of other measures. For example, links could be changed and calculations made for differences in centralization, transitivity, clustering, and so on. In this way, change centralization, transitivity, clustering could be created to provide additional structural measures. Aggregating these changes to the nodes provides individual structural measures not captured by other network indicators. In addition, these different change indicators can be aggregated to create a composite indicator of structural importance or used separately depending on the application.

It is also expected that bridging individuals might exhibit certain types of behaviors. For example, bridging nodes are more likely to connect otherwise disconnected groups. In an adolescent sexual contact network, for example, bridging nodes may engage in earlier romantic relationships thus connecting boys and girls. In friendship networks, bridging nodes may be less likely to have friends who know one another and this may create some cognitive dissonance and interpersonal tension. Regardless of cause or motivation, occupying a bridging position may put individuals at risk for certain deleterious behaviors. Conversely, bridging may put individuals at an advantageous position to learn about or adopt certain behaviors earlier than their non-bridging peers.

The measures and results reported here may have particular application to the prevention of disease transmission. To prevent diseases from spreading within communities, researchers advocate immunizing central nodes as that would have the greatest effect on preventing further spread. To prevent disease from spreading between communities, however, bridging nodes should be immunized. From a local perspective it makes sense to focus on central nodes, but from a global or macro perspective, bridging is critical. Thus, while targeting central nodes may seem a more rational strategy for disease prevention, it also requires the severing or management of many more relationships while bridging nodes may have few relationships that need severing to disrupt transmission.

One limitation of this approach is that it considers only one link change at a time. In many settings, multiple changes both in links and nodes will occur and there is some artifice to holding an entire network constant while changing one link. It is hoped that further developments with this approach can devise analytic solutions that measure the sensitivity of this approach to multiple network changes. Another limitation is that it is somewhat sensitive to network size and structure. Link deletions cannot be made to isolated nodes and link additions cannot be made to completely connected nodes. Finally, it may be that these bridging measures are helpful in understanding network structure in large networks, yet the methodology proposed here is computationally intensive and thus more appropriate for small or modest sized networks in which it is possible to delete every link in the network. In terms of time to compute, assuming the standard geodesic path calculation complexity is $N^3$, and there are a potential $N(N-1)$ edges in a network (i.e., when density = 1.0), then the $B_i$ estimate has a worst-case complexity of $N^5$. However, many real-world networks are sparser, and the calculation takes advantage of that by deleting only existing links.

In sum, the bridging measures provide an approach to understanding network structure and dynamics directly derived from Granovetter's (1973) insight that ties which decrease overall network distances are potentially strong. The bridging measures are intuitively appealing as they tap into important structural positions that are not at the center of the network but rather bridge different segments. We hope the bridging measures are useful to scholars interested in studying network structure, network dynamics, and network effects on substantive behavioral phenomenon.

## Acknowledgements

Support for this research was provided by NIDA grants P50-DA16094 and RC AA019239. We thank Rebecca Davis, Erik Lindsley, Martin Everett, Phil Bonacich, Chris Weare, Laura Koehly, Valdis Krebs and several anonymous reviewers for comments on earlier drafts.

## Appendix A.

Derivation of formula (6)

$$
\begin{aligned}
\Delta C &= (C - C_{ij'})|_{i=perip-} \\
&= \frac{1}{N}\sum_{i}^{N}\frac{1}{N-1}\sum_{j}^{N-1}\left(\frac{1}{d_{ij}}-\frac{1}{d_{ij'}}\right)\Bigg|_{i=perip-} \\
&= \frac{(((1/d_{ij})-(1/d_{ij'}))|_{i=perip-,j=central}) + \left(\sum_{j\in perip+}((1/d_{ij})-(1/d_{ij'}))|_{i=perip-}\right)}{N(N-1)} \\
&= \frac{(((1/d_{ij})-(1/d_{ij'}))|_{i=perip-,j=central}) + ((N-2)((1/d_{ij})-(1/d_{ij'}))|_{i=perip-,j=perip+})}{N(N-1)} \\
&= \frac{((1/1)-(1/\infty)) + ((N-2)((1/2)-(1/\infty)))}{N(N-1)} \\
&= \frac{(1)+(1/2)(N-2)}{N(N-1)} = \frac{1}{2(N-1)}
\end{aligned}
$$

In the above equation, the nodes in the expression were decomposed into two subsets, the central node ($d_{ij} = 1$) and the peripheral nodes ($d_{ij} = 2$). The $(N-2)$ arises from the number of $B_{perip+}$ nodes.

Derivation of formula (7)

$$
\begin{aligned}
\Delta C_{central} &= (C - C_{ij'})|_{i=central} = \frac{1}{N}\sum_{i}^{N}\frac{1}{N-1}\sum_{j}^{N-1}\left(\frac{1}{d_{ij}}-\frac{1}{d_{ij'}}\right)\Bigg|_{i=central} \\
&= \frac{(1)\sum_{j}^{N-1}((1/d_{ij})-(1/d_{ij'}))}{N(N-1)} \\
&= \frac{\left(\sum_{j}^{N-1}1/d_{ij}\right)-\left(\sum_{j}^{N-1}1/d_{ij'}\right)}{N(N-1)}
\end{aligned}
$$

$$= \frac{((N-1)(1/d_{ij})) - ((N-2)(1/d_{ij'}))}{N(N-1)}$$

$$= \frac{((N-1)(1/1)) - ((N-2)(1/1))}{N(N-1)}$$

$$= \frac{(N-1-N+2)}{N(N-1)} = \frac{1}{N(N-1)}$$

In this equation, the central node initially has links to $(N-1)$ peripheral nodes. After removal for the link from the central node to $B_{perip-}$, there are only $(N-2)$ links (all directed to the $B_{perip+}$ nodes).

Derivation of formula (8)

$$\Delta C_{perip+} = (C - C_{ij'})|_{i=perip+} = \frac{1}{N}\sum_{i}^{N}\frac{1}{N-1}\sum_{j}^{N-1}\left(\frac{1}{d_{ij}} - \frac{1}{d_{ij'}}\right)\Bigg|_{i=perip+},$$

$$\forall perip+ = \sum_{\forall perip+} \frac{(((1/d_{ij}) - (1/d_{ij'}))|_{i=perip+,j=central}) + (((1/d_{ij}) - (1/d_{ij'}))|_{i=perip+,j=perip-}) + \left(\sum_{j\neq i, j\in perip+}((1/d_{ij}) - (1/d_{ij'}))|_{i=perip+}\right)}{N(N-1)}$$

$$= \sum_{\forall perip+} \frac{((1/1) - (1/1))|_{i=perip+,j=central} + ((1/2) - (1/\infty))|_{i=perip+,j=perip-} + \left(\sum_{j\neq i, j\in perip+}((1/2) - (1/2))|_{i=perip+}\right)}{N(N-1)}$$

$$= \sum_{\forall perip+} \frac{(0) + (1/2) + (0)}{N(N-1)} = (N-2)\frac{(1/2)}{N(N-1)} = \frac{(1/2)(N-2)}{N(N-1)}$$

In Eq. (8), all nodes have been grouped into subgroups of equal geodesic distances as in Eq. (7). Unlike Eq. (7), however, we must account for multiple reference nodes ($B_{perip+}$) by summing over all nodes, and so the $(N-2)$ arises from the number of $B_{perip+}$ nodes.

## References


Auerbach, D.M., Darrow, W.W., Jaffe, H.W., Curran, J.W., 1984. Cluster of cases of the acquired immune deficiency syndrome: patients linked by sexual contact. American Journal of Medicine 76, 487–492.
Becker, M.H., 1970. Sociometric location and innovativeness: reformulation and extension of the diffusion model. American Sociological Review 35, 267–282.
Borgatti, S., 2006. Identifying key players in a social network. Computational and Mathematical Organization Theory 12, 21–34.
Borgatti, S., 2005. NetDraw: Graph Visualization Software. Analytic Technologies, Lexington, KY.
Borgatti, S.P., Everett, M.G., 2006. A graph-theoretic perspective on centrality. Social Networks 28, 466–484.
Borgatti, S.P., Everett, M.G., Shirey, P.R., 1990. LS sets, lambda sets and other cohesive subsets. Social Networks 12, 337–357.
Burt, R.S., 1992. Structural Holes: The Social Structure of Competition. Harvard University Press, Cambridge, MA.
Cancian, F., 1979. The innovator's situation: Upper-middle-class conservatism in agricultural communities. University Press, Palo Alto, CA: Stanford.
Darrow, William (June, 2007). Personal communication.
Doreian, P., Fujimoto, K., 2004. Identifying linking-pin organizations in inter-organizational networks. Computational and Mathematical Organization Theory 10, 45–68.
Everett, M.G., 1983. EBLOC: a graph theoretic blocking algorithm for social networks. Social Networks 5, 323–346.
Freeman, L., 1979. Centrality in social networks: 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.
Girvan, M., Newman, M.E.J., 2002. Community structure in social and biological networks. Proceedings of the National Academy of Science 99 (12), 7821–7826.
Granovetter, M., 1973. The strength of weak ties. American Journal of Sociology 78, 1360–1380.
Granovetter, M., 1982. The strength of weak ties revisited. In: Marsden, P.V., Lin, N. (Eds.), Social Structure and Network Analysis. Sage, Newbery Park, CA.
Harary, F., Norman, R., Cartwright, D., 1965. Structural Models: An Introduction to the Theory of Directed Graphs. Wiley, New York.
Holme, P., Ghoshal, G., 2008. The diplomat's dilemma: maximal power for minimal effort in social networks. 2008/05/26, DOI: 0805.3909, arXiv.
Kirke, D.M., 2005. Teenagers and Substance Use: Social Networks and Peer Influence. Palgrave, New York.
Kirke, D.M., 2004. Chain reactions in adolescents' cigarette, alcohol and drug use: similarity through peer influence or the patterning of ties in peer networks? Social Networks 26, 3–28.
Klovdahl, A., 1985. Social networks and the spread of infectious diseases: the AIDS example. Social Science and Medicine 21, 1203–1216.
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, Berlin.
Moody, J., White, D.R., 2003. Structural cohesion and embeddedness: a hierarchical concept of social groups. American Sociological Review 68, 103–127.
Motter, A.E., Nishikawa, T., Lai, Y.C., 2002. Range-based attack on links in scale-free networks: are long-range links responsible for the small-world phenomenon? Physical Review E 66, 065103.
Rogers, E.M., 2003. Diffusion of Innovations, 5th ed. The Free Press, New York.
Schwartz, J.E., Sprinzen, M., 1984. Structures of connectivity. Social Networks 6, 103–140.
Valente, T.W., 2010. Social Networks and Health: Models, Methods, and Applications. Oxford University Press, New York.
Valente, T.W., Davis, R.L., 1999. Accelerating the diffusion of innovations using opinion leaders. The Annals of the American Academy of the Political and Social Sciences 566, 55–67.
Wasserman, S., Faust, K., 1994. Social Networks Analysis: Methods and Applications. Cambridge University Press, Cambridge, UK.
Watts, D., 1999. Small Worlds: The Dynamics of Networks Between Order and Randomness. Princeton University Press, Princeton, NJ.
White, D.R., Harary, F., 2001. The cohesiveness of blocks in social networks: node connectivity and conditional density. Sociological Methodology 31, 305–359.