Social Networks 33 (2011) 88–97

Contents lists available at ScienceDirect

Social Networks

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

# Spine segments in small world networks

Noah E. Friedkin*

*Department of Sociology, University of California, Santa Barbara, CA 93107, United States*

ARTICLE INFO

*Keywords:*
Small worlds
Networks

ABSTRACT

Investigations of small world contact networks, defined as networks with a short characteristic path length and a substantial local clustering of contacts in the neighborhood of each node, have emphasized the process performance of such networks. The argument that large-scale, small world, contact networks are structures with startlingly efficient process performance is premised on the existence of shortcuts, without which the characteristic path lengths of the networks would be substantially larger. No doubt, given a high probability of transmission in each contact of a network, such shortcuts are a potential structural basis of reliable flows of information, influence, material and disease. However, interpersonal contacts are often markedly unreliable transmission conduits, and the average shortcut contact may be a more unreliable, episodic, transmission conduit than the average contact of cliques. With markedly unreliable contacts, fundamental helix substructures, that are parallel-transmission subsystems of the contact network, importantly enter into the analysis of network performance. These substructures of disjoint path redundancies are based on the local clustering of contacts in the neighborhoods of each node. Drawing on network reliability theory, this article presents an approach in which intersecting cliques of contact networks are a theoretically important construct in the specification of the transmission implications of observed contact networks. Clique intersections are a structural basis of path redundancies that enable reliable transmission among the nodes of contact networks consisting of contacts that may or may not be active conduits of transmissions during some period of time. The strong contacts that occur among clique members further enhance the contributions of these path redundancies.



## 1. Introduction

Contact networks transmit information, influence, material, and disease. The performance of contact networks, i.e., their structural enabling of more or less reliable and rapid transmissions, is an enduring subject of inquiry.[1] Information on local events, witless and insightful opinions, new practices, valuable commodities, and debilitating illnesses are exchanged and spread via interpersonal contacts. With such transmissions, individuals' information, attitudes, behaviors, resources, and health are not independent of the other individuals' information, attitudes, behaviors, resources, and health. A fundamental contribution of social network theories, concerned with processes that unfold in contact networks, are the explanations that they provide of origins of the states of individuals on these variables. Network-based interdependence arises in small groups with 2–10 members, in meso-groups with 11–$10^2$ members, and groups that are orders of magnitude larger. The investigation of network-based interdependencies in small- and meso-groups has been ongoing. Investigations of large-scale groups have accelerated with work on small-world contact networks (Newman, 2000; Newman et al., 2006; Strogatz, 2001; Watts, 1999; Watts and Strogatz, 1998). This article is focused on the latter line of work and the thesis that particular structural features of large-scale networks enable reliable transmissions between the nodes of low density contact networks.

Watts and Strogatz (1998) illustrate the structural properties of a small-world contact network with Fig. 1. Each node of the graph (Fig. 1A) is situated in a maximal complete clique of size 3, and these cliques intersect in a regular manner: each clique intersects with two other cliques, based on two shared nodes, in single cycle of intersections. Fig. 1B is based on an algorithm, a constrained random rewiring of the edges in Fig. 1A, which produces shortcuts in the graph. While the algorithm is ad hoc, it serves to realize a graph with high clustering of contacts in the neighborhood of each node, and the same number of edges as the graph of Fig. 1A. The resulting Fig. 1B graph has a short characteristic path length and large clustering coefficient. The random shortcuts dramatically reduce the characteristic path length of the graph without dramatically reducing its clustering coefficient.

* Tel.: +1 805 259 8389.
*E-mail address:* friedkin@soc.ucsb.edu

[1] The term "performance" will have specific well-defined definition in the formalization employed in this article—reliability theory, where it refers to success or failure of an edge as a transmission conduit during some period of time, and the reliability of a network as a basis of transmissions in the particular ordered pairs of nodes of the network.

0378-8733/$ – see front matter 
doi:10.1016/j.socnet.2010.11.004

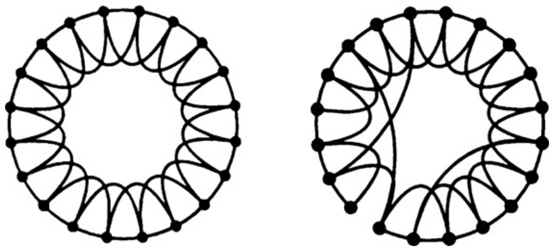



**Fig. 1.** Watts–Strogatz small-world ring. (A) Each node is situated in a maximal complete clique of size 3 that intersects with two other cliques, based on two shared nodes, in single cycle of intersections. (B) The result of a constrained random rewiring of the edges of (A). The number of nodes of the ring, here 20, and number of shortcuts, here 3, may both be increased to generate a large-scale small-world network.

The rewiring algorithm employed in the Watts and Strogatz (1998) small-world generation model may be, and has been, modified (Strogatz, 2001). In its revised form the construction of a small-world model starts with a ring of $n$ nodes in which each node is adjacent to its nearest and next-nearest neighbors out to some range $k$. Shortcut edges are then added – rather than rewired – between randomly selected pairs of nodes. Increasing the range $k$ uniformly increases the adjacencies of nodes to $2k$, and generates larger cliques of size $k+1$ that sequentially intersect in a ring of intersections, where each intersection is based on $k$ shared members.

Based on the existence of shortcuts, the key potential implication of the short paths of small-world networks is their provision of efficient communication channels between different parts of a system, which allow the dynamical processes unfolding in the network to quickly generate ramifying information flows, behavioral cascades, and global coordination of behavior. The hypothesis that shortcuts may facilitate such dynamical processes is, of course, plausible. Whether shortcuts are reliably able to do so is more problematic. The importance of shortcuts and other paths of a contact network depend not only on the properties of the process that unfolds in a network, but also on the probability of edge failures and mitigating path redundancies. Edge failure refers to the binary state of a particular edge as either an active or inactive basis of transmission during a period of time, and path-redundancy refers to the existence of some multiplicity of alternative paths by which transmission may occur.

Drawing on threshold models of behavioral diffusion, Centola and Macy (2007) illustrate how properties of the process that is assumed by such models condition the transmission implications of structural features of a contact network. Moody (2002) illustrates how information on the temporal sequence of disease transmissions in interpersonal contacts conditions the implications of a contact network for contagions. The present article is consistent with such work, in which a fixed contact network is assumed and the implications of the network for transmissions among the network's nodes is taken as ambiguous in the absence of additional specification. Here, I employ a specification of a contact network in terms of its edge-failure probabilities.

For large-scale networks of interpersonal contacts, a credible premise is that the average edge of an interpersonal contact is a more or less reliable transmission conduit; each edge of the network may or may not be an active conduit that enables a transmission during some period of time. A further credible premise is that the average shortcut edge of a contact network is a more unreliable, episodic, transmission conduit than the average edge of the cliques of a contact network (Granovetter, 1973, 1983). Thus, shortcuts provide opportunities that may or may not be realized. The high clustering coefficients of small-world networks also has an ambiguous status. On the one hand, the presence of such clustering is an acknowledgement that the edges of many empirical contact networks are clustered and, therefore, should be a structural feature of small-world models. On the other hand, such clustering has an important status as the structural basis of path redundancies that may have a substantial effect on the probability of transmissions and, perhaps, in turn, on the emergence of global coordination. In the ring network (Fig. 1), path redundancy increases as the range $k$ of adjacencies is increased.

The present article considers the implications of edge-failure probabilities and path redundancies in large-scale contact networks. The article draws on network reliability theory to present a theoretical analysis of the implications of edge-failure probabilities and path redundancies in large-scale "small world" networks (Colbourn, 1987; Hillier and Lieberman, 1980; Ross, 2007). The article highlights the transmission implications of sequences of intersecting cliques, and it assesses the implications of shortcuts that are *added* to a contact network in which all nodes are linked by one or more sequences of intersecting cliques. Based on a strong form of such sequences, the article presents an analysis that generally applies to rings, trees and other networks in the domain of networks with this strong form of clique sequencing. An approach is employed that deals with conservative lower bounds for network transmission reliability values based on edge disjoint paths in idealized "small world" networks with a regular form of clustered edges.

While the main body of this article is focused on a theoretical analysis that bears on idealized "small world" networks, the network reliability theory that is employed has broader applications to the investigations of observed contact networks, whether they are small or large. These applications draw on the familiar constructs of a network's reachability matrix and strong components. In empirical investigations, it is not unusual to find that a network is a single strong component, or that it contains a giant strong component that includes a large proportion of the nodes of the network. Such components, in which each node is linked to each other node by one or more paths, are *opportunity structures* for transmissions. In addition to detailed analyses of their internal structural features or analyses of particular social processes that may unfold in them, network reliability theory provides a potentially useful framework for refining the construct of an opportunity structure.

The approach begins with a valued network, with edge values that are taken as corresponding to the probability of the state of each edge, during some period time, as an active or inactive transmission conduit. Thus, a network with $|E|$ edges may be viewed as having $2^{|E|}$ realizations of active edges, including a realization with no active edges, and in each of these realizations a particular node $i$ is either linked to a particular node $j$ by one or more paths of active edges that allow a transmission from $i$ to $j$, or not. The network's transmission reliability values, with respect to each of its ordered pairs of nodes, are the probabilities of the existence of one or more paths of active edges connecting node $i$ to node $j$, on which basis a transmission from $i$ to $j$ may occur. Conditional on a measurement model for the edges' probability values, networks with otherwise identical structural features may present startlingly different implications for transmissions. These implications rest on either analytically derived transmission reliability values, analytically derived bounds for these reliabilities, or numerically derived estimates of the exact reliabilities. Hence, lines of work on structural models of cohesion (e.g., Moody and White, 2003) and interpersonal influence networks (e.g., Friedkin, 1998) may be advanced with this approach. Appendix A of the article describes and illustrates some the available techniques for obtaining network transmission reliabilities. The numerical technique, also outlined and illustrated in Appendix A, is readily applicable to observed complexly configured contact networks and may serve as a useful adjunct to standard structural analyses.

## 2. Intersecting cliques and their spines

We need some definitions to proceed with the development pursued in the present article:

- *Simple connected graphs and their paths*. Let $G(V, E)$ be a nontrivial simple connected graph, where $V = \{v_1, v_2, \ldots, v_n\}$ is the vertex set of $G$ and $E$ is the edge set of $G$. With $n = |V| \geq 2$ vertices (nodes), the graph is nontrivial, and with $|E| \geq 1$ edges (lines) it is not empty. For an $v_i$, $v_j$ edge in the edge set of $G$, the nodes $v_i$ and $v_j$ are its *endpoints*. A *path* on $G$ is a sequence of nodes $(v_i, v_k, v_u, \ldots, v_w, v_j)$ in which no node occurs more than once with an edge sequence of intersecting endpoints as follows $(v_i, v_k)(v_k, v_u), \ldots, (v_w, v_j)$. A path *joins* the nodes that it contains. The graph is *simple* when there are no loops on the nodes ($v_i$, $v_i$ edges are absent for all $i$), when the edges are undirected and unweighted, and when at most one edge exists for each of the $n(n-1)/2$ pairs of nodes (multiple edges joining $i \neq j$ do not exist for all $i$ and $j$). The graph is *connected* when at least one path joins every $i \neq j$ pair of nodes. The distance $d_{ij}$ separating two $i \neq j$ nodes of $G$ is the number of edges of a shortest path that joins them. The *characteristic path length* of $G$ is the average distance separating the $n(n-1)/2$ pairs of nodes of $G$.
- *Subgraphs, cliques, and densities*. An edge-induced subgraph is a subset of the edges of $G$ together with any vertices that are their endpoints. A vertex-induced subgraph of $G$ is a subset of the vertices of $G$ together with any edges whose endpoints are both in this subset. A *clique* is a complete subgraph of $G$, i.e., each pair of nodes in the subgraph is joined by an edge. The clique is *maximal* when its subgraph cannot be enlarged with the addition of other nodes. A *k-clique* is a subgraph of $G$ in which the maximum distance separating any pair of nodes in the subgraph less than or equal to $k$. The *density* of a subgraph of $G$, is the fraction of the number of possible edges among the nodes of the subgraph that exist among them.
- *Adjacency, structural equivalence, node degrees, neighborhoods, and clustering coefficient*. Two nodes $i \neq j$ are *adjacent* in $G$ if they are joined by an edge. These adjacencies may be represented as an $n \times n$ symmetric adjacency matrix $G = [g_{ij}]$, where $g_{ii} = 0$ for all $i$, $g_{ij} = 1$ for all $i \neq j$ if $i$ and $j$ are joined by an edge, and $g_{ij} = 0$ for all $i \neq j$ if $i$ and $j$ are not joined by an edge. Two nodes $i \neq j$ are *structurally equivalent* in $G$ if $\sum_{m=1}^{n}(g_{im} - g_{jm})^2 = 0$. The *degree* of a node $i$ is the number of nodes adjacent to $i$, $d_i = \sum_{m=1}^{n} g_{im}$. The *neighborhood* of node $i$ is this subset of adjacent nodes. The *clustering coefficient* of $G$ is the average density of the neighborhoods of $G$. Each node of $G$ has $d_i$ adjacent nodes, among whom $d_i(d_i - 1)/2$ edges may exist. Let $f_i$ be the fraction of these possible edges that exist in $G$. The clustering coefficient for $G$ is $(1/n)\sum_{i=1}^{n} f_i$.
- *Shortcut edges*. A *shortcut* is an $i, j$ edge whose *removal* from the edge set of $G$ produces a $G^-$ graph in which $d_{ij} \geq 3$, or in which $d_{ij} = \infty$ if there is no path in $G^-$ that joins $i$ and $j$. Alternatively, given $d_{ij} \geq 3$ in $G$, the *addition* of an $i, j$ edge to the edge set of $G$ produces a $G^+$ in which the $i, j$ edge is a shortcut.

For the theoretical analysis that is presented in this article, I will assume a contact network $G(V, E)$, with edges that are independently open to transmissions with probability $\pi$ and closed to transmissions with probability $1 - \pi$. For this network $G$, i.e., a stationary system of Bernoulli random variables, the edges may be represented as $\{X_1, X_2, \ldots, X_{|E|}\}$, for which there are $2^{|E|}$ realizations $(X_1 = x_1, X_2 = x_2, \ldots, X_{|E|} = x_{|E|})$, where

$$x_k = \begin{cases} 1 \text{ if the } X_k \text{ edge is open,} & E(X_k = 1) = \pi \text{ for all } k \\ 0 \text{ if the } X_k \text{ edge is not open,} & E(X_k = 0) = 1 - \pi \text{ for all } k \end{cases} \tag{1}$$

Each realization presents an edge-induced subgraph in which node $i$ is either joined to node $j$ by a path, or not. For each pair of nodes, $i$ and $j$ of $G$, the following *structure function* for the system is defined:

$$\phi(\mathbf{x}) = \phi(X_1 = x_1, X_2 = x_2, \ldots, X_{|E|} = x_{|E|}) = \begin{cases} 1 \text{ if } i \text{ is joined to } j \\ 0 \text{ if } i \text{ is not joined to } j \end{cases} \tag{2}$$

Simply put, each realization has a reachability matrix in which node $i$ either reaches node $j$ or not. The system's reliability for a particular pair of nodes of $G$, i.e., the probability of a transmission from $i$ to $j$, for each $i$ and $j$ of $G$, is the expectation that $i$ reaches $j$, i.e.,

$$\rho_{ij} = \sum_{\mathbf{x}:\phi(\mathbf{x})=1} \prod_{i=1}^{|E|} \pi^{x_i}(1-\pi)^{1-x_i} \tag{3}$$

Within the constraints of the NP-complete problem of determining the $G$ system's reliability values, it is feasible to analytically derive exact reliabilities, or upper and lower bounds for these values. In Appendix A, I provide a textbook illustration of how this is accomplished. A close look at what is involved is instructive. The analysis is non-trivial. Given a particular contact network of sufficient importance, an analysis may be justified that takes into account the full edge set of $G$. However, it may or may not be feasible to do so. In the present analytical context, the signature of a NP-complete problem appears when there is no known algorithm that is able to solve the analytical problem on a particular network in a practical amount of time. However, numerical techniques are available that, while approximate in nature, are feasible to implement for a broad domain of complexly configured networks.

An intensive focus on particular networks does not provide a platform for the development of fundamental theory on social network structure in the following *specific* sense: the more information on a particular social network that is taken into account, the more constrained are conclusions about the transmission implications of generic structural features of networks that exist in each member of a broad class of networks. The present investigation focuses on an elementary idealized substructure of social networks – spine segments – that occur in the domain of a class of $G$ in which all pairs of nodes at distance $d_{ij} \geq 2$ are joined by sequentially intersecting cliques. The networks in this domain have theoretical foundations in the semipath structures of generalized balance theory (Johnsen, 1985) and role structure theory (White et al., 1976) in which the global structure of a contact network is conditioned by basic premises concerning balanced and unbalanced elementary triadic structures and structural equivalence, respectively. Large empirical literatures in social psychology, anthropology, and sociology have developed on these theories. More broadly, networks consisting of sequences of intersecting cliques and high-density neighborhoods have been emphasized in work concerned with the coordination of large-scale highly differentiated social structures (Friedkin, 1998), where such sequential intersections are denoted as the spines of ridge structures. Shortcuts appear as structural anomalies in this line of work in which the global organization of structural cohesion is emphasized.

A strong form of sequentially intersecting cliques $R(C, B)$ exists for a network $G(V, E)$ that may be partitioned into $n_C = |C|$ *positions*, $C = \{c_1, c_2, \ldots, c_{n_C}\}$, each of which is occupied by some number of elementary nodes of $G$, $s_k \geq 2$, $k = 1, 2, \ldots, n_C$, that is a clique of structurally equivalent nodes in $G$, where each $c_u$, $c_v$ edge of the edge set of $R$ is *bond* composed of $s_u s_v$ edges of $G$. Assuming that the vertex-induced subgraph for the elementary nodes of each position of $R$ is not empty of elementary edges, the position's subgraph must be complete, i.e., a clique, to satisfy the structural equivalence condition of the elementary nodes located in the position. Similarly, since each bond of $R$ is not empty of elementary edges, each bond also must be complete to satisfy the structural equivalence condition of the elementary nodes of each position.

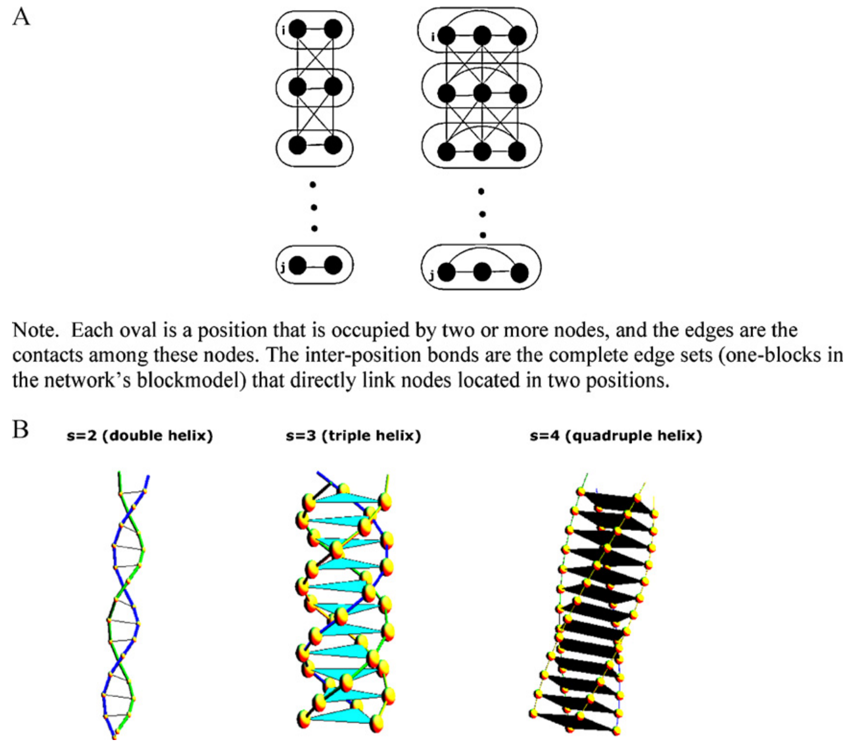



**Fig. 2.** The substructure of spine-segments. The substructure exist for all nodes $i$ and $j$ at distance $d_{ij} \geq 2$ in the $G$ of $R$ and contains a subset of size $s$ of the shortest $i-j$ paths in $G$ joining nodes $i$ and $j$. (A) Segments with two ($s = 2$) and three ($s = 3$) elementary nodes in each position of the segment that join nodes $i$ and $j$. (B) Isomorphic representation of segments as helical structures with a backbone of composed of undisplayed edges of (A) that join the nodes of adjacent positions. *Note*. Each oval is a position that is occupied by two or more nodes, and the edges are the contacts among these nodes. The inter-position bonds are the complete edge sets (one-blocks in the network's blockmodel) that directly link nodes located in two positions.

Any simple connected graph may be taken as the image of a $R(C, B)$ structure in which the vertices are positions occupied by a number of elementary nodes, and in which the edges are bonds between the positions composed of elementary edges. For each such $R$, given the specification of the number of elementary nodes that are located in each position, $S = \{s_1, s_2, \ldots, s_{n_C}\}$, there is a corresponding network $G$ with the structural equivalence and completeness conditions entailed in the definition of $R$. For the class of $R(C, B)$ with $|C| \geq 3$ positions and $|B| \geq 2$ bonds, the properties of the corresponding $R$ and $G$ graphs include the following:

(a) $G(V, E)$ is a simple connected graph with $|V| = \sum_{k=1}^{|C|} s_k$ nodes and $|E| = \sum_{k=1}^{|C|} \binom{s_k}{2} + \sum_{(c_u, c_v) \in B}^{|B|} s_u s_v$ edges. Each position contains $\binom{s_k}{2}$ elementary edges. The number of edges joining the members of two positions $c_u \neq c_v$ is either $s_u s_v$ or 0, in correspondence to the bonds of $R$. The elementary nodes involved in each $c_u$, $c_v$ bond of $R$ are a *maximal* clique of $G$ containing $s_u + s_v$ elementary nodes and $\binom{s_u + s_v}{2}$ elementary edges.

(b) Paths of $R$ with three or more positions, $c_u, c_v, c_w, \ldots, c_k$, involve a sequence of intersecting maximal cliques: the maximal clique of elementary nodes in $c_u \cup c_v$ intersects with the maximal clique of elementary nodes in $c_v \cup c_w$, and so on. In general, the elementary nodes involved in every path of positions of length $k \geq 2$ are a $k$-clique. All elementary nodes, $i$ and $j$, at distance $d_{ij} \geq 2$ in $G$ are joined by at least one path of positions, $c_u, c_v, c_w, \ldots, c_k$, with $s_u, s_v, s_w, \ldots, s_k$ elementary nodes in the respective positions and with $i \in c_u$ (source position) and $j \in c_k$ (terminal position). The vertex-induced subgraph of $G$ for the union of the elementary nodes on this path is denoted as an $ij$-spine segment of $G$ (henceforth, simply spine segment). This spine segment contains a subset of the shortest paths joining nodes $i$ and $j$ in $G$. It is a generic substructure that exists for all nodes $d_{ij} \geq 2$ in all $G$ in the domain of $R$.

(c) Now note that the subgraphs of elementary nodes and edges of a spine segment may be configured in *various* ways depending on the number of elementary nodes, $s_u, s_v, s_w, \ldots, s_k$, that are located in the positions of the segment. These variations generate complex arrays of minimal cut sets and minimal paths in the substructure, as a function of the distribution of the elementary nodes among the positions. These cut sets and minimal paths are the analytic grist for determining transmission probabilities (Appendix A). However, within all spine segments there exists a still more fundamental substructure that is based on the minimum number of elementary nodes located in the positions of the segment that joins $i$ and $j$, $s = \min(s_u, s_v, s_w, \ldots, s_k) \geq 2$, that is a vertex-induced subgraph for the union of $s$ elementary nodes taken from each position of the segment. All constructions of this more fundamental substructure, based on $s$, are isomorphic, i.e., all realizations of the $\left\{ \binom{s_u}{s}, \binom{s_v}{s}, \binom{s_w}{s}, \ldots, \binom{s_k}{s} \right\}$ combinations of elementary nodes present isomorphic subgraphs with a clique of $s$ nodes in each position and $2s$ edges joining the nodes of each pair of adjacent positions. This substructure contains $s$ of the shortest paths of $G$ that join the elementary nodes in the $c_u$ and $c_k$ positions.

Fig. 2 illustrates the generic form of the spine-segment substructure that connect all pairs of $d_{ij} \geq 2$ elementary nodes in all $R$ for a node $i$ located in the source position of the segment and a node $j$ located in the terminal position of the segment. Fig. 2A presents the special case of a $s = 2$ segment, the minimal segment for the class

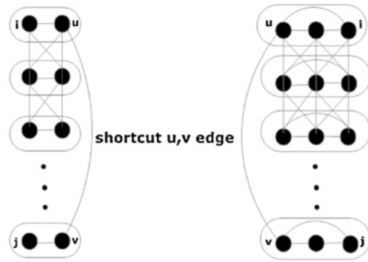



**Fig. 3.** Adding a shortcut edge to a spine-segment. For the labeled nodes in the two positions of the $s=2$ and $s=3$ segments, the shortcut $u$, $v$ edge reduces the $u$, $v$ distance from $d_{uv} \geq 2$ to $d_{uv} = 1$, and an $i$, $j$ distance $d_{ij} \geq 4$ to $d_{ij} = 3$, for all $i \neq u$ in the same of position as $u$ and all $j \neq v$ in the same position as $v$. In general, for $d_{ij} \geq 2$, the $u$, $v$ shortcut contributes a path of length 3 to the PTS that is shorter than $d_{ij}+2$. *Note*. Each oval is a position that is occupied by two or more nodes, and the edges are the contacts among these nodes. The inter-position bonds are the complete edge sets (one-blocks in the network's blockmodel) that directly link nodes located in two positions.

of $R$ with $|C| \geq 3$ positions and $|B| \geq 2$ bonds, and a $s=3$ segment. Fig. 2B illustrates the helical representation of the substructure, i.e., a double-helix in the case of a $s=2$ substructure, a triple-helix in the case of a $s=3$ substructure, and so on. The nodes at the same level of the helix are the $s$ elementary nodes of a position, and they are joined by $s(s-1)/2$ edges. The string of nodes on each of the $s$ helices join sequentially a node in one position to a node of an adjacent position. The backbone of the structure is comprised of the unrepresented edges among the nodes of adjacent levels, e.g., the crossed edges of Fig. 2A.

(d) Moving still deeper in the pursuit a closed-form generic expression of the transmission implications of spine segments, it may be noted that the substructure of all spine-segments contains, at its core, a *parallel transmission subsystem* (PTS) with $2s-1$ pairwise edge-disjoint paths of which $s$ are paths of length $d_{ij}$ (the shortest paths) and $s-1$ paths of length $d_{ij}+2$. A parallel transmission subsystem is, by definition, a subsystem based on pairwise edge-disjoint paths that, by virtue of their disjoint edge sets, make independent contributions to the probability of a transmission from node $i$ to node $j$. The reliability of this subsystem, i.e., the probability of a transmission from $i$ to $j$ on the basis of the core PTS, is

$$\rho'_{ij} = 1-(1-\pi^{d_{ij}})^s(1-\pi^{d_{ij}+2})^{s-1} \tag{4}$$

The edge set of the subsystem is obviously a subset of the edge set of $G$. Under the natural assumption of a monotone system, the full system's *i-to-j* transmission reliability is $\rho_{ij} \geq \rho'_{ij}$ and $\rho'_{ij}$ presents a conservative lower bound for the probability of a transmission from $i$ to $j$ in $G$. The conservative feature of this elementary bound allows an analysis of transmission probabilities from $i$ to $j$ for all $i$ and $j$ at distance $d_{ij} \geq 2$ in $G$, for all $G$ in the domain of $R$. Here, the global structure of $G$ is ignored in order to present a fundamental set of implications of the core PTS of the spine segments that are embedded in each $G$ in the domain of $R$. Alternative lower bounds, which take into account more or all of the edges of $G$, present values that are contingent on the global structure of $G$. Depending on the condition $\{\pi, d_{ij}, s\}$ of the PTS, $\rho'_{ij}$ may be elevated to a value that is only modestly increased by the inclusion of the other $i-j$ minimal paths of $G$.

(e) Fig. 3 illustrates a $G^+$ circumstance in which a $u$, $v$ shortcut edge has been added to the edge set of $G$. Shortcuts are structural anomalies in the $G$ of $R$. The $u$, $v$ pair of nodes that are the endpoints of the shortcut edge are not structurally equivalent with the nodes in their respective positions. With the addition of this $u$, $v$ shortcut, the maximum number of pairwise edge-disjoint paths connecting $i$ and $j$ in the spine-segment remains $2s-1$, the distribution of path lengths is altered, and transmission probability for $i$ and $j$ in the PTS of $G^+$ is

$$\rho'_{ij}(G^+) = 1-(1-\pi^3)(1-\pi^{d_{ij}})^s(1-\pi^{d_{ij}+2})^{s-2} \tag{5}$$

where $d_{ij}$ is the length of the shortest paths joining $i$ and $j$ in $G$ (n.b., not in $G^+$). The additional $u$, $v$ edge, which generates a path of length 3 joining $i$ and $j$, must increase the PTS transmission probability for $i$ and $j$. The amount of the increase is

$$\begin{aligned}\Delta_{ij} &\equiv \rho'_{ij}(G^+) - \rho'_{ij}(G)\\ &= \pi^3(1-\pi^{d_{ij}-1})(1-\pi^{d_{ij}})^s(1-\pi^{d_{ij}+2})^{s-2} > 0 \quad (0<\pi<1)\end{aligned} \tag{6}$$

The effect of the shortcut is its setting of a threshold $\pi^3$ transmission probability below which the PTS transmission probability for $i$ and $j$ cannot fall, for an $i$ and $j$ at any distance $d_{ij} \geq 2$ in $G$, which applies to all $i \neq u$ in the same of position as $u$ and all $j \neq v$ in the same position as $v$.

## 3. Results

Fig. 4A presents the transmission probabilities from a node $i$ to node $j$ in the $G$ of $R$ that are separated by some distance, $d_{ij}=2, \ldots, 25$, based on the PTS of the spine-segment that joins them. Each curve describes the PTS transmission probability in all $R$ with $|C| \geq 3$ positions and $|B| \geq 2$ bonds with a minimum of $s=2, \ldots, 100$ occupants in each position of a segment that joins the two nodes. The three plots presented in Fig. 4A are based, respectively, on edge failure probabilities of 0.50, 0.30, and 0.10. Presented in this form, the curves are generally applicable, i.e., all elementary pairs of nodes $i$ and $j$, at distance $d_{ij} \geq 2$ in the class of $G$ under consideration, have the minimum transmission probability presented on the curve for that $s$ on the plot.

The effects of path redundancy and edge failure probability are evident in these plots. Ceteris paribus, as the probability of edge failure decreases, the PTS reliability value increases. It should be evident that as the edge failure probability approaches 0, all structural features of $G$, except the simple connectivity of $i$ and $j$, become increasingly irrelevant. These features include $G$'s size, density, characteristic path length and clustering coefficient. Ceteris paribus, as path redundancy increases, so does the PTS reliability value. The marginal positive contributions of the path redundancies of parallel $s \geq 2$ subsystems decline with increasing distance. In general, for a given $\pi$ and $d_{ij} \geq 2$, there is some value of $s$ that will satisfy a criterion lower bound $\rho'_{ij} \geq \kappa$ based on $\rho'_{ij} = 1-(1-\pi^{d_{ij}})^s(1-\pi^{d_{ij}+2})^{s-1}$. Below, we have the solution

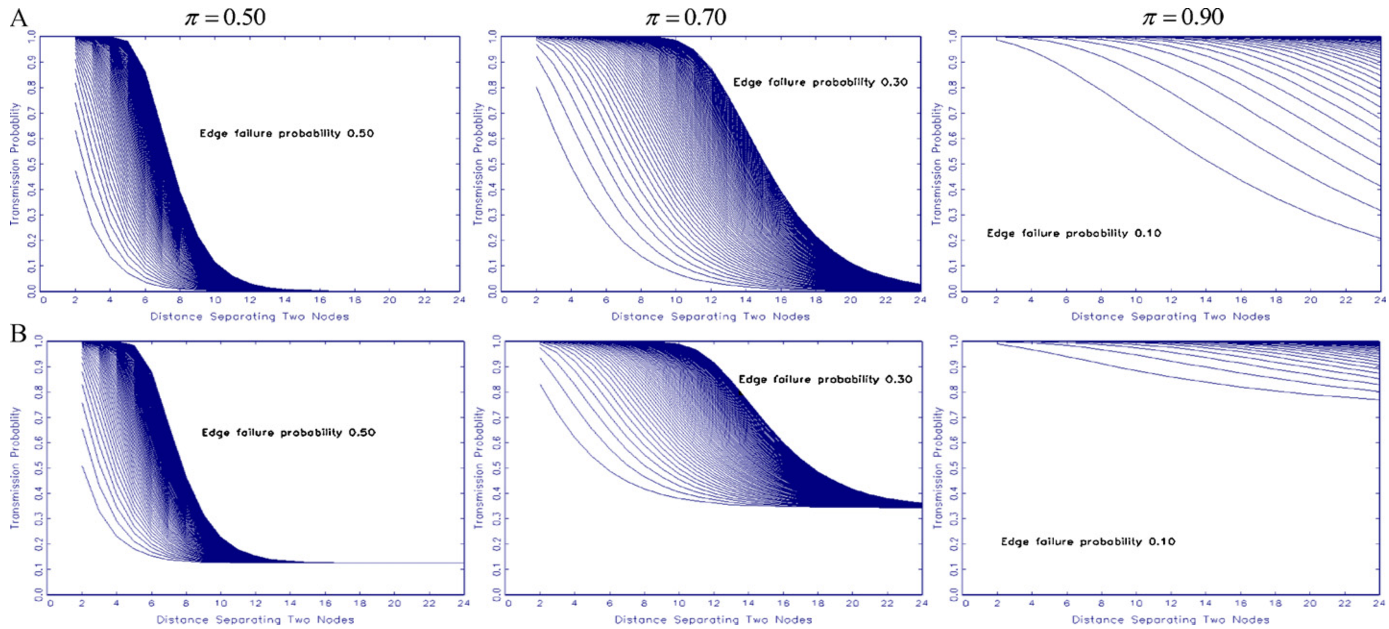



**Fig. 4.** Probabilities of a transmission for the parallel transmission subsystem (PTS) of spine-segments. This subsystem exits for all nodes $i$ and $j$ at distance $d_{ij} \geq 2$ in the $G$ of $R$. The curves on each plot are based on position sizes ($s$) 2–100. The elevation of the curves increase with $s$. (A) The transmission probability is $\rho'_{ij} = 1-(1-\pi^{d_{ij}})^s(1-\pi^{d_{ij}+2})^{s-1}$. (B) With the addition of the $u$, $v$ shortcut edge, the transmission probability is $\rho'_{ij} = 1-(1-\pi^3)(1-\pi^{d_{ij}})^s(1-\pi^{d_{ij}+2})^{s-2}$. The effect of the shortcut is a $\Delta_{ij} = \pi^3(1-\pi^{d_{ij}-1})(1-\pi^{d_{ij}})^s(1-\pi^{d_{ij}+2})^{s-2} > 0$, $0<\pi<1$, increase of the $i$, $j$ transmission probability $\rho'_{ij}$.

of this equation for those values of $s$ that satisfy $\rho'_{ij} \geq \kappa$, where $\kappa$ is some specified minimum criterion reliability value:

$$s > \frac{\ln(1-\pi^{d_{ij}+2}) + \ln(1-\kappa)}{\ln(1-\pi^{d_{ij}}) + \ln(1-\pi^{d_{ij}+2})}, \quad 0 < \pi < 1 \qquad (7)$$

To illustrate, for $\pi = 0.70$ and $d_{ij} = 16$, $s > 342$ satisfies the criterion $\rho'_{ij} \geq \kappa = 0.80$. The implication is that the joint condition of unreliable edges and long $i-j$ paths is a powerful constraint on reliable transmissions that may be mitigated by the path redundancies generated by a sequential intersection of very large cliques.

The plots of Fig. 4B present an analysis of the addition of the shortcut edge (Fig. 3) to a spine-segment. The plots for $s = 2, \ldots, 100$ present the PTS probability $\rho'_{ij}(G^+)$ for a node $i \neq u$ located in the same position as $u$ and a node $j \neq v$ located in the same position as $v$, for $d_{ij} = 2, 3, \ldots, 25$ in $G$. The main effect of the shortcut is a baseline transmission value $\rho'_{ij}(G^+) \geq \pi^3$ below which the PTS transmission probability cannot fall. For $\pi = 0.70$, this baseline value is modest $\rho'_{ij}(G^+) \geq 0.343$. Although the baseline becomes substantial for high values of $\pi$, it bears noting again that, for sufficiently high values of $\pi$, all structural features of $G$, apart from its simple connectivity, become increasingly irrelevant. Thus, it is in systems with unreliable edges that a shortcut may make a substantial contribution, but it does not suffice to secure a highly reliable transmission.

## 4. Discussion

In this article, network reliability theory is brought to bear on the analysis of "small world" contact networks. Given the rich mathematical structure and related body of work on reliability engineering, the present investigation may usefully serve to further build the interdisciplinary intersection of developments on small-world networks in the engineering, biological, physical, and social sciences. While *markedly unreliable* edges typically do not appear in the engineering applications of network reliability theory, such edges are, arguably, the reality for networks composed of interpersonal contacts. The present article draws on network reliability theory to present a theoretical analysis of the implications of unreliable edges and path redundancies in large-scale "small world" networks. The analysis occurs in the framework of idealized "small world" networks with structures that are linked to sociological work on role structures, and their representation as blockmodels composed of structurally equivalent positions and bonds between such positions.

Work on "small world" networks has emphasized that the short characteristic path length of such networks is a structural basis of efficient communication channels between distant parts of a system, which allow the dynamical processes unfolding in the network to quickly generate ramifying information flows, behavioral cascades, and global coordination of behavior. The work highlights the importance of shortcuts. The high clustering coefficients of small-world networks enters into this work as an acknowledgement that the edges of many empirical contact networks are clustered and, therefore, also should be a structural feature of small-world models. The present analysis suggests that the theoretical importance of clustering is at least as great as the theoretical importance of shortcuts in "small world" networks. Clustering does not set up the "small world" problem; clustering is part of its *solution* in providing a multiplicity of alternative transmission paths. If contacts are stochastic conduits of transmissions, with probabilities of being active (open) or inactive (closed) during some period of time, then path redundancies enter into the analysis of transmissions as an important theoretical construct and determinant of reliable transmissions among the nodes of a network. In contact networks composed of unreliable edges, the existence of parallel transmission subsystems in the network is a core structural foundation of reliable transmissions between the nodes of the network. When the local clustering of edges in the neighborhoods of nodes generates a sequence of intersecting cliques, then path redundancy is an implicated structural feature of such sequences. Path redundancy mitigates the unreliability edges and may substantially elevate the probability of transmissions between pairs of nodes. This familiar terrain of network reliability engineering has an application in research on "small world" networks. Clique intersections are one structural basis of such redundancies, and the strong contacts of cliques further enhance the contributions of these path redundancies. Thus, the pattern and composition of intersecting cliques warrant close attention in specifying the transmission implications of observed contact networks.

In the present work, I allow for the special case implicitly assumed in the literature on "small world" contact networks – the special case of a contact network in which all contacts have exceedingly high probabilities of being in an active state. I examine the broader implications of "small world" structural features relaxing the assumption that these probabilities are high. I advance the premise that the probability of an active state for the average contact of a large-scale network is substantially less than 1.0 for many types of transmissions, i.e., flows of particular types of information, interpersonal influences on particular issues, and transmissions of particular types of diseases. I note that the contributions of shortcuts to network transmission probabilities, via the short paths that they create, may be modest relative to the contributions of path redundancies enabled by intersecting clique formations. Moreover, I note that sociological research on *local bridges* in contact networks, stemming from Granovetter's (1973, 1983) seminal work on them, does not support the emphasis that "small world" investigations have placed on shortcuts. The available theoretical and empirical investigations of contacts that are local bridges support the conclusion that such contacts are weak, more unreliable, episodic, transmission conduits than the average contacts of cliques. Granovetter's "strength of weak ties" argument does not assert that weak ties are important transmission conduits; in this argument, he treats weak ties as unreliable transmission conduits that are infrequently activated. The argument is that a local-bridge contact is more likely to be a weak contact than a within-clique contact and that, *when* a local-bridge contact is active, new or useful information is more likely to be transmitted in a local-bridge contact than in an active within-clique contact. The available empirical evidence supports this nuanced argument and erodes the idea that paths involving local bridges are reliable structural bases of transmissions.

I conclude with discussion of some of the broader applications of the network reliability approach that has been employed. In Appendix A of this article, I have described and illustrated the classical analytical approach to determining network transmission reliability values. I also have outlined and illustrated a numerical approach that returns estimates of these values, which may be applied in empirical investigations of contact networks. This numerical approach is not constrained by the assumption of the idealized role structure upon which the present theoretical analysis of "small world" networks has been conducted. It works with the full edge set of an empirical contact network, and generates estimates of the network transmission reliability values for each of the network's ordered pairs. This application of network reliability theory may be a useful addition to the "tool kit" of social network analysis. Its most basic contribution is a refinement of the opportunity-structure interpretation of a strong component in which "opportunity" is replaced with "reliability" values that are probabilities of the occurrence of at least one *active* path, during some period of time, connecting node $i$ to node $j$ for each of the ordered pairs of nodes of the network. The implementation of the refinement requires a measurement model of a valued network in which the edge values are taken as corresponding to the probability of an active vs. inactive edge, for each edge. These values may be based on a suitably scaled measure of tie strength or proximity for each contact. A matrix of estimated contact network transmission reliabilities is obtained from a random sample of the realizations of active contacts, consistent with the measurement model of the probability matrix for active contacts; that is, for each of the sampled realizations, the reachability matrix for the realization is obtained, and the estimated transmission reliabilities, for each of the $(i, j)$ pairs of nodes of the contact network, are the proportions of the sampled realizations in which node $i$ reaches node $j$; see Appendix A for an illustration. This approach may be employed on a graph or digraph.

Conditional on the investigator's specification of the probability matrix for the active states of the contacts of a network, the matrix of numerical estimates of the network's transmission reliabilities may be employed in hierarchical cluster analyses, blockmodel analyses, point-centrality analyses, and analyses of selected subgroups of a network. When interest is focused on the structural implications of the subnets of subgroups, a reliability analysis of the subnets may be a useful adjunct to standard structural analyses. Most definitions of structurally cohesive subgroups are consistent with variable within-group structural features and tie strengths; for the particular subnet of a subgroup, the reliability matrix for the subgroup may be obtained. Under the assumption of monotonic positive contributions of the components of reliability values, ignoring edges that involve nodes outside a particular subgroup does not diminish the reliability values obtained strictly on the basis of the subgroup's subnet.

The application of reliability theory also extends to analyses of process models of social diffusion unfolding in a network of contacts. Consider the class of deterministic diffusion models in which a specified diffusion process generates a prediction that some equilibrium fraction of the contact network members will have adopted a particular behavior, practice, or innovation. From the perspective of network reliability theory, if a contact network is employed as the structural basis of transmission, then some specification of the uncertain active vs. inactive status of each contact of the network may be invoked. The reliability of the predictions of deterministic models of social diffusion that are operationalized with contact-network constructs may be assessed with a sample of the possible realizations of the contact network, based on a specification of the uncertain status of each contact as an active or inactive conduit of transmissions during some period of time. With respect to a particular diffusion process, a particular contact network may enable a reliable outcome that is insensitive to the failures of various contacts as transmission conduits, or components of the pressures to adopt the behavior, practice, or innovation. Here, again, network reliability theory may serve as a useful adjunct to structural analyses that rest on the assumption of a stable set of constraints (the edges of an observed contact network), but allow that these edges may be active or inactive during some period of time. It should be noted that the formalization of reliability theory presented in this article does not assume that the observed edges of a contact network are independent events. It does, however, assume that the probability that a particular given edge is active or inactive is independent of the active or inactive state of other given edges. Relaxing this assumption is feasible, but invokes a far more complex analysis. Whether the increased complexity entailed with a relaxation of this assumption is worth dealing with is a judgment that requires careful consideration.

I close with a comment on the helix representation of the spine-segments that were analyzed in this article. The graph-analytic adjacency of nodes does not necessarily imply spatial proximity. The two-dimensional representation of the spine-segments that have been analyzed in this article does not exclude their three-dimensional representation, as illustrated in Fig. 2 in terms of helices with a backbone that completes their edge set. The helix representation is suggestive. Perhaps spatially constrained biological and physical structures take this compact form when the reliability of transmissions among elementary nodes is an important factor in the coordination of complex systems.

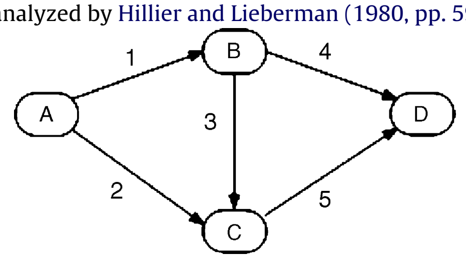
## Appendix A.

This appendix describes and illustrates, in separate sections: (a) the analytical approach with which exact transmission reliabilities are obtained, (b) the classic analytical approach with which upper and lower bounds of transmission reliabilities are obtained, (c) a numerical approach with which estimates of exact transmission reliabilities may be obtained, and (d) a conservative lower bound approach, based on edge-disjoint paths, that enables the theoretical analysis presented in the present article.

*Analytical approach to exact reliabilities*

Within the constraints of the NP-complete problem of determining a network's reliability values, it is feasible to obtain exact reliabilities for these values. Here, I provide a textbook illustration of how this is accomplished. A close look at what is involved is instructive, since analytical problem is non-trivial. The NP-complete problem appears when there is no known algorithm that is able to solve the analytical problem on a particular network in a practical amount of time. This problem arises as a function of the number of edges involved in the network.

I illustrate the approach with a simple digraph; graphs may be analyzed with the same approach. Consider the following small network analyzed by Hillier and Lieberman (1980, pp. 599–605):




In this network, a transmission from A to D will occur during some period of time if either (a) edges 1 and 4 are both open, or (b) edges 2 and 5 are both open, or (c) edges 1, 3, and 5 are all open. If edge 1 is closed to transmission, then the conditions (a) and (c) of a transmission are precluded; if edge 2 is closed, then (b) is precluded; and so on. Given probability values for the open-closed (active-inactive, success-failure) states of each of the five edges of the network, the probability of a transmission from A to B is equivalent to the probability that that at least one of the conditions {(a), (b) or (c)} is satisfied. This probability value is the network's reliability value with respect to a transmission from A to B.

To determine the exact value of the network's reliability value, with respect to a transmission from A to B, the states of each of the five edges of the network are described by the binary random variables, $X_1=x_1, X_2=x_2, X_3=x_3, X_4=x_4$, and $X_5=x_5$, for which there are $2^5=32$ possible realizations. One such realization is $X_1=1$, $X_2=0$, $X_3=1$, $X_4=0$, and $X_5=1$, in which a transmission from A to D will occur based on the open edges 1, 3, and 5, i.e., $A\xrightarrow{1}B\xrightarrow{3}C\xrightarrow{5}D$.

Each realization presents an edge-induced subgraph in which node A is connected to node D via a path, or not. For *each* of the 32 realizations, the *structure function* for the system generates a binary outcome

$$\phi(x)=\phi(X_1=x_1, X_2=x_2, \ldots, X_5=x_5) = \begin{cases} 1 & \text{if a path from A to B exists} \\ 0 & \text{if no such path exists} \end{cases}$$

Thus, in the simplest case of Bernoulli variables, we have

$$x_k = \begin{cases} 1 & \text{if the } X_k \text{ edge is open}, \quad E(X_k=1)=\pi_k \text{ for each } k \\ 0 & \text{if the } X_k \text{ edge is not open}, \quad E(X_k=0)=1-\pi_k \text{ for each } k \end{cases}$$

for each of the 5 edges of the network, $k=1, 2, \ldots, 5$. On this basis, if a path from A to B exists in a particular realization, then the probability of that realization may be determined. For example, the probability of the following realization, $A\xrightarrow{1}B\xrightarrow{3}C\xrightarrow{5}D$, in which a transmission from A to D will occur based on the open edges 1, 3, and 5, is

$$P(X_1=1, X_2=0, X_3=1, X_4=0, X_5=1)=\pi_1(1-\pi_2)\pi_3(1-\pi_4)\pi_5$$

Summing the probabilities of those realizations among the 32 in which a path from A to D exists, the probability of a transmission from A to D is

$$\rho_{AD}=\sum_{\mathbf{x}:\phi(\mathbf{x})=1}^{32}\prod_{k=1}^{5}\pi_k^{x_k}(1-\pi_k)^{1-x_k}$$

where $\rho_{AD}$ is the exact probability value of a transmission from A to B. This probability value is the network's reliability with respect to a transmission from A to B during some period of time. The calculation of this reliability value entails the evaluation of each of the 32 edge-induced subgraphs to determine the outcome of the structure function (whether or not a particular realization presents at least one path from A to D) and the calculation of the probability of each realization. With these values in hand, the expected value of the outcome of the structure function is determined.

A challenging computational problem arises when the number of realizations (subgraphs) that must be evaluated is large, e.g., 1,073,741,824 realizations in the case of a network with 30 edges, for a particular pair of nodes. If the network with these 30 edges has 10 nodes, and if the investigator is analyzing the reliability of a transmission from each node *i* that is a transmitter to each other node *j* that is a potential receiver of *i*'s transmission, then the number of realizations involved in the evaluation may increase by a factor of 45 or 90, depending on whether the network is a graph or diagraph.

Exact reliability values also may be obtained from the minimal paths or minimal cuts of the network. A minimal path is a minimal set of open (active) edges that enable a transmission from one node to another. A minimal cut is a minimal set of closed (inactive) edges that disable a transmission from one node to another. For the (A, D) pair in the illustrated network, the minimal paths are $X_1X_4$, $X_1X_3X_5$, and $X_2X_5$, and the minimal cuts are $X_1X_2$, $X_4X_5$, $X_2X_3X_4$, and $X_1X_5$. This approach also is computationally intensive. Both methods reduce to the same result. For the illustrated network, the result is

$$\begin{aligned}\rho_{AD} &= P\{\phi(X_1, X_2, X_3, X_4, X_5)=1\} \\ &= \pi_1\pi_4+\pi_1\pi_3\pi_5+\pi_2\pi_5-\pi_1\pi_3\pi_4\pi_5-\pi_1\pi_2\pi_4\pi_5 \\ &\quad \pi_1\pi_2\pi_3\pi_5+\pi_1\pi_2\pi_3\pi_4\pi_5 \\ &= 2\pi^2+\pi^3-3\pi^4+\pi^5, \quad \text{if} \quad \pi_k=\pi \quad \text{for all} \quad k\end{aligned}$$

The expansion of latter involves $2^z-1$ terms, where $z$ is either the number of minimal paths or the number of minimal cuts. In the simplest case of homogeneous probabilities, with $\pi=0.60$, the reliability of transmission from A to B is $\rho_{AD}=0.6250$.

*Classic reliability bounds*

When obtaining an exact reliability value via analysis is impractical, as it frequently is, reliability bounds $L\leq\rho_{ij}\leq U$ may be obtained more or less readily, depending on the type of bound. The literature on network reliability theory presents alternative approaches for obtaining analytical reliability bounds and the problem of obtaining them is the subject of ongoing research. The classical approach, with which optimal lower and upper bound values are obtained, also is based on the minimal path set and the minimal cut set. The upper bound is based on the minimal path set and the lower bound is based on the minimal cut set.

For a minimal path set, the approach to the upper bound is based on the result that the reliability of a transmission from $i$ to $j$ is equivalent to the probability that least one of the minimal paths is active, i.e., $\rho_{ij} = 1 - P(\text{all paths fail})$, and that the value $P(\text{all paths fail})$ must be greater than or equal to the product of the probabilities of each minimal path's failure. For the illustrated network,

$$\begin{aligned}\rho_{AD} &= P\{\phi(X_1, X_2, X_3, X_4, X_5) = 1\}\\ &= 1 - P(X_1X_4 = 0, X_1X_3X_5 = 0, X_2X_5 = 0)\end{aligned}$$

and

$$\begin{aligned}\rho_{AD} &\le 1 - P(X_1X_4 = 0)P(X_1X_3X_5 = 0)P(X_2X_5 = 0)\\ &= 1 - (1 - \pi_1\pi_4)(1 - \pi_1\pi_3\pi_5)(1 - \pi_2\pi_5)\\ &= 1 - (1 - \pi^2)^2(1 - \pi^3), \quad \text{if} \quad \pi_k = \pi \quad \text{for all} \quad k\end{aligned}$$

under the assumption of independent random variables. Thus, in the simplest case of homogeneous probabilities, with $\pi = 0.60$, the above calculation presents the upper bound $\rho_{AD} \le 0.6789$.

For a minimal cut set, the approach to the lower bound is based on the result that the reliability of a transmission from $i$ to $j$ is equivalent to the probability that least one of the edges in *each* minimal cut is active, in which case all of the potential cuts fail to eliminate all paths from $i$ to $j$, i.e., $\rho_{ij} = 1 - P(\text{all cuts fail})$, and on the result that the probability value $P(\text{all cuts fail})$ must be greater than or equal to the product of the probabilities of each minimal cut's failure. For the illustrated network,

$$\begin{aligned}\rho_{AD} &= P\{\phi(X_1, X_2, X_3, X_4, X_5) = 1\}\\ &= P\left\{\begin{matrix}[1-(1-X_1)(1-X_2)] = 1, \quad [1-(1-X_4)(1-X_5)] = 1,\\ [1-(1-X_2)(1-X_3)(1-X_4)] = 1, \quad [1-(1-X_1)(1-X_5)] = 1\end{matrix}\right\}\\ &= P\left\{\begin{matrix}[(1-X_1)(1-X_2)] = 0, \quad [(1-X_4)(1-X_5)] = 0\\ [(1-X_2)(1-X_3)(1-X_4)] = 0, \quad [(1-X_1)(1-X_5)] = 0\end{matrix}\right\}\end{aligned}$$

and

$$\begin{aligned}\rho_{AB} &\ge P\{[(1-X_1)(1-X_2)] = 0\}P\{[(1-X_4)(1-X_5)] = 0\}\\ &\quad P\{[(1-X_2)(1-X_3)(1-X_4)] = 0\}\\ &\quad P\{[(1-X_1)(1-X_5)] = 0\}\\ &= [1-(1-\pi_1)(1-\pi_2)][1-(1-\pi_4)(1-\pi_5)]\\ &\quad \times [1-(1-\pi_2)(1-\pi_3)(1-\pi_4)][1-(1-\pi_1)(1-\pi_5)]\\ &= [1-(1-\pi)^2]^3[1-(1-\pi)^3], \quad \text{if} \quad \pi_k = \pi \quad \text{for all} \quad k\end{aligned}$$

under the assumption of independent random variables. Thus, in the simplest case of homogeneous probabilities, with $\pi = 0.60$, the above calculation presents the lower bound $\rho_{AD} \ge 0.5548$.

In sum, for the illustrated network, in the simplest case of homogeneous independent probabilities, with $\pi = 0.60$, the exact reliability of transmission from A to B is $\rho_{AD} = 0.6250$, and, if it were impractical to calculate this exact value (which is obviously not the case here) and practical to calculate bounds (which is obviously the case here), the bounds would be $0.5548 \le \rho_{AD} \le 0.6789$, with the lower bound based on the minimal cut set of the network and the upper bound based on the minimal path set.

*A numerical approach to exact reliabilities*

The network transmission reliability from node A to B is equivalent to the probability of a realization in which A reaches B via one or more paths. Given the adjacency matrix $\mathbf{G} = [g_{ij}]$ for an observed network, a network reliability analysis requires a measurement model the $n \times n$ probability matrix $\mathbf{\Pi} = [\pi_{ij}]$, where $\pi_{ij} = 0$ if $g_{ij} = 0$, and $0 \le \pi_{ij} \le 1$ if $g_{ij} = 1$, for all $i$ and $j$. Estimates of exact network transmission reliabilities may be obtained directly from the specified probability matrix. A numerical approach to approximate reliability values is feasible with Monte Carlo simulations of random realizations, where in each such realization, $\mathbf{R}_s = [r_{sij}]$, $r_{sij} = 0$ if $g_{ij} = 0$, and $r_{sij} = 1$ if $\nu_{ij} \le \pi_{ij}$, where $\nu_{ij}$ is a randomly selected value from the uniform distribution. Thus, the sample of realizations, $\mathbf{R}_s$ ($s$ = 1, 2, . . . ), are random networks consistent with the specified probability matrix $\mathbf{\Pi}$ for the contact network. The reachability matrix for each realization will indicate whether or not at least one path exists from a particular node $i$ to a particular node $j$. Based on $T$ Monte Carlo generated random realizations, a numerical estimate of the reliability of a transmission from node $i$ to node $j$ is the proportion of the $T$ trials in which $i$ reaches $j$.

With the above approach implemented on the illustrated network, and

$$\mathbf{G} = \begin{bmatrix}0&1&1&0\\0&0&1&1\\0&0&0&1\\0&0&0&0\end{bmatrix}, \quad \mathbf{\Pi} = \begin{bmatrix}0&0.60&0.60&0\\0&0&0.60&0.60\\0&0&0&0.60\\0&0&0&0\end{bmatrix}$$

a numerical estimate of $\rho_{14} = \rho_{AD} = 0.6150$ is obtained based on $T$ = 2, 000 Monte Carlo trials. Note that this estimate is close to the exact analytically determined value of 0.6250. Five additional estimates (replications based on the same numerical procedure) are also close to this exact value: 0.6155, 0.6285, 0.6085, 0.6270 and 0.6130.

This numerical approach may be implemented on a heterogeneous probability matrix. For example, based on

$$\mathbf{G} = \begin{bmatrix}0&1&1&0\\0&0&1&1\\0&0&0&1\\0&0&0&0\end{bmatrix}, \quad \mathbf{\Pi} = \begin{bmatrix}0&0.30&0.45&0\\0&0&0.35&0.50\\0&0&0&0.75\\0&0&0&0\end{bmatrix}$$

a numerical estimate of $\rho_{14} = \rho_{AD} = 0.4685$ based on $T$ = 2000 Monte Carlo trials. This estimate is close to the exact analytically determined value

$$\begin{aligned}\rho_{AD} \quad & P\{\phi(X_1, X_2, X_3, X_4, X_5) = 1\}\\ &= \pi_1\pi_4 + \pi_1\pi_3\pi_5 + \pi_2\pi_5 - \pi_1\pi_3\pi_4\pi_5 - \pi_1\pi_2\pi_4\pi_5\\ &\quad - \pi_1\pi_2\pi_3\pi_5 + \pi_1\pi_2\pi_3\pi_4\pi_5\\ &= 0.4585\end{aligned}$$

Five additional estimates (replications based on the same numerical procedure) are also close to this exact value: 0.4620, 0.4530, 0.4580, 0.4755, and 0.4505.

Clearly, since this numerical approach is based on the reachability matrices of the sampled realizations, estimates of the reliabilities of *all* the ordered pairs of the network are available. For example, with

$$\mathbf{G} = \begin{bmatrix}0&1&1&0\\0&0&1&1\\0&0&0&1\\0&0&0&0\end{bmatrix}, \quad \mathbf{\Pi} = \begin{bmatrix}0&0.60&0.60&0\\0&0&0.60&0.60\\0&0&0&0.60\\0&0&0&0\end{bmatrix}$$

we obtain the estimates

$$\mathbf{P} = [\rho_{ij}] = \begin{bmatrix}0&0.5985&0.7510&0.6275\\0&0&0.6205&0.7530\\0&0&0&0.5940\\0&0&0&0\end{bmatrix}$$

and with

$$\mathbf{G} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{\Pi} = \begin{bmatrix} 0 & 0.30 & 0.45 & 0 \\ 0 & 0 & 0.35 & 0.50 \\ 0 & 0 & 0 & 0.75 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

we obtain the estimates

$$\mathbf{P} = [\rho_{ij}] = \begin{bmatrix} 0 & 0.3175 & 0.5155 & 0.4575 \\ 0 & 0 & 0.3605 & 0.6240 \\ 0 & 0 & 0 & 0.7430 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

in which the $\rho_{14} = \rho_{AD}$ estimate is an embedded element. Moreover, since efficient algorithms exist for obtaining reachability matrices, obtaining such transmission reliability estimates is practical for large networks.

*Theoretical analyses*

The present article draws on network reliability theory to present a theoretical analysis of the implications of unreliable edges and path redundancies in large-scale "small world" networks. For this analysis, an approach is employed that deals with conservative lower bounds. These bounds are based on edge disjoint paths in idealized networks with a regular form of clustered edges. The approach is detailed in the body of the body of the article and is not repeated here. Given the techniques and illustrations presented in the previous sections of this appendix, the reader may appreciate why a closed-form general expression for reliabilities was enabled by the idealized role structure considered in this article, and the generic spine-segment components of this role structure.

## References

Centola, D., Macy, M., 2007. Complex contagions and the weakness of long ties. American Journal of Sociology 113, 702–734.

Colbourn, C.J., 1987. The Combinatorics of Network Reliability. Oxford University Press, New York.

Friedkin, N.E., 1998. A Structural Theory of Social Influence. Cambridge University Press.

Granovetter, M.S., 1973. The strength of weak ties. American Journal of Sociology 78, 1360–1380.

Granovetter, M.S., 1983. The strength of weak ties: a network theory revisited. Sociological Theory 1, 201–223.

Hillier, F.S., Lieberman, G.J., 1980. Introduction to Operations Research, 3rd ed. Holden-Day, San Francisco.

Johnsen, E.C., 1985. Network macrostructure models for the Davis-Leinhardt set of empirical sociomatrices. Social Networks 7, 203–224.

Moody, J., 2002. The importance of relationship timing for diffusion: indirect connectivity and STD infection risk. Social Forces 81, 25–56.

Moody, J., White, D.R., 2003. Structural cohesion and embeddedness: a hierarchical concept of social groups. American Sociological Review 68, 103–127.

Newman, M.E.J., 2000. Models of the small world: a review. Journal of Statistical Physics 101, 819–841.

Newman, M.E.J., Barabási, A.-L., Watts, D.J., 2006. The Structure and Dynamics of Networks. Princeton University Press, Princeton.

Ross, S., 2007. Introduction to Probability Models. Academic Press, New York.

Strogatz, S.H., 2001. Exploring complex networks. Nature 410, 268–276.

Watts, D.J., 1999. Small Worlds. Princeton University Press, Princeton.

Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of 'small-world' networks. Nature 393, 440–442.

White, H.C.S., Boorman, S.A., Brieger, R., 1976. Social structure from multiple networks. I. Blockmodels of roles and positions. American Journal of Sociology 81, 730–781.