7

THE COHESIVENESS OF BLOCKS IN SOCIAL NETWORKS: NODE CONNECTIVITY AND CONDITIONAL DENSITY

*Douglas R. White**
Frank Harary†

This study shows several ways that formal graph theoretic statements map patterns of network ties into substantive hypotheses about social cohesion. If network cohesion is enhanced by multiple connections between members of a group, for example, then the higher the global minimum of the number of independent paths that connect every pair of nodes in the network, the higher the social cohesion. The cohesiveness of a group is also measured by the extent to which it is not disconnected by removal of 1, 2, 3, . . . , k actors. Menger's Theorem proves that these two measures are equivalent. Within this graph theoretic framework, we evaluate various concepts of cohesion and establish the validity of a pair of related measures:

This paper was prepared under NSF grant #BCS-9978282, "Longitudinal Network Studies and Predictive Cohesion Theory" (White 1999), and revised with Santa Fe Institute support from John Padgett. It provides the foundational theoretical basis on which the hypotheses of the large-scale longitudinal network studies of the research grant are based. We thank James Moody and anonymous *American Journal of Sociology* reviewers of Moody and White (2000) and White et al. (2001) for a great many useful comments and suggestions made in early revisions of this paper. Two anonymous reviewers of *Sociological Methodology* made extensive suggestions for intermediate and final revisions, and we are indebted to Phil Bonacich and Santa Fe Institute participants in the Working Group on Co-Evolution of Markets and the State—John Padgett, Michael Heaney, Walter Powell, David Stark, Sander van der Leeuw and Sanjay Jain—for useful commentary.

*University of California, Irvine
†New Mexico State University

1. *Connectivity—the minimum number k of its actors whose removal would not allow the group to remain connected or would reduce the group to but a single member—measures the social cohesion of a group at a general level.*
2. *Conditional density measures cohesion on a finer scale as a proportion of ties beyond that required for connectivity k over the number of ties that would force it to k* + 1.

Calibrated for successive values of k, these two measures combine into an aggregate measure of social cohesion, suitable for both small- and large-scale network studies. Using these measures to define the core of a new methodology of cohesive blocking, we offer hypotheses about the consequences of cohesive blocks for social groups and their members, and explore empirical examples that illustrate the significance, theoretical relevance, and predictiveness of cohesive blocking in a variety of substantively important applications in sociology.

Solidarity is a generic concept encompassing multiple ways that individuals coalesce into groups. We can distinguish several kinds of bonds that contribute to solidarity: *members to group*; *members to group norms*; *members to leaders*; and *members to members*. We can conceive of these bonds as having *forms* such as moral rules, norms, incentives, or contexts, and *contents* such as various types of relationships. Form and content, social "facts" of relationships versus norms that might govern them, and other oppositions are not removed from one another. Rather, they are performed, enacted, and understood in networks of interactions. Hence, it is useful to partition solidarity into its *ideational* and *relational* components (Fararo and Doreian 1998), the former referring to the psychological identification of members within a collectivity,[1] and the latter to the connections among the collectivity's members, which can be visualized as graphs.

Within the relational component of solidarity, we can further distinguish two aspects of the form of relations of a group that help to hold it together. What we call *cohesion* is the contribution made by (adding or subtracting) individual members of a group, together with their ties, to

[1]What we call *attachment* of members to groups often involves complex interactions among psychological, dispositional, moral, normative, and contextual concerns, and they are often for this reason difficult to measure and to depict as graphs. To simplify measurement, researchers often try to elicit from individuals indicators of their attachments to groups. The same is true of what we call *adherence to leadership*. Common research questions in this domain are: What are the attractive or charismatic qualities of leaders (or attractions *to* their followers) that create weaker or stronger many-to-one ties or commitments?

holding it together. What we call *adhesion* (edge cohesion) is the contribution made to holding a group together—keeping membership constant—by (adding or subtracting) ties between its members.[2] We ask one of the fundamental questions of sociology: How and when do groups, norms, leaders, and commitments emerge out of cohesive clusters? Alternatively, we can ask: How and when does the formation of groups and the emergence of leaders lead to the transformation of cohesive clusters? Because cohesion often spills over the boundaries of formal groups, dynamic reconfiguration of groups and alliances can be studied in the interplay between these two questions.[3] Alternately, as different groups emerge and overlap, and groups interact at another level of organization, dynamic reconfigurations of cohesiveness can be studied, oscillating between different levels involving questions about top-down and bottom-up effects.[4]

The *content* of ties is also important to how we view the dynamics of cohesion and group transformations. At the simplest level (Harary 1953), negative ties tend to repel and positive ties attract. We are principally concerned here with positive dyadic bonds and the concept and measurement of cohesion as a relational component of social solidarity, where the ties in question are ones that can bond pairs of people in nonexclusive ways that could constitute a basis for positive relations that hold a group together. We do not try to deal here with more subtle refinements of content as they might affect solidarity, but we try to isolate the contribution of different structural forms of connection, given the simplifying assumption of relatively homogeneous "positive" content of ties. The model of cohesion presented here would not be appropriate, for example, if the relation under study was that of conflict or antagonism, where "negative" ties occur (see Harary 1953).

Section I develops a series of assertions of increasing precision that provide intuitive foundations for sociological models of cohesion

[2] In the way we will operationalize these two concepts, they will be closely related, but cohesion will turn out to be the stronger measure since removal of individuals from a group automatically removes their ties, while removal of ties does not entail removal of individuals.

[3] In heterarchic systems, such as a government that derives its legitimacy from "We the people" to guarantee empowerment against intrusions at intermediate levels (Morowitz 2000:11–12), multiple relations contend for and oscillate in their salience for regulatory processes. Such oscillations include centralized systems that are hierarchically organized from upper to lower levels. Alternatives include decentralized systems that are emergent, often hierarchically, from lower to upper levels.

[4] Powell, White, Koput, and Owen-Smith (2001) develop analysis for bottom-up effects in the emergence of new structural forms in collaborative networks among biotechnical firms, for example, as distinct from the top-down effects of co-evolving government policies and agencies affecting the biotech industry.

and, secondarily, adhesion. Sections I.B and I.C provide some general expectations and hypotheses about cohesion and adhesion in terms of their sociological antecedents and consequences. We show the significance of defining structural cohesion as resistance-to-taking-a-group-apart and path cohesion as stick-togetherness (our definitions 1.1.2. and 1.2.2) and discuss the sociological implications of these two facets of cohesion, and similarly for two parallel aspects of adhesion. We show how these two aspects meld in each case into a single equivalent concept that defines the boundaries of social groups at different levels of cohesion or adhesion. Section I.D shows the utility of defining density and closeness of ties within bounded cohesive blocks, which will lead to a later section (III) on conditional density. Section II provides the graph theoretic foundations for concepts and measurements of cohesion and adhesion, and gives the definitions needed for the proofs of equivalence of the twin aspects of structural and path cohesion or adhesion. Section III defines conditional density and a scalable measure of cohesion that combines connectivity and conditional density within nested patterns of subgroup cohesion and subgroup heterogeneity. Section IV examines a case study of a factional dispute in a karate club to exemplify our measures of both cohesion and adhesion, and shows how it is useful as well to take into account relative density and closeness within the bounded context of connectivity subsets. Section V describes how the proposed measures of cohesion have been tested in larger scale sociological studies than the karate example, and Section VI summarizes and concludes our study.

I. TOWARD SOCIOLOGICAL MODELS OF COHESION AND ADHESION

A. Some Basic Intuitions and Concepts

The following series of sociological assertions may help to give the reader some intuitive underpinnings for the models of cohesion and adhesion that follow in Section II. Intuitively, cohesion begins with the role of individuals in holding a group together:

1. *A group is cohesive to the extent that its members possess connections to others within the group, ones that hold it together.*

If cohesion begins with individuals who are connected, higher levels of group cohesiveness should entail that the removal of some one (two, three. . .) actor(s) should not disconnect the group. Simmel (1908 [1950:123]) noted the fundamental difference in this respect between a solitary dyad and a triad:

> The social structure [of the dyad] rests immediately on the one and on the other of the two, and the secession of either would destroy the whole . . . as soon, however, as there is a sociation [clique] of three, a group continues to exist even in case one of the members drops out.

We identify this as the first of two facets of cohesion:

1.1. *A group is cohesive to the extent that it is resistant to being pulled apart by removal of its members.*

Generalizing Simmel's intuition as a structural feature of cohesion, we introduce the following definition:

1.1.1. A group is *structurally cohesive* to the extent that it is resistant to being pulled apart by the removal of a subset of members.

The concept of the robustness of connections under removal of members of a group is closely related to the graph theoretic concept of connectivity given in Section II. There we review the graph theoretic foundations of cohesion and adhesion. Harary et al. (1965) anticipated the approach of utilizing connectivity as a measure of cohesiveness. Wasserman and Faust (1994:115–17) cite his definition of connectivity (Harary 1969) as one way to measure the cohesion of a graph, but they do not apply it to finding cohesive subgroups. White (1998; White et al. 2001) develops the latter idea, and Moody and White (2000) provide and apply an algorithm for measuring maximal subsets of nodes in a graph at different levels of connectivity. We implement this approach here, defining the structural aspect of the cohesion of a group in quantitative terms as follows (the formal exposition here parallels Moody and White):

1.1.2. A group's *structural cohesion* is equal to the minimum number of actors who, if removed from the group, would disconnect the group.

This is the minimum number k of its actors whose removal would not allow the group to remain connected or would reduce the group to a single member. It allows hierarchies of cohesive blocks to be identified. At the highest level, for a clique with n members, all but one member must be removed to get an isolate, so the structural cohesion is defined as $n - 1$. Cohesiveness may be viewed as "the resistance of a group to disruptive forces" (Gross and Martin 1952:553), and structural cohesion provides a relational basis for a group to resist disruption by defection or removal of members.

If resistance to being pulled apart is an aspect of cohesiveness, however, we must pay equal attention to a second, more integrative aspect of cohesion commonly discussed in the literature.[5] This integrative aspect is emphasized in such definitions of cohesiveness as "the forces holding the individuals within the groupings in which they are" (Moreno and Jennings 1937:371); "the total field of forces that act on members to remain in the group" (Festinger et al. 1950:164); and "a dynamic process that is reflected in the tendency for a group to stick together and remain united in pursuit of its goals and objectives" (Carron 1982:124).

This second aspect of the cohesive integration of a social group can be defined and measured—independently of robustness to disruption by removal or defection of members—by the number of distinct ways that members of a group are related. Cohesion increases, for example, when members have multiplex bonds, such as two people who are classmates, friends, and neighbors. Two people may also be connected by multiple independent paths that have no intermediate members in common. We identify such possibilities in the second of two related facets of cohesion:

1.2. *A group is cohesive to the extent that pairs of its members have multiple social connections, direct or indirect, but within the group, that pull it together.*

The integrative aspect of group cohesion that we will examine is the number of independent paths linking pairs of members. We define this aspect of cohesion as follows:

1.2.1. A group is *path cohesive* to the extent that its members have a multiplicity of independent paths between them, *within the group*, that pull it together.

[5] French (1941:370), for example, discussed how a group exists as a balance between "cohesive" and "disruptive" forces, including responses to disruptive forces.

This can be defined quantitatively as follows:

1.2.2. A group's *path cohesion* is equal to the minimum number of its independent paths taken over all pairs of members.

When we define structural and path cohesion formally, as we do in Section II, one of the purposes of the formal language of graphs is to derive as a theoretical result or mathematical proof that the two graph theoretic concepts of path (1.2.2: stick-togetherness) and structural (1.1.2: won't-pull-apart) cohesion are equivalent. This is important to sociological theory, and the mathematics can contribute at the conceptual level to sociological explanation, because it demonstrates that two measurement constructs—each of which is central to the study of social cohesion—can be reduced to one by virtue of their formal equivalence. We do not claim that these are the only constructs relevant to understanding cohesion, but simply that when our substantive and intuitive sociological conceptions converge on these two constructs of structural and path cohesion, they are formally equivalent. Both measures that are thereby unified are highly relevant to the relational component of group solidarity. Although the measurement of connectivities or node-independent paths in social networks is a complex problem computationally, accurate approximation techniques are now available (e.g., White and Newman 2001) for large networks.

Besides the removal of members, the only other way in which a group is vulnerable to disconnection is by removal of ties between members. This possibility defines what we call *adhesion* within a group (which is thus a logically weaker concept than cohesion because removal of individuals is excluded). We arrive at a definition of adhesion by elaborating a series of sociological assertions that run parallel to those about cohesion, except that here the members of a group are held constant, and we consider, between group members, only the ties, whose removal can separate the group:

2. *A group is adhesive to the extent that its members' ties hold it together.*

Facet 1

2.1. *A group is adhesive to the extent that it is resistant to being pulled apart by removal of ties between members.*

2.1.1. A group is *structurally adhesive* to the extent that it is resistant to being pulled apart by removal of a subset of its ties between members.

Facet 2

2.2. *A group is adhesive to the extent that the ties between its members or indirect connections within the group pull it together.*

To count the number of such ties or indirect connections, we refer in this context to paths between two members within a group as being disjoint (technically speaking: edge-disjoint) if none of the dyadic ties that make up the respective paths are the same:

2.2.1. A group is *path adhesive* to the extent that it is held together by (edge-) disjoint paths between each pair of members.

Our two facets of adhesion can be precisely defined as follows:

2.1.2. A group's *structural adhesion* is equal to the minimum number of direct links between group members that, if removed, would disconnect the group.
2.2.2. A group's *path adhesion* is equal to the minimum number of (edge-) disjoint paths among different pairs of members.

Thus we define group adhesion to refer to stick-togetherness/don't-pull-apart in relation to the edges or paths that connect the group but holding constant the members of the group itself.

High adhesion may be obtained without a concomitant rise in cohesion by subgroups that link through one or a few central nodes. Figure 1 gives two examples where in each case the cohesion is minimal (each group is vulnerable to disconnection by removal of a single member) but adhesion is significantly higher (three ties must be removed to disconnect any of the members of these two groups). In the first example there is a structure of many small cliques that have one member in common (e.g., a common "leader" unifying the cliques). In the second, there is the same number of small cliques, but the central member who links them is not a member of the cliques. The examples differ in the structure of cohesion in the subgroups (overlapping versus nonoverlapping cohesive subgroups). The converse, high cohesion with low adhesion, is impossible by defini-

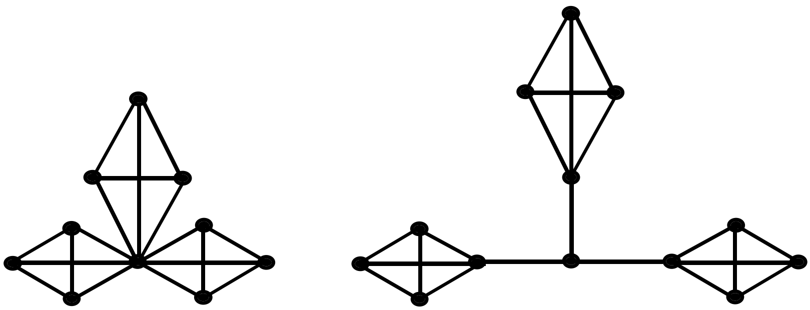


FIGURE 1. Two graphs with low cohesion (one node-removal separates each graph) but high adhesion (a minimum of three edge-removals separates each graph).

tion, which conveys the sense in which adhesion is a weaker measure of cohesion, or perhaps a measure of something quite different, necessary but not sufficient for cohesion.

Our graph theoretic formalization in Section II takes up the two facets of structural and path cohesion, then those of adhesion, shows the mathematical equivalence of these two facets for both cohesion and adhesion and how cohesion and adhesion are mathematically related. Before doing so, under the heading of hypotheses, we examine some of the implications of these two concepts for group structure and social processes. We will not consider the concepts of attachment-to-group, adherence to leadership, or adherence to norms, but these themes might be developed by treating the group, leader, or normative elements and constructs, as distinguished nodes, and investigating individual-to-group or follower-to-leader ties within the broader network, or one- and two-mode networks of individual and normative elements.

B. Hypotheses About Cohesion

Once we develop a graph theoretic methodology that provides formal measures for the intuitive concepts of cohesion, it will be possible in the long run and over many studies to evaluate the social consequences at the group and individual level that are thought to follow from different aspects of cohesion. Differences in cohesiveness, including finer levels of differences in density within each group (see Section I.D), should have recurrent antecedents and/or predictive consequences for social groups and their members across different social contexts.

Our goal here is to develop methodology and not to test hypotheses, so we will not attempt to define relevant measures for possible consequences, at the group or individual level, of our measures of cohesion, but some relevant ideas and examples may be briefly sketched. At the level of formal theory, for example, if the relevant units of time or cost are normalized for the networks of two groups (such as longevity of group members, time-decay of ties, etc.), it is a testable hypothesis that the network with higher structural cohesion (other things being equal) will, in a set period of time, be less likely to separate into two disconnected groups. Formal theory might be developed to predict outcomes such as relative stability of groups, or the relative duration of a group as a social configuration.

Similarly, if information transmission in a network is noisy or unreliable, then compensatory gains in path cohesion should give an initial rapid benefit from higher capacity to transmit redundant information. Further benefits will at some point begin to have diminishing marginal returns, analogous to the declining marginal benefits to reliable measurement of averaging more independent measures of the same variable. This, and the fact that adding links in a network typically has a cost, leads to the hypotheses that, when transmission is unreliable (or quality decays with distance), measurable benefits to gains in connectivity will be high initially, but growth of cohesion at higher levels will tend to be self-limiting because of a rising cost/benefit ratio. As compared to suboptimal levels of cohesion in a network in this context, near-optimal levels are hypothesized to occur at very low densities for large networks and, given sufficient stability, to predict higher congruence in information transmittal, higher levels of consensus among group members, more rapid emergence of group norms, and higher levels of effective coordination in mobilizing group action or exerting group-level influence. Relations for which path cohesion might be especially predictive of relevant outcome variables are those that serve as conduits for items that are transmitted in social networks, such as information, gossip, disease, or favors or goods exchanged.

Engineering applications of network concepts of transmission generally emphasize designing networks so that distance decay in signal transmission is compensated by intermittent amplifiers that dampen noise and boost coherent signals. It is often assumed by social theorists who utilize concepts of transmission or flow in social networks that such amplifiers are absent in naturally occurring social interactions. The feedback circuits required for amplification, however, are found in cohesive blocks. Path cohesive blocks of such networks thus might be hypothesized to serve as natural amplifiers in social networks, boosting signal by creating internal patterns of coherence. For the phenomenon of network externalities in which a product's value is enhanced by additional users, the early exponential rise in adoption may be accelerated by path cohesion, or intermittently decelerated if the cohesion is concentrated in distinct pockets.[6]

The Internet is an example of a type of network in which redundancies facilitate transmission and the emergence of cohesive pockets and hierarchies of users and sites. Cohesion, not adhesion, was the object of the packet-switching transmission design through multiple pathways.[7] The physical elements of the ARPANET/Internet system, such as links among servers, also required cohesion and not just adhesion. At a third level, links between Web pages, cohesion has been an emergent phenomenon of potential benefit to users.

Path cohesion also operates as resistance to a group's being pulled apart through bonding effects that are independent of distance rather than subject to distance decay. It is k times more difficult to break apart two nodes if they have k independent chains of connections than it is to break them apart if they have a single chain of connections. Hence higher path cohesion is an indicator of a group's resistance to being pulled apart even with transmission decay or its absence altogether.

Long-range bonding effects may operate through chains of connectivity even in sparse networks. Grannis (1998), for example, found that the best predictor of contiguous zones of homogeneity in urban neighborhoods is not closeness of ties or walking or driving distance, but chaining of neighbor relations along residential streets. These bonding chains do not imply that members of the homogeneous sets have a high density of neighbor relations, or high door-to-door transmission rates, but that they have chains of neighboring by which members of the homogeneous group

[6] In this case, however, connectivity theory would suggest a critical threshold where the rise in added value goes from linear (because component sizes grow linearly with adoption up to this threshold) to exponential (because after the threshold is reached, sizes of component and/or cohesive sets begin to grow exponentially), to dampened marginal returns.

[7] If the designers of ARPANET, the military forerunner of the multiple-path Internet packet technology, had only been concerned with adhesion, the system might have been designed around a single central hub that would have left it vulnerable to attack.

are neighbors of neighbors of neighbors, etc., without constraint on path length. Homogeneity tends to be transitive through these local bonds rather than decaying with distance. Grannis (personal communication) hypothesized that structural cohesion contributes to neighborhood homogeneity but did not test the hypothesis directly.

For individuals, membership in one or more groups with differing levels of structural and path cohesion might also have predictive consequences for levels of attachment and participation in the group or the larger community in which the individual is embedded.

Are cohesive blocks in social networks equivalent to what Granovetter (1973, 1983) identified as “strong” as opposed to “weak” ties—namely, ties of high multiplexity where dyadic interaction is frequent? If strong ties tend to cluster due to greater transitivity than weak ties, structural cohesion within clique-like structures would follow as a consequence. The two concepts are not equivalent, however, and structural cohesion does not imply such transitivity. It is an open question as to whether the relevant circumstances in which weak ties are more effective than strong ones for network reachability include effects of structural cohesion.

C. Hypotheses About Adhesion

Structural and path adhesiveness in social networks is not our primary concern here, but hypotheses about the effects of adhesion should be considered alongside those concerning cohesion. Higher levels of adhesion imply more channels of connections between pairs of members of a group, even if the channels are not strictly parallel but may run through the same (e.g., central) nodes. High adhesion networks with low cohesion (e.g., Figure 1), like graph centralization, may lead to vulnerability of strategic points of control in social networks.[8] Hence, adhesion is a direct measure of the potential for flow between nodes, without considering the potential for congestion or vulnerability through central nodes. Centralized polities and

[8]When adhesion and cohesion in a graph are minimal, but the graph is connected, and the maximum distance in the graph is two, the graph is maximally centralized as an egocentric star. In general, high adhesion relative to low cohesion within a social group implies that the network is relatively centralized, but the pattern of centralization may be one of a central node connecting a number of cohesive outliers. The relation between centrality measures and graph theoretic measures of adhesion and cohesion (edge and node connectivity) is complex and deserves separate treatment.

bureaucracies, for example, may require high adhesion networks if the latter entail the potential for establishing control points internal to social groups.

Social fragmentation is a domain where adhesive groupings have been used to hypothesize about lines of cleavage, given the presence of internal conflict within a previously solidary group. Zachary (1975, 1977) argued that when conflicts between rival leaders within a group are sufficiently intense, the group would segment through dissolution of the minimum number of links needed to separate into two. Anticipating our analysis of this case in Section IV, we take issue with this adhesion-based hypothesis for several reasons. First, the idea that a minimum number of edges can be removed to disconnect a graph does not imply that a unique set of edges will be identified by this criterion. The graph in Figure 2, where nodes 1 and 2 represent two leaders in conflict, helps to illustrate the critique. The removal of pairs of edges a, b will disconnect 1 from 2, but so will removal of the two edges c, d, or the edges e, f, or e, h. Second, we would argue that the persons who are more cohesively linked to leaders, with closeness of ties as a secondary factor, are more likely to remain as members of their respective factions. Hence in Figure 2 persons 3 and 5 will be more likely to remain with 1 and persons 4 and 6 with leader 2. This illustrates that minimum edge-removal is not the only way to effectively divide the group. Further, there is a single node in Figure 2 whose removal would disconnect the graph. Hence the social pressure to take sides might be greatest for this person, node 4, who is also the only member of the two most cohesive sets, 1-3-4-5 versus 2-4-6, each of which is circled in Figure 2. On the basis of cohesiveness and tie density, we would expect that person 4 would remain in leader 2's faction, hence expected to decouple from 3 and from 5 rather than from 2 and 6. The end result is the same as edge

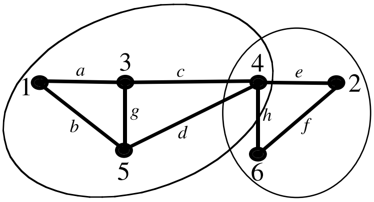


FIGURE 2. Cohesive blocks in a factional dispute between leaders 1 and 2.

removal, in which nodes remain and edges are removed. We would argue, however, that the underlying social dynamics of segmentation are more likely to involve the agency of individual decision-makers, hence favoring hypotheses about effects of cohesion, closeness, and the density of cohesive blocks rather than effects of adhesion, or minimum removal of edges. Since edges do not autoregulate independently of the agents they connect, we might hypothesize that while cohesive blocks are likely to influence the development of norms and sanctions about what kinds of edges with outsiders are favored or disfavored, adhesive subsets that are not also cohesive are less likely to do so.

D. Hypotheses that Integrate Other Aspects of Cohesion: *Conditional Density and Other Variables*

Given a first level of structural and path cohesion at which graphs can be ordered by their invulnerability to disruption (connectivity) and inversely to the ties that hold them together (equivalent definitions 1.1.2 and 1.2.2), additional finer levels of cohesion can be ordered by density, closeness of ties, and other factors.[9] We can define a strict measure of conditional density (Section III) because a certain density and minimum number of ties within a group are already implied by its level of connectivity.[10] Within the boundaries of cohesive blocks established by structural and path cohesion, our graph theoretic formalization in Section II allows us to describe a finer level of cohesion reflecting aspects such as density, that make a secondary contribution to cohesion.[11] At this second level of cohesion, a *conditional definition* related to cohesion is one based on a feature of cohesion that lends itself to a finer level of measurement, where the feature has a necessary minimum and maximum value associated with the level of structural and path cohesion, and is measured within the bounded cohesive set associated with that level.

Conditional density will be a primary focus of our development of a scalable measure of cohesion that takes an integer value k for a certain

[9] We shift here to refer to the equivalent concepts of structural cohesion and path cohesion simply as “cohesion.” Our assertion in doing so is that these concepts capture two of the most important facets of many-to-many ties among clusters of individuals as they form into cohesive blocks. Throughout this section, we are concerned with how other variables also contribute to cohesion, with density as a case in point.

[10] This is not the case for multiplexity (multiple types of ties between group members) or for frequency of interaction between members of a group, which can vary independently of structural and path cohesion.

[11] This could also be done for diameter, closeness of connections, or adhesion as secondary aspects of cohesion.

level of structural and path cohesion for a uniquely defined subgroup of a social network, plus a decimal value (between 0 and 1) for an added contribution to the robustness of cohesion within the group made by within-group density. It provides a useful example of how complementary aspects of cohesion can be measured. These ideas toward a general methodology for the cohesive blocking of social networks must remain intuitive until conditional density is given formal explication in Section III.

One purpose of our graph theoretical formalization of the concepts of structural and path cohesion and conditional density is to show the advantages of a scalable aggregate measure of cohesion in social groups over other approaches that use relative densities of different clusters of nodes in social networks to try to identify the boundaries of cohesive social groups. Clique-finding algorithms, for example, give a unique inventory of cohesive clusters whose relative density is maximal, but the denser regions of social networks will often exhibit many intersecting cliques. The intersections among cliques form a lattice that typically defines a welter of intersecting social boundaries (Freeman 1996). As a density threshold for overlapping subgroups is relaxed, the overlap of subgroups rises exponentially. Forcing social group detection into a framework of mutual exclusion, on the other hand, makes little substantive sense when the concept of structural cohesion provides a meaningful framework for detecting and interpreting multiple group memberships. Structural and path cohesion are also able to detect more distributed patterns of cohesion in social networks than the unions of intersecting cliques.

In contrast to approaches that use a density criterion alone, different levels of structural and path cohesion will typically define hierarchical nesting of relatively few bounded cohesive subgroups, with severe limits on the overlap between hierarchical clusters, and the lower the level of cohesion, the less the overlap, hence high coherence of structure. Without offering detailed hypotheses,[12] but judging from case studies presented in Section IV or reviewed in Section V, along with those in Moody

[12] "Invisible colleges" in intellectual and citation networks, for example, are likely to be predicted from cohesive blocks and conditional densities, with downstream predictions from measures of cohesion to other group and individual level sociological effects. Similar models of effects might be applied to the idea of cohesive blocks of infectious sites in epidemiology. Identification of cohesive blocks and conditional densities in networks of economic exchange may provide a means of identifying cores and hierarchies in economic systems, or intensive markets for particular product clusters. In analysis of distributed cohesion in large networks such as these, the study of overlapping cliques will usually fail to identify cohesive sets that are anywhere near as large as those identified by structural and path cohesion.

and White (2000), we have reason to think that structural and path cohesion, along with additional density criteria within cohesive blocks, will be found to have important consequences for many different types of social groups, and that the graph theoretic concepts of cohesion and conditional density will find a useful explanatory niche in sociological theorizing. We would insert a caveat, however—namely, that consideration of the type of social relations being studied ought to guide how the concepts of adhesion and cohesion might be used, either separately or in combination with conditional density or other measures such as distance on shortest paths.

The emergence of trust in social groups, for example, might depend on both level of structural cohesion and the relative compactness of cohesive blocks—in terms of interpersonal distances on shortest paths—in the following way. In such groups, each individual A might receive information concerning each other group member B through a variety of paths that flow through distinct sets of intermediaries, and if the distances from A to B are sufficiently short, then A will be able to interpret these multiple independent sources of information about B's characteristics or identity as a person but as seen or filtered by a variety of others. This ability to compare independent perspectives on each of the others in the group is conducive to discriminations concerning trust and distrust. Not everyone in such groups will necessarily be trusted, but the conditions fostered by comparisons within such groups would provide a reliable basis for informed judgments as to trust. Hence more elaborated discriminations about trust (and hence the emergence of high-trust networks) might be expected to be more frequent within such groups.

As a hypothesized basis for interpersonal trust, the model of connectivity plus conditional density is one in which conditional distance also needs to be taken into account. As in the case of density, we can determine the minimum and maximum possible diameter of a group (the largest shortest-path distance for any two members) and the corresponding minimum and maximum average shortest-path distances between members, if we know the number of its members and the structural cohesion of the group. Watts (1999a, b) defines a "small world" as a large network with local clustering of ties but relatively low average distance between members.[13] We do not develop here a model of conditional distance analogous

[13] He shows that, for model networks with many nodes, high local clustering, and very high average distances between nodes, successive random rewiring of edges produces a small world rather quickly by creating shortcuts that shorten the average internode distances. Small worlds and conditional distance are well worth further investigation in relation to structural and path cohesion.

to that of conditional density, although it would be possible to do so. Rather, we note that within our model of conditional density, as edges are added randomly to a graph at a certain level of structural and path cohesion (thereby adding to conditional density), average internode distances will also fall quickly, resulting in the closeness of a "small world." Assuming that increases to conditional density for groups at a given level of structural cohesion occur by random addition of edges, conditional density becomes a proxy for conditional distance. Precise measures of conditional distance could be constructed to develop further the methodology of cohesive blocking, and our conditional density measure is instructive as to how one might proceed to do so.

II. THE GRAPH THEORETIC FOUNDATIONS OF COHESION AND ADHESION

For clarity of presentation in formalizing definitions and theorems about social cohesion and adhesion, a social relation hypothesized or assumed to contribute to cohesion or adhesion is considered as a graph. This allows us to equate the sociological definitions 1.1.2 and 1.2.2 for cohesion and 2.1.2 and 2.2.2 for adhesion with corresponding graph theoretic definitions of node and edge connectivity and, using the theorems of Menger (1927), to establish the equality between the two fundamental properties of cohesion and adhesion: resistance to being pulled apart (definitions 1.1.2 and 2.1.2), and stick-togetherness (1.2.2 and 2.2.2). Our goal in this section is to provide a formal methodology with appropriate graph theoretic terminology for the cohesive (or adhesive) blocking of social networks.

A *graph* $G = (V, E)$ consists of a set V of n *nodes* or *vertices* and a set E of m *edges* each joining a pair of nodes. We say G has *order* n and *size* m. The two nodes in each unordered pair (u, v) in E are said to be *adjacent* and constitute an edge that is *incident* with nodes u and v.[14] A *path* in G is an alternating sequence of distinct nodes and edges, beginning and ending with nodes, in which each edge is incident with its preceding and following node. A graph is *connected* if every pair of nodes is joined by a path. The *distance* between two nodes in G is the minimum

[14] A group with nonsymmetric relations is representable by a digraph $D = (V, A)$ consisting of a set V of *nodes* and a set A of *arcs* (directed edges) consisting of ordered pairs of nodes in V. A more complex but also more general derivation of our results regarding measures of cohesion, applicable to digraphs, was done by Harary, Norman and Cartwright (1965, ch. 5).

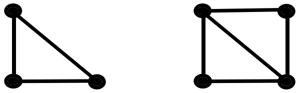


FIGURE 3. A disconnected graph with two components and three cliques, each a K_3.

size of a path of G that connects them. A *subgraph* of a graph G is a graph having all of its nodes and edges in G.

Since we regard a social group with nondirected interpersonal relations as part of a social network—as a subgraph of a larger graph—it is useful to provide some definitions about subgraphs. A set S is *maximal* (*minimal*) with respect to some property if no proper superset (subset) of S, containing more (fewer) elements than S, has the property but S does. A *component* of G is a maximal connected subgraph. A *complete graph* K_n of order n has every pair of nodes adjacent. A *clique* of a graph G is a maximal complete subgraph of G of order at least 3, hence a maximal subgraph K_n of G of order $n \geq 3$. Figure 3 shows a disconnected graph with two components and three cliques, each a K_3.

A. Connectivity and Resistance to Pulling Apart by Removal of Nodes

Two primary references on the node and edge connectivities of G, denoted by κ (kappa) and κ', respectively, are Harary (1969, ch. 5) and Tutte (1966). The *removal of a node v* from G leaves the subgraph $G - v$ that does not contain v or any of its incident edges. The (*node*-) *connectivity* $\kappa(G)$ is defined as the smallest number of nodes that when removed from a graph G leave a disconnected subgraph or a trivial subgraph.[15] The connectivity of a disconnected graph is zero as no nodes need to be removed; it is already not connected. Our definition 1.1.2 (section I.A) corresponds, in the terminology of graph theory, to that of the (node) connectivity of a graph. The *trivial graph* K_1 of one node and no edges ($n = 1$ and $m = 0$), or a disconnected graph, has cohesion 0. A solitary dyad has cohesiveness 1, a triad has 2, and a 4-clique has 3.

[15] This two-part definition is needed because no matter how many nodes are removed from a complete graph, the remaining subgraph remains complete and hence connected until the trivial graph with one node is obtained, and we do not remove it since its removal leaves emptiness. Thus connectivity is defined as $n - 1$ for the complete graph K_n.

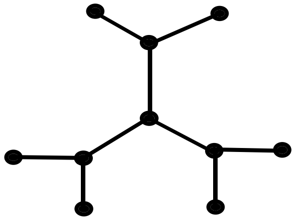


FIGURE 4. A graph G that is a star.

A *cutnode* of a connected graph G is one whose removal results in a disconnected graph. A set S of nodes, edges or nodes and edges *separates* two nodes u,v in a connected graph G if u and v are in different components of $G - S$. A *node cut set* (*cutset*) is a set of nodes that separates a connected graph into two components. An *endnode* is one with a single incident edge. Its removal does not separate a graph. A *cycle* of order n nodes, designated C_n, is obtained from a path P_n with $n \geq 3$ by adding an edge joining its two endnodes. Two paths are (node-) *disjoint* or *node-independent* if they have no nodes in common other than their endnodes. A cycle containing nodes u,v entails that u and v are joined by two (node-) disjoint paths. A *tree* is a connected graph with no cycles, as in Figure 4. It is easy to see that each node in a tree is either an endnode or a cutnode.

A connected graph has connectivity 1 if and only if it has a cutnode. Thus a tree has connectivity 1 but a cycle does not; it has $\kappa = 2$. In Figure 5, which shows the eleven graphs of order 4, the first five graphs are disconnected (the first graph is *totally disconnected*), while the remaining six are connected and thus consist of a single component. The first two connected graphs are the trees of order four. The last is the complete graph K_4. The second of size 4 is the cycle C_4. The graph before C_4 has a cutnode and hence connectivity 1.

A maximal connected subgraph of G with connectivity $k > 0$ is called a *k-component* of G, with synonyms *component* for 1-component, *bicomponent* for 2-component (called a *cyclic component* by Scott 2000:105) and *tricomponent* for 3-component (called a *brick* by Harary and Kodama 1964). In Figure 5 graphs 5 and 8 have bicomponents of order 3—namely, triangles; the three graphs from C_4 to K_4 have bicomponents of order 4; and the complete graph K_4 is itself a tricomponent.

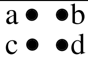
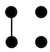
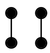
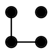
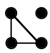
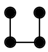
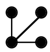
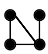
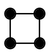
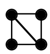
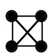
	Graph	Size	Number of Components	Connectivity	Type
1		0	4	0	
2		1	3	0	
3		2	2	0	
4		2	2	0	
5		3	2	0	
6	P_4 	3	1	1	Path (tree)
7	$K_{1,3}$ 	3	1	1	Star (tree)
8		4	1	1	
9	C_4 	4	1	2	Cycle
10		5	1	2	
11	K_4 	6	1	3	Complete graph

FIGURE 5. The eleven graphs of order 4.

A *block* of G is a maximal connected subgraph with no cutnodes (Harary 1969; Even 1979; Gibbons 1985).[16] The blocks of a graph give a partition of its edges. In Figure 5 there are three graphs that are single blocks: C_4 and the last two graphs. Graphs 2 and 5 contain a single block plus isolated nodes. There are two K_2 blocks in graphs 3 and 4 and two blocks in the graph before C_4. Three blocks are contained in each of the two trees of order 4, since each edge is a block. A block may contain a solitary dyad (not contained in a cycle) whereas a bicomponent is a block in which there are 3 or more nodes.

[16] Scott (2000:108,187 fn. 9), owing to the fact that block has another meaning in network analysis, uses the unnecessary and unfortunate term *knot*, easily confounded with the established term with another meaning in topology. Everett (1982a, 1982b) deals separately with both types of block.

A *cohesive block* of a graph *G* (a term we define here for use in sociological analyses of cohesion) is a *k*-component of *G* where the associated value of connectivity defines the cohesion of the block. We use *cohesivity* to refer to cohesive blocks of $\kappa = 2$ or more.[17] Some of the most commonly used network measures of cohesion, as we will show below, lack a guarantee of cohesivity, or even of connectedness. Within blocks of connectivity 1 will be nested more cohesive blocks, if any, of higher connectivity.[18] We may use the term *cohesive groups* to refer to substantive contexts where this concept has been applied to identify social groups on the basis of their network connectivities. We use *cohesive subsets* to refer to subgraphs of a graph *G* that may be cohesive in some respects but do not necessarily correspond to cohesive blocks defined by connectivities of subgraphs.

B. Edge Connectivity, and Resistance to Pulling Apart by Removal of Edges

As distinct from node removal, the *removal of an edge e* from *G* leaves the subgraph $G - e$ that contains all the nodes of *G*. Edge removal presents a lesser vulnerability to a graph being pulled apart than node removal, which removes all incident edges. An edge of a connected graph whose removal results in a disconnected graph is called a *bridge*.[19] The *removal of a set of edges* in *G* is the successive removal of each edge *e* in the set. An *edge-cutset* (or *edge-cut*) of a connected graph *G* is a set of edges whose removal results in a disconnected graph. The *edge connectivity* κ'

[17] The motivation for coining the technical term *cohesivity* is that blocks with $\kappa = 1$ are connected but easily disconnected by removal of single nodes, and thus only weakly structurally cohesive (Moody and White 2000). They are not really cohesive if we mean by that term "difficult to break apart" and "held together by coordinate bonds" rather than by adherence to single connecting nodes. Thus a group in which everyone was connected only to a single leader would lack cohesivity.

[18] To clarify the subtle difference between the graph theoretic and our sociological vocabulary once again, the *blocks* of order $n \geq 3$ are 2-component *cohesive blocks* (which may contain higher order *k*-components) while *blocks* of order $n = 2$ (single edges that are not contained in cycles) are contained within 1-component blocks that lack cohesivity. The *blocks* of a 1-component consist either of dyads not contained in cycles, or of proper subgraphs of the 1-component that have cohesivity (connectivity 2 or more). A 2-component is both a *block* and a *cohesive block*.

[19] Another characterization of a tree is that it is connected and that each edge is a bridge. A connected graph has $\kappa' = 1$ if and only if it has a bridge. Connected graphs 6, 7, and 8 in Figure 5 have bridges.

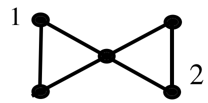


FIGURE 6. The bow tie graph. The lowest order of graph ($n = 5$) at which edge and node connectivities ($\kappa = 1$, $\kappa' = 2$) differ.

(G) of G is the smallest number of edges in an edge-cutset. Thus a disconnected graph has $\kappa' = 0$. Our definition 2.1.2 (Section I.A) corresponds to that of edge connectivity.

Edge connectivity does not differ from node connectivity for the graphs in Figure 5, where both types of connectivity are equal: zero for the first five graphs; one for the next three; two for the ninth and tenth, and three for K_4. Only at order 5 do node and edge connectivity of graphs begin to diverge, as exemplified in Figure 6, where the edge connectivity is 2 but the connectivity is 1.[20]

An *adhesive block* of a graph G is a k-edge-component of G where the associated value of edge connectivity defines the adhesion of the block. Figure 6 contains a single adhesive block with $\kappa' = 2$, but two overlapping cohesive blocks with $\kappa = 2$. Adhesive blocks may also overlap: An adhesive block with edge connectivity κ' may have at most $\kappa' - 1$ edges in common with a second block of equal or higher connectivity. Borgatti et al. (1990) define LS and lambda sets based on edge connectivity but restrict them to mutually exclusive subsets of adhesive blocks. This does not assure that such sets are cohesive, however, as they may contain cutnodes.

Two paths are *edge-independent* if they have no edges in common. Nodes 1 and 2 in Figure 6 are joined by just one node-independent path but two edge-independent paths. Graph G_1 in Figure 7 also illustrates the difference between node and edge connectivity: Nodes 1 and 2 are joined by two node-independent and three edge-independent paths, and five nodes (all but nodes 1, 2, and 3) are separable by either two edges or two nodes. Node and edge connectivity both equal 2 for the total graph (adhesive and cohesive blocking of $\kappa' = \kappa = 2$), but there is a greater surplus of path adhesion between nodes 1 and 2 than path cohesion.

[20] Node and edge connectivity are equal for any graph in which the minimum degree $\delta(G)$ is sufficiently large so that $\delta(G) \geq n/2$ (Harary 1969:44).

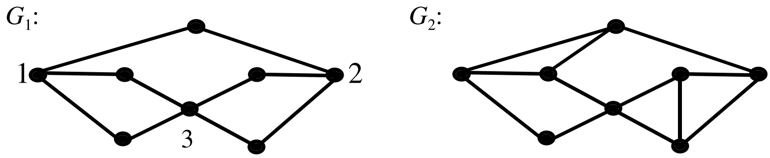


FIGURE 7. Two graphs with the same cohesion but different levels of adhesion.

C. Egocentric (Degree), Dyadic (Size), and Density Criteria as Partial but Insufficient Indicators of Cohesion, and Their Relation to Node and Edge Connectivity

The *degree* of a node u, denoted deg u, is the number of nodes to which u is adjacent. The *minimum degree* $\delta(G)$ is the smallest degree of a node in G. In an attempt to define cohesive subsets, Seidman (1983) defines a k-*core* of graph G as a maximal subgraph with $\delta \geq k$. Doreian and Woodward (1994) prove that k-cores form hierarchical series—i.e., for $k' > k$, a k'-core is a subgraph (possibly empty) of a k-core. The k-core, however, is no guarantee of cohesivity. The graphs in Figure 1 have $\delta = 3$ and thus are 3-cores, but are minimally cohesive because each has a cutnode (connectivity $\kappa = 1$) and lack cohesivity ($\kappa \geq 2$), (connectivity 0). For larger k, the same observation holds: There is no necessary concomitant increase in cohesion. The bow tie graph in Figure 6 is a similar example, with $\delta = 2$—hence a 2-core, lacking cohesivity.

The *density* $\rho(G)$ is the ratio of m edges of a graph G of order n and the number m_1 of edges of the complete graph K_n. As $m_1 = m(K_n) = n(n-1)/2$ we have

$$\rho(G) = m/m_1 = 2\text{m}/\text{n}(\text{n}-1). \tag{1}$$

Increases in size (number of edges, implying increased density) for a fixed n do not necessarily increase connectivity, and connectivity can vary independently of them. For example, the graph in Figure 3 has $n = 7$, $m = 8$, and connectivity 0. Other graphs with the same order and size have connectivity 1 or 2. There are, however, some dependencies between connectivity, degree, size, and density. We will make use of these dependencies later, in defining conditional density.

Whitney's Theorem (1932; cf. Harary 1969:43) states the inclusion relations between connectivity $\kappa(G)$ at the stronger end of a scale

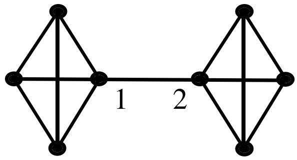


FIGURE 8. Graph with $\delta(G) = 3$ (a 3-core and 5-plex) that lacks both 3-connectivity and 3-edge connectivity. It is not even 2-connected, and thus lacks cohesivity ($\kappa \geq 2$), nor is it 2-edge-connected ($\kappa = \kappa' = 1$).

of cohesiveness, edge connectivity $\kappa'(G)$ at the middle, and minimum degree $\delta(G)$ at the weaker end (more inclusive, but less or at best equally cohesive):

Whitney's Theorem: For any graph G, $\kappa(G) \leq \kappa'(G) \leq \delta(G)$. (2)

From Whitney's Theorem it follows that every k-component is nested in a k-edge-component that is contained in a k-core, but not conversely. It follows that $\delta/(n-1)$, $\kappa'/(n-1)$, and $\kappa/(n-1)$ are the minimum densities, respectively, given δ and the connectivities κ' and κ of a graph G.[21]

Seidman and Foster (1978) attempted another measure of cohesion that is inadequate for similar reasons as the k-core. A *k-plex* is a maximal subgraph of order n where every node has degree $n - k$ or greater. Not every k-plex is an $(n - k)$-component. Figure 4 contains a 3-plex of order 5 (hence $n - k = 2$) that lacks cohesivity ($\kappa \geq 2$) because it contains a cutnode. For increases in $k \geq 2$, a k-plex may still have a cutnode or even be disconnected and thus there is no necessary concomitant increase in cohesion. In general, k-plexes and k-cores do not entail either respective $n - k$ or k node or edge connectivity. Figure 8 shows a graph of order 8 with a bridge between nodes 1 and 2. This graph is a 3-core and 5-plex that lacks both connectivity 3 and edge connectivity 3. Because of the bridge, the graph has edge (and node) connectivity 1. A connected graph

[21] In 1736, Euler proved the first theorem in graph theory, that the sum of the degrees of the nodes of any graph G is $2m$, twice the size of the graph (see Harary 1969:14). Letting $\underline{d}$ denote the average degree, this shows that $\underline{d} = 2m/n$. Since the smallest degree cannot be bigger than the average degree—i.e., $\delta \leq \underline{d}$, we have $\delta \leq 2m/n$ so $2m \geq n\delta$. Given the ordering of values $\delta \geq \kappa' \geq \kappa$ of G, Whitney's Theorem implies that $2m \geq n\delta \geq n\kappa' \geq n\kappa$. Recall that the density of a graph G is $\rho(G) = 2m/n(n-1)$. It follows that the minimum density $\rho(G)$ of a graph with minimum degree δ, substituting the inequality $m \geq n\delta/2$, and canceling, is $\delta/(n-1)$. Similarly, it follows that $\kappa'/(n-1)$ and $\kappa/(n-1)$ are the minimum densities, respectively, given connectivities κ' and κ of a graph G.

with no bridges (e.g., graphs 9–11 in Figure 5) has κ' at least 2. We shall not refer to a k-core or a k-plex further, as neither lend themselves to useful theorems or measures relating to cohesion in groups.[22]

D. Connectivity and Multiple Independent Paths as Cohesion, Menger's First Equality

Some of the deepest theorems in graph theory concern the equivalence between structural properties of graphs, such as connectivity (based on cutnodes), and how graphs are traversed. A graph G is *k-connected* if its connectivity is at least k. It is *k-edge-connected* when its edge connectivity is at least k. Karl Menger (1927) proved the equality of k-connectedness and the minimum number of node-independent paths between every pair of nodes,[23] which is a property of how a graph can be traversed. Consider graph G_1 in Figure 7: There is no pair of nodes with fewer than two node-independent paths, such as join nodes 1 and 2. The graph also cannot be disconnected by removal of fewer than 2 nodes. The proof of Menger's Theorem is found in Harary (1969:47), but its relevance to connectivity as a measure of cohesion (Wasserman and Faust 1994:115–17; Scott 2000:100–20) does not seem to be recognized in current sociological literature.[24] Menger's formulation and characterization of k-connected graphs, given below, is one of the most useful results in all of graph theory in that it establishes an equivalence between a structural and a traversal property of graphs, properties that happen to be the two most salient attributes of cohesion. Hence the structural cohesion in a group (definition 1.2.1) is equivalent to the path cohesion of the group (1.2.2).[25]

[22] The unions of intersecting cliques are another attempt to define cohesive sets (Freeman 1996), but the unions of cliques that have only one node in common also lack cohesivity.

[23] He accomplished this as an abstract result in the study of point-set topology. Note that by definition a graph of connectivity k is k-connected, but a k-connected graph may have connectivity k or greater. Likewise for edge connectivity k and k-edge-connected.

[24] Scott (2000:13) notes that "Harary developed powerful models of group cohesion" but does not develop what these ideas were in his chapter on cohesion.

[25] Alba and Kadushin (1976) define the cohesion of two nodes as the number of cycles in which they are contained. Since two cycles may differ but have edges in common, k cycles containing two nodes do not imply $k - 1$ (node-) disjoint paths between them, so this measure of cohesion does not identify clear boundaries of cohesive subsets. As noted, Harary, Norman, and Cartwright (1965) were the first to propose the connectivity of a graph (for the digraph case) as the primary measure of cohesiveness.

The *local connectivity* of two nonadjacent nodes u,v of a graph G is written $\kappa(u,v)$ and is defined as the minimum number of nodes needed to disconnect u and v. When u and v are adjacent they cannot be separated by removal of any number of nodes. Therefore local connectivity $\kappa(u,v)$ is not defined when u and v are adjacent. A complete graph, in which every pair of nodes is adjacent, does not have any local connectivities, and its (global) connectivity is defined as $n-1$, which corresponds to the number of node-independent paths that join each pair of nodes. But when G is not complete, the connectivity $\kappa(G)$ is the minimum value of the local connectivity taken over all nonadjacent pairs of nodes. Local and global edge connectivities are similarly defined, but with no exception for adjacent nodes.

Local Menger's Theorem A: The minimum node cut set $\kappa(u,v)$ separating a nonadjacent u,v pair of nodes equals the maximum number of node-independent u-v paths.[26]
Global Menger's Theorem A: A graph is k-connected if and only if any pair of nodes u,v is joined by at least k node-independent u-v paths.

Hence, for sociology, Menger's Theorem states the equivalence of our two parallel series of definitions of cohesion: connectivity (structural cohesion) and number of independent paths (path cohesion), which can be combined into a single concept of cohesion. Algorithms for computing numbers of node-independent paths between all pairs of nodes are given in White and Newman (2001). We now begin to expand on these two aspects of cohesion and reach a fuller appreciation of their sociological interpretation and implications.

E. (Edge-) Flow Connectivity and Node-Flow Connectivity

A *multigraph M* is obtained from a graph G when some of the edges are converted to two or more edges. An (integer valued) *network* is obtained from a graph G by assigning natural numbers, called *weights, values,* or *capacities,* to the edges of G. Therefore, when each edge with value t in a network is replaced by t edges joining u and v, we have a multigraph. Multigraphs are especially useful when the number of edges between two nodes (or corresponding values of the weighted graph or network) repre-

[26] The local and global theorems of Menger are examples of minimax theorems in mathematics.

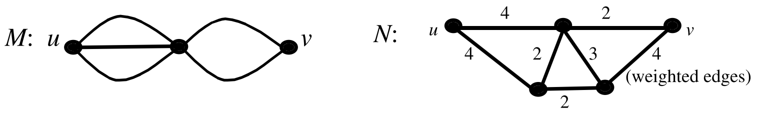


FIGURE 9. Multigraph M illustrating node-flow connectivity $\kappa''(u,v) = \kappa''(M) = 2$, where the connectivity is $\kappa(u,v) = \kappa(M) = 1$, and a network N (weighted graph) where node-flow connectivity $\kappa''(u,v) = 4$ but $\kappa'(u,v) = 6$.

sent flow capacities such as an ordinal limit on how much of some item can be transmitted from one node to another. The flow along a u-v path in which there are multiple edges in a multigraph is the minimum number of multiple edges joining adjacent pairs of nodes in the path. More generally, in a multigraph M, the flow from u to v is the number of edge-independent single-edge u-v paths. In the first graph of Figure 9, for example, there are two edge-independent single-edge u-v paths.

Extrapolating from Menger's Theorem, the local (*edge*-) *flow* from u to v in a multigraph M equals $\kappa'(u,v)$, the minimum number of edges in a cutset that separates u and v. This result illustrates how the concept of adhesion can be extended to graphs with weighted edges to capture the idea of differential strengths or capacities of edges. The (edge-) flow measure is widely used in the study of capacitated networks and flows (Ford and Fulkerson 1956), and will be discussed shortly.

We now consider only node-independent u-v paths: The local u-v *node-flow* (i.e., node-independent flow) is the maximum sum for a set of u-v flows for node-independent u-v paths. For graphs, this is simply the number of node-independent paths and by Local Theorem A is equal to $\kappa(u,v)$, the minimum number of nodes that must be deleted to disconnect u and v. To capture the idea of node-flow for a multigraph (equivalently, for a network), we expand the concept of connectivity to consider node-independent flow through multiple edges. The local *node-flow connectivity*, $\kappa''(u,v)$, in a multigraph M is the smallest number of edges in a set of node-independent u-v paths whose removal disconnects u and v.[27] The (global) *node-flow connectivity*, $\kappa''(M)$, is the smallest number of edges in a set of node-independent paths connecting any pair of nodes, whose removal disconnects M. This is not edge connectivity, because we consider only at the edges on node-independent paths. Figure 9 shows a mul-

[27] Two or more edges are *parallel* in a multigraph M if they join the same two nodes u,v. To disconnect a connected multigraph, one or more sets of parallel edges must be removed.

tigraph M where $\kappa(u,v) = \kappa(M) = 1$ but $\kappa''(u,v) = \kappa''(M) = 2$. The networks M and N (with weighted edges) in Figure 9 show $\kappa''(u,v) = 2$ and $\kappa''(u,v) = 4$, respectively.

By restatement of Menger's Local and Global Theorems,[28] we derive new corollaries of the celebrated Ford-Fulkerson Theorem (1956):

Local Ford-Fulkerson (Node-Flow Edge-Cut) Corollary. The *u-v* node-flow in a multigraph M equals the minimum number of edges in a cutset $\kappa''(u,v)$, within a set of node-independent paths, that separates u and v.
Global Ford-Fulkerson (Node-Flow Edge-Cut) Corollary. The minimum of the (maximum) *u-v* node-flows for all u,v pairs in a multigraph M equals $\kappa''(M)$, the minimum number of edges in a cutset, within a set of node-independent paths, whose removal disconnects M.

Because removing a node v of a connected G removes deg v edges and deg v is 1 or more, we also obtain the equivalent of Whitney's Theorem:

$$\kappa''(u,v) \leq \kappa'(u,v) \leq \min(\deg(u), \deg(v)). \quad (3)$$

Network N of Figure 9 (a graph with weighted edges), for example, shows flow or edge connectivity $\kappa'(u,v) = 6$, compared with node-flow connectivity $\kappa''(u,v) = 4$.

The concept of node-flow, as defined for the first time here, is not a single minimum value over a graph or multigraph but defines instead a matrix of values between each pair of nodes. Hence it provides a more detailed account of how cohesion is distributed in a group or network.[29] Flows through multiple node-independent paths are especially important in considering influences or effects as they spread through a network, and in compensating for distance decay. Higher redundancy—i.e., node independence—in flow may compensate for transmission decay at larger distances, and blocks of actors connected by node-independent flows may act as amplification systems for boosting the coherent signals transmitted in social interactions.

[28] Insofar as we know, the definition of node-flow is a new concept, and the restatements are new, but its proof is obvious from Menger's Theorem and the definition of node-flow.

[29] Scaling techniques applied to node-flow matrices should give a more detailed analysis of differential cohesion in a group, but node-flow has been more difficult to compute for large networks than connectivity, hence White and Newman's (2001) results will provide a fruitful avenue of research.

F. Edge Connectivity and Edge-Independent Paths as Adhesion, Menger's Second Equality

Like cohesion, we defined social adhesion both in terms of resistance to disconnection through edge removal (2.1.2—structural adhesion) and of multiple paths (2.2.2—path adhesion). These two definitions are also unified as equivalents by the edge version of Menger's Global and Local Theorems:

Local Menger's Theorem B: The minimum number of edges in a cutset separating u and v equals the maximum number of edge-independent paths that join u and v.
Global Menger's Theorem B: A graph G is k-edge-connected if and only if every pair of nodes u,v in G are joined by at least k edge-independent u-v paths.

When the edges of graphs are weighted, or we convert a weighted graph to an equivalent multigraph, we can make use of Ford and Fulkerson's (1956) maximum flow–minimum cut theorem, one of the most widely used results in all of operations research:[30]

Local Ford-Fulkerson (Edge-Cut) Theorem. The maximum flow u-v in a multigraph M equals the minimum edge-cut $\kappa'(u,v)$ that separates u and v.
Global Ford and Fulkerson (Edge-Cut) Theorem. The minimum flow between any pair of nodes in a multigraph M (equivalently, a network or integer-weighted graph) equals the minimum number of edges $\kappa'(M)$ whose removal disconnects M.

It is this theorem that Zachary (1975, 1977) uses to partition his karate club network, whose leaders are in conflict, into two halves: those who are most adhesively connected (by the greatest number of edge-independent paths = the least-edge cutset) with the club's administrator and those most adhesively connected to the club's karate instructor. We discussed this case in Section I.C, in which our critique of Zachary's adhesion-based minimum

[30] Dirac (1960) showed that this result, in which each edge e has a numerical weight $w(e)$, is a straightforward corollary of Menger's Theorem B. Several variations on Theorems A and B are presented in Harary (1969, ch. 5).

edge-cut method began with its weakness as a sociological basis for identifying the relational component of solidarity as compared with a cohesion-based minimum removal of nodes. Our critique noted that it is easier to assert mechanisms whereby social pressures operate on individuals—due to their agency—to take sides in disputes, than to understand why group segmentation should occur along a minimum edge-cut.[31] In Section IV, following Section III where we develop the measure of conditional density, we exemplify and compare the two methods (adhesive edge connectivity and cohesive node connectivity) for the Zachary karate club study.

G. The Hierarchical Properties of Cohesive and Adhesive Blocks

At the first level of analysis of structural and path cohesion and adhesion, which reduce to the concepts of node and edge connectivities in graphs, we can complete the series of intuitive assertions of Section I as follows:

1.3. Relationally cohesive groups can be regarded as multiply nested in terms of connectivity values in the following sense: a connected graph can contain several 2-components, each of which can contain 3-components, and so forth. Likewise for a multigraph or graph in which edges are weighted: cohesive groups are multiply nested in terms of their node-flow-connectivity values.

2.3. Relationally adhesive groups are multiply nested in terms of edge connectivity values in the sense that a connected graph can contain several 2-edge-components, each of which contain 3-edge-components, etc. Likewise for a multigraph or weighted graph: adhesive groups are multiply nested in terms of their flow-connectivity values.

Graphs and social groups at the same levels of connectivity can be further ordered at a second level, according to conditional density, taken up in the next section.

[31] Ties have no autonomous agency, so that if a social network has a certain edge-flow between two conflicting individuals (and there may be many edge-flow equivalent edge-cuts that will partition the group into disconnected factions), how would the "network" know to partition along an edge-cut? Alternatively, where a cutnode or cutset of individuals exists whose removal would disconnect the graph, social pressure to choose sides is likely to fall on this individual or set of individuals, who have the agency to alter their ties according to principles of cohesion, density, and closeness of ties to the leaders and their cohesively closer faction members, agency which is lacking to the network as a whole.

III. CONDITIONAL DENSITY

The idea of conditional density is that if some property of a graph is held constant—such as connectivity—then density may vary only within a limited range and can be rescaled from zero to some maximum within that range. Since connectivity is an integer number, a minimum value for a group, and density above the minimum required for connectivity at that level makes an additional contribution to cohesion, a rescaling of density as a fractional number allows us to add the two together to get an aggregate measure of cohesion. By adding conditional density in such a measure, we can account for additional cohesion that connectivity cannot capture alone.

To define precisely the conditional density of a graph G with respect to some structural property, we need some preliminary definitions. Let P be a generic property of graphs, such as connected, or bipartite, etc. We always denote the order of G by n (nodes) and its size by m (edges). Let $m_0(G{:}P)$ be the minimum size of a graph G of order n that has property P, and let $m_1(G{:}P)$ be the maximum size. Then the *conditional P-density*, $\rho(G{:}P)$, is defined by

$$\rho(G{:}P) = (m - m_0)/(m_1 - m_0).$$

If m_1 is the maximum m in a graph of order n with a given κ, and m_2 is the smallest m that forces that graph to surpass property P of connectivity κ, then $m_2 = m_1 + 1$ is the *upper size limit* on the number of edges at which a less than complete graph G of order n cannot retain property P, but below which G has property P. The *conditional P-density*, $\rho_2(G{:}P)$, of an (n, m) graph, which is always less than one, is defined within the lower and upper size limits m_0 and m_2:[32]

$$\rho_2\,(G{:}P) = (m - m_0)/(m_2 - m_0) < 1. \qquad (4)$$

If P is omitted from either of these formulas, so that $m_0 = 0$ and we let $m_1 = m_2 = m(K_n) = n(n-1)/2$, the size of a complete graph K_n, then ρ_2 and ρ reduce to the usual graph density formula:

$$\rho(G) = \rho_2(G) = m/m_1 = m/(n(n-1)/2) = 2m/n(n-1).$$

[32] The notation for m_0 and m_1 designates that density normally varies between 0 and 1; conditional P-density ρ_2 approaches but never reaches 1 unless $m_1 = m(Kn)$. This latter characteristic will be useful when we define cohesion as an aggregate measure consisting of the sum of connectivity κ plus conditional κ-density $\rho_2(G{:}\kappa)$.

For graph 8 in Figure 5, for example, the ordinary density $\rho(G) = .67$, and where the property P is that of connectivity ($\kappa = 1$), graph 8 has a surplus of one edge beyond those needed for $\kappa = 1$, while the maximum such surplus is two edges for a connected graph with 4 nodes, hence the conditional P-density of this graph is $\rho_2\ (G{:}\kappa = 1) = 0.5$. Three surplus edges beyond those required for $\kappa = 1$ are needed to force a graph with 4 nodes to have $\kappa = 2$, as in graph 10. In general, conditional P-density ρ_2 is the ratio of surplus edges, beyond those minimally needed for a graph of order n to have property P, to the upper size limit m_2 at which a graph with order *n cannot* still retain property P.[33]

A. Connectivity and Conditional Density: *A Unified Approach to Measuring Cohesion*

Graph connectivity and density are two aspects of cohesion that are tightly bound together. We take advantage of their interdependence to combine and unify them into a single measure of cohesion.

To apply conditional density to the property of connectivity requires the values of $m_0(G{:}\kappa)$ and $m_1(G{:}\kappa)$ or $m_2(G{:}\kappa)$ for a graph G of size m and order n with connectivity κ.[34] These are known from extremal graph theory.[35] Let $\lceil x \rceil$ be the fraction x rounded up to the nearest integer, and for conciseness, let $m^* = m(K_{n-1})$, the size of the complete graph of order $n-1$. The limiting size $m_2(G{:}\kappa)$ of a graph of order n with connectivity κ, where $0 \leq \kappa < n-1$, is $1 + \kappa + m^*$. For $\kappa = n-1$ we define $m_2(G{:}\kappa) = n(n-1)/2$, the maximum size of a graph, giving maximum conditional density of 1 only for K_n. The minimum numbers of edges m_0 of G with connectivity $\kappa = 0$, 1, and >1 are 0, $n-1$, and $\lceil n\kappa/2 \rceil$, respectively. In general, $m_0 = \lceil \mathrm{n}\kappa/2 \rceil$ rises linearly with n, while m_1 and m_2 rise quadrat-

[33] Conditional P-densities ρ and ρ_2 differ in that the denominator of the former limits density to an interval [0,1] relative to the number of edges at which property P cannot be retained. Conditional P-density, $\rho(G{:}P)$, is the number of surplus edges divided by the maximum number of surplus edges at which a graph with order n can still retain property P.

[34] See also Harary (1983) on conditional connectivity, and Harary and Cartwright (1961) on the number of arcs in each connectedness category of a digraph.

[35] See Harary (1969:17–19) for an introduction to extremal graphs. The result $(m{:}n,\ \kappa > 1) = \lceil n\kappa/2 \rceil$ agrees with a minimum density of $\kappa/(n-1)$ for a graph of connectivity $\kappa > 1$. A problem opposite to that of conditional density, covered in extremal graph theory, is *conditional connectivity*: What are the minimum and maximum connectivities for a graph of order n and size m? A 4-node graph with 5 edges, for example, must be 2-connected.

ically, so that conditional densities are more tightly limited when there are fewer nodes and higher values of κ.

When $\rho_2(G{:}\kappa)$ is close to zero, the connectivity structure is fragile, in that the removal of a randomly chosen edge is likely to reduce the connectivity of G. The minimum size of a graph G of order n for $\kappa = 2$, for example, is n, realized only by the cycle C_n. The removal of any one of these n edges reduces the connectivity to $\kappa = 1$. If G contains one surplus edge, the chance that random removal of an edge will reduce the connectivity of G to 1 is $n/(n+1)$. As the surplus density $\rho_2(G{:}P = \kappa)$ increases, more nodes will have extra edges, and the graph becomes less vulnerable to a lowered connectivity with the removal of a random edge. As conditional density approaches 1, the connectivity structure is more robust: Many randomly chosen edges can be removed with less chance of reducing connectivity.

B. Cohesion: A Scalable Aggregate Measure (Connectivity κ Plus ρ_2 Density)

The sum of the connectivity κ and the conditional density $\rho_2(G{:}\kappa)$ of a graph G is not the only possible measure of its cohesion but is a better measure than any of the other cohesive subset algorithms discussed in Sections II.C and I.D (k-plex, k-core, and intersecting cliques) because none of the higher values on these measures is a guarantee of our minimum criterion of cohesivity ($\kappa \geq 2$; for any value of k, a k-core or k-plex may even have $\kappa = 0$). Connectivity and conditional density each contribute independently to cohesion, according to the two criteria reviewed earlier: the structural cohesion integer and the conditional density fraction. We now consider how density plays into the criteria for cohesion.

For the first criterion of cohesion—namely, that a cohesive block stays together—the value k of $\kappa(G)$ is the guarantee that a graph G cannot be disconnected without removal of at least k nodes. In addition, higher values of conditional density reduce the likelihood that removal of a random edge will diminish the value of κ.

For the second criterion of cohesion, that the nodes of a cohesive block should be strongly tied, the value k of κ is also the guarantee, by Menger's Theorem A, that every pair of nodes in a graph with $\kappa(G) = k$ has k or more independent paths connecting them. In addition, the higher the value of conditional density, $\rho_2(G{:}\kappa)$, the less the likelihood that the removal of a random edge will diminish the minimum number k of independent paths that join every pair of nodes.

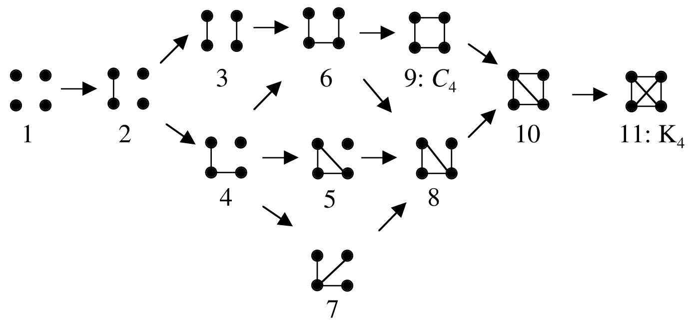


FIGURE 10. The 11 graphs of order 4 showing transitions by graph evolution (addition of edges).

The probabilistic framework for considering the contributions of conditional density to cohesion—invoking the *expected* impact of edge removal as a supplement to the *greatest* impact as measured by connectivity (structural cohesion)—provides one of the major advantages of this approach over alternatives such as conditional distance or conditional adhesion.[36]

C. A Well-Constructed Measure of Cohesion

A well-constructed measure requires a demonstration of the *unit* of measurement that gives a monotone increase in magnitude of the quantity measured. The unit for which the aggregate measure of cohesion is monotone increasing is the addition of an edge within a connected graph of order n. Hence an essential criterion for measurement is satisfied. The aggregate cohesion measure is monotonically increasing in any sequence of graphs of a given order (n) in which edges are successively added. The sequences of graphs of order 4 shown in Figure 10, for example, satisfy this criterion. These are the same 11 graphs as in Figure 5, along with directed arrows showing which graphs are transformed to another by addition of

[36] Future developments in the theory of random graphs (Palmer 1985; Kolchin 1999) might also help to establish whether the distribution of distances between nodes in a graph with connectivity κ and conditional density $\rho_2(G{:}\kappa)$ is less than, equal to, or greater than expected in a "random" distribution. Approximation methods have not yet been developed to answer these questions within random graph theory, although a simulation approach might be useful to bring the emergence of small world cohesivity into the kind of simulation framework developed by Watts (1999a).

an edge; their aggregate cohesion measures are given in the fourth row of Table 1: For each successive graph under edge addition, there is an increase in aggregate cohesion.

Only certain sequences of adding edges will give maximum increases in aggregate cohesion. Cohesion is increased maximally in the sequences in Figure 10, for example, that lead from the totally disconnected graph (1) to the graph with all nodes connected by a single path (6), to one with all nodes connected by a single cycle (9), to the graph consisting of a single clique (11). The principle of attaining successive maximum cohesion by adding edges in a graph of any order n is always to build first a single connecting path, then a cycle (C_n of connectivity 2), then to place in a distributed way the minimum edges needed to build connectivity 3, similarly for connectivity 4, and so forth.

Adding edges so as to build cliques of maximum order as a subgraph, in contrast, does not maximize increase in overall connectivity or aggregate cohesion, although it does increase subgroup density and heterogeneity. Nor does adding an edge to a graph increase cohesion if one or two new nodes are also added incident to the edge. Unless cohesion was initially zero, this will decrease the cohesion of the graph.

Table 2 shows in columns 2 and 3, for the various values of n and κ given in column 1, the values m_0 and m_2 for graphs of order $n > 1$ with

TABLE 1
Connectivity κ, Conditional Density $\rho_2(G{:}\kappa)$, and Aggregate Cohesion $\kappa + \rho_2(G{:}\kappa)$, for the 11 Graphs in Figure 5

Graph G	1	2	3	4	5	6	7	8	9	10	11
Size (edges)	0	1	2	2	3	3	3	4	4	5	6
κ	0	0	0	0	0	1	1	1	2	2	3
$\rho_2(G{:}\kappa)$	0.0	.25	0.5	0.5	.75	0.0	0.0	.5	0.0	0.5	0.0
$\kappa + \rho_2(G{:}\kappa)$	0.0	.25	0.5	0.5	.75	1.0	1.0	1.5	2.0	2.5	3.0
Component S		a, c	a, c	a, c, d		G	G	G	G	G	G
κ		1	1	1		1	1	1			
$\rho_2(S{:}\kappa)$		0.0	0.0	0.0		0.0	0.0	0.5			
$\kappa + \rho_2(S{:}\kappa)$		1.00	1.00	1.00		1.0	1.0	1.5			
Bicomponent S					a, c, d			a, c, d	G	G	G
κ					2			2	2	2	
$\rho_2(S{:}\kappa)$					0.0			0.0	0.0	0.5	
$\kappa + \rho_2(S{:}\kappa)$					2.0			2.0	2.0	2.5	
Tricomponent S											G
κ											3
$\rho_2(S{:}\kappa)$											0.0
$\kappa + \rho_2(S{:}\kappa)$											3.0

TABLE 2

Ranges of m_0 and m_2 for Computing Conditional Density $\rho_2(G{:}\kappa)$ at Connectivity κ

Graph Sizes Given: Order n, Connectivity κ	At κ min $m_0 = [n\kappa/2]^+$	Forcing $\kappa + 1$ $m_2 = 1 + \kappa + m^*$ for $\kappa < n - 1$
$n > 1$, $\kappa = 0$ Fig. 3	0	$1 + \kappa + m^*$
$n > 1$, $\kappa = 1$ Figs.1,2,4,6,8 and 5.8	$n - 1$	$1 + \kappa + m^*$
$n = 4$, $\kappa = 2$ Fig. 5.9	4	6
$\kappa = 3$ Fig. 5.11	6	6
$n = 5$, $\kappa = 2$	5	9
$\kappa = 3$	8	10
$\kappa = 4$	10	10
$n = 6$, $\kappa = 2$	6	13
$\kappa = 3$	9	14
$\kappa = 4$	12	15
$\kappa = 5$	15	15
$n = 7$, $\kappa = 2$ Fig. 7	7	18
$\kappa = 3$	11	19
$\kappa = 4$	14	20
$\kappa = 5$	18	21
$\kappa = 6$	21	21
$n = 40$, $\kappa = 2$	40	744
$\kappa = 3$	60	745
$\kappa = 4$	80	756
$\kappa = 5$	100	757
$\kappa = 6$	120	758
$\kappa = 39$	780	780

$m^* = m(K_{n-1})$, the size of the complete graph of order $n - 1$.

$\kappa = 0$ or 1, and for graphs with $4 \leq n \leq 7$ and $n = 40$ nodes for various values of $\kappa > 1$. Illustrative graphs from Figures 1–8 are referenced in the table. The bow tie graph ($n = 5$, $\kappa = 1$) in Figure 6, for example, has a conditional density of (6-4)/(8-4) = 0.5 and an ordinary density of 0.6.

As noted above, the denominator of conditional density $\rho_2(G{:}\kappa)$ ensures that it cannot reach 1.0 for a given κ. This allows the aggregate measure of cohesion—as a connectivity integer plus a conditional density decimal (<1)—to correctly distinguish between the case of maximum density at connectivity κ and minimum density at connectivity $\kappa + 1$, where the aggregate cohesiveness of the former is always less than that of the latter.[37]

[37] This is not the case for the sum of connectivity and $\rho(G, \kappa)$ conditional density, which does not give a measure of cohesion because the sum $\kappa + \rho(G, \kappa)$ for a graph with connectivity κ and size $m_1(G, \kappa)$ is the same as the sum for a graph with connectivity $\kappa + 1$ and size $m_0(G, \kappa) = 0$.

D. Subgroup Cohesion

The boundaries and measures of each of the *k*-components of a graph provide a convenient way to study the structure of social cohesion. One of the problems in previous measures of social cohesion such as intersecting cliques, *k*-cores, and *k*-plexes—apart from the fact that they are no guarantee of cohesivity ($\kappa \geq 2$; or even that $\kappa > 0$)—is that there is so much overlap in the cohesive subsets they identify. The measure of cohesion based on the sum of the connectivity $\kappa(S)$ of a subgraph S and its conditional density $\rho_2(S{:}\kappa)$ typically yields more interpretable cohesive subgroups, with very little overlap, and a hierarchy of nested *k*-components, each with successively higher levels of cohesion.

To illustrate measurement of subgroup cohesion, relationships among the cohesion measures are shown in Table 1 for each *k*-component of the eleven 4-node graphs in Figure 5. Rows two to four of the table show the connectivity κ, conditional density $\rho_2(G{:}\kappa)$, and aggregate cohesion $\kappa + \rho_2(G{:}\kappa)$ for each of the graphs in their entirety. The second, third, and fourth sets of four rows each show these values for the largest component, bicomponent, and tricomponent, if any, of each of the 11 graphs.

E. Subgroup Inhomogeneities

Social groups with networks of high connectivity have high cohesion, but they may be highly inhomogeneous if they have high conditional densities as well. Groups with low conditional densities have relatively fewer surplus edges with which to create local subgroup inhomogeneities. Thus some of the problems in the study of nested subgroups, their relative homogeneities and inhomogeneities, and the relation between cohesion and social solidarity (Markovsky and Lawler 1994; Markovsky and Chaffee 1995; Markovsky 1998) can be studied by means of connectivity and conditional density.

IV. AN EMPIRICAL EXAMPLE: ZACHARY'S KARATE CLUB

Zachary's (1975, 1977) two-year ethnographic network study of 34 members of a karate club is a good proving ground to examine concepts, measures and hypotheses involving relational aspects of social solidarity. This section provides an illustration of how measures of cohesion are used to predict the outcome variable of sides taken in a factional dispute from the boundaries of nested cohesive sets. The disputants were the karate teacher (T, #1, Mr. Hi) and the club administrator (A, #34, John) and the dispute was about whether to improve the solvency of the club by raising fees (teacher) or by holding costs down (as A insisted). This resulted in each calling meetings at which they hoped to pass self-serving resolutions by encouraging attendance of their own supporters. The formation of factions was visible to the ethnographer and evident in meeting attendance, which varied in factional proportions according to the convener. Ultimately Mr. Hi (T) was fired, set up a separate club, and the factional split became the basis for each student's choice of which of the new clubs they would join. The prediction tested here is that when two "sociometric centers" of a group force a division into two, the cohesion measure will predict how members of the old group will distribute among the new ones.

A. Global View of the Karate Network

Figure 11 shows the network of friendships among the 34 members. Zachary weighted the strength of each friendship by the number of con-

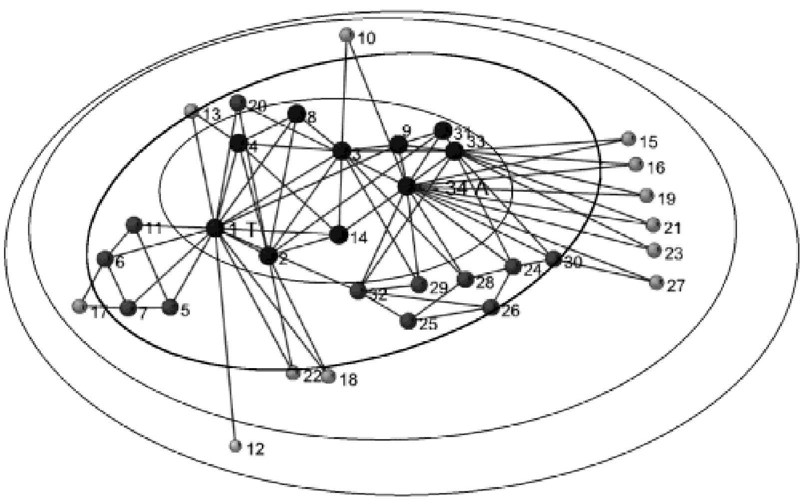


FIGURE 11. Nested adhesive sets for edge connectivities of 1, 2, 3, and 4. This and the following figures are drawn with Batagelj and Mrvar's (1997, 1998) Pajek software.

TABLE 3
Edge Connectivity Sets for the Karate Club

Sets	Members [1,34 leaders]	Nested in Set	n	κ'
1	1–34		34	1
2	1–11,13–34	1	33	2
3	1–9,11,14,20,24–26,28–34	2	22	3
4	1–4,8–9,14,31,33–34	3	10	4

texts (karate and other classes, hangouts, tournaments, and bars) in which the pair met, but the weights are not shown in the figure. Instead, concentric rings of adhesive subsets are circled in Figure 11 according to their 1-, 2-, 3-, and 4-edge connectivity. Table 3 lists the members and gives the number (n) in each set, the concentric nesting of the sets, and the edge connectivity κ' of each. The set with highest adhesion, which consists of 10 members and includes A and T, is separable by four edges ($\kappa' = 4$) but A and T are separated by a minimum of 10. There are, however, many different edge-cuts of size 10 that separate A and T. Hence unweighted edge-cuts, as well as adhesive sets, fail to predict faction membership. Each of the nested 1-, 2-, 3-, and 4-edge-connected subsets contains a cutnode (T) and lacks cohesivity since $\kappa < 2$.

Zachary used weighted minimum edge-cuts between A and T (the Local Ford-Fulkerson max flow–min cut theorem) to predict the separation of the two factions. Except for three persons who did not take sides, this gave a near-perfect prediction of the split. The particular distribution of weights on the edges, however, contributed to a unique-cut solution, pushed somewhat away from T since weights were highest for those close to him. Zachary did not utilize criteria for subgroup cohesion, but the dynamics of the dispute gives us the opportunity to examine cohesive blocks before the split and the role they played in mobilizing the taking of sides.

Looked at in terms of cohesion (Figure 12) the network has five cohesive blocks of connectivity 2 or greater, each enclosed in Figure 12 by one of the concentric circles. Only node T is common to them all. Two exclude node A (a 3-component within a 2-component) and three (a 4-component within a 3- within a 2-component) include node A. Table 4 shows the cohesive blocks 1–5 circled in Figure 12, their members, number of nodes, the hierarchical nesting of each block, its connectivity, conditional density, and aggregate cohesion. An additional

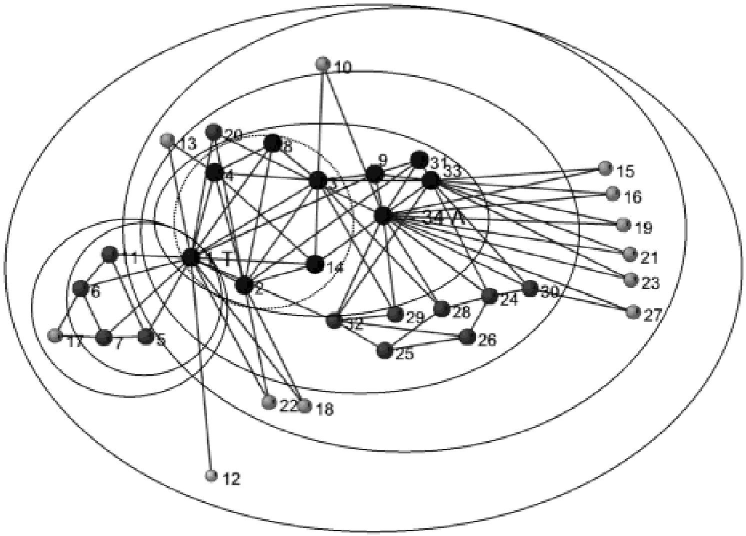


FIGURE 12. Cohesive blocks hierarchically ordered by connectivity into two nests (the outer dotted circle nests them all in a connected graph with connectivity 1).

subset with the highest aggregate cohesion of 4.75 is also shown—the six people within the dotted circle in Figure 12—which is not a maximal cohesive block but part of block 5 (with cohesion 4.24). This subset forms the core of support for Mr. Hi, while the remnants of block 5

TABLE 4
Connectivity Block and Subset Characteristics for Karate Club

Blocks and Sets	Members [1,34 leaders]	Nested in Set	n	κ	ρ_2	Aggregate Cohesion
1	1,5–7,11,17		6	2	.2	2.20
2	1,5–7,11	2	6	3	.54	3.54
3	1–4,8–10,13–16,18–34		28	2	.12	2.12
4	1–4,8–9,14,20,24–26,28–34	3	18	3	.12	3.12
5	1–4,8–9,14,31,33,34	4	10	4	.24	4.24
6*	1–4,8,14	5	6	4	.75	4.75
7*	9,31,33,34	5	4	3	.00	3.00

*Sets 1–5 are cohesive blocks; set 6 is the densest cohesive subblock within 5 and set 7 is the residual within 5 after taking out set 6.

after removing this subset, shown as set 7 in Table 4, are supporters of A.

A first and approximate prediction of factions uses the number of node-independent paths (node-flow) joining pairs of nodes, and then takes the maximum spanning tree of the edges in the original network selected in order of largest number of node-independent paths (White and Newman 2001: the spanning tree portion of the algorithm, in general, breaks ties in favor of pairs of nodes separated by least distance). The result is depicted in Figure 13, in which the vertical line is a good predictor of the initial factional alignment, with followers of T to the left and those of A to the right. Person 9, on A's side of the prediction line, initially aligns with A but later switches to T. After removal of the cutnode between T and A, which is also person 9 (which also removes the two dotted lines in the figure), two trees remain with T and A at their respective centers. Except for those who do not take sides (three nodes labeled with a question mark), and two others, 28 and 29, the trees predict the factional alignments. The spanning tree algorithm, however, introduces some noise to Figure 13 as

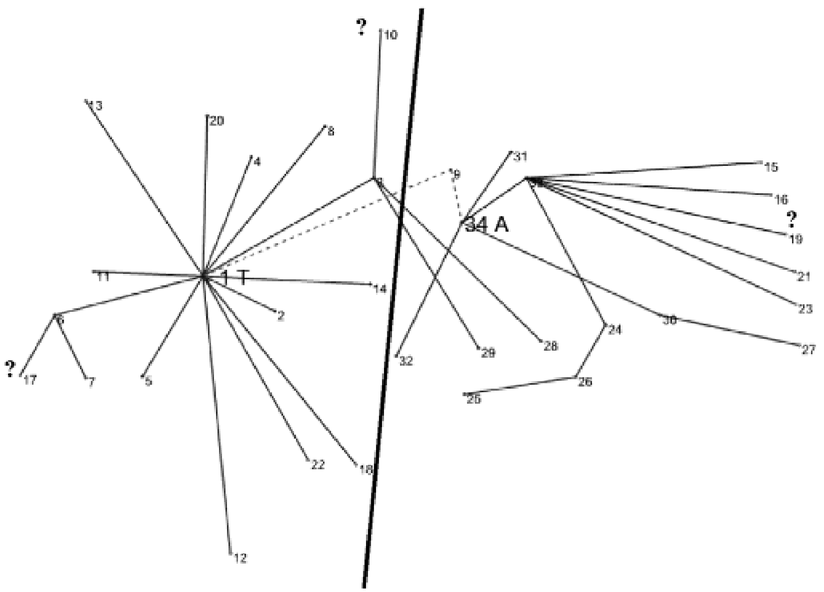


FIGURE 13. Maximum spanning tree of the numbers of node-independent paths between pairs of nodes, where solid lines predict faction members except for 28 and 29 and those nodes labeled with a question mark.

a predictor variable because choice is arbitrary among edges that are equally well qualified for the spanning tree. Persons 28 and 29 are an example, and could equally well be linked by the algorithm to A, thereby improving the prediction.

B. Closeup of Cohesion in the Karate Network

Because cohesive blocking is a deterministic procedure, it makes more precise predictions than White and Newman's approximation algorithm. If we situate the problem of determining factional divisions in the context of the opposition between leaders, as Zachary did, there are four persons—9, 14, 20, and 32—who had friendships with both leaders and thus had to make up their minds which leader to follow as the club split. Their membership in cohesive blocks and subsets provides a determinate prediction as to their decisions about which leader to follow. The choices they made corresponded not to the number of contexts in which they had friendships with T (Mr. Hi) or A (John), as Zachary would suggest, but to the pull of cohesive ties with others in core group of T (set 6 in Table 4) versus A (set 7). For each of these four people, who must decide between T and A, Table 5 contains four labeled rows: In the three columns under Mr. Hi's faction are the subset size (n), number of edges (e), and aggregate cohesion ($\kappa + \rho_2$) within Mr. Hi's faction (set 6); and similarly for A's faction, set 7. In the center of the table is a column that shows whether cohesion is greater with T ($>$) or A ($<$). In the rightmost columns are each person's predicted and actual choice of faction, showing that each of these people chose to align with the faction in which they have highest cohesion.

TABLE 5
Aggregate Cohesion (AC) with Leadership Factions for Persons Tied to Both Leaders and Obliged to Choose Between Them

Member	Mr. Hi's Faction		AC_1	AC_1	A's Faction		AC_2	Predicted Choice	Actual Choice
	n	e	$\kappa + \rho_2$	AC_2	n	e	$\kappa + \rho_2$	of Faction	of Faction
14	6	14	4.75	$>$	5	7	1.75	Mr. Hi	Mr. Hi
20	4	5	2.5	$>$	5	7	1.75	Mr. Hi	Mr. Hi
9	2	1	2.0	$<$	4	6	4.0	A	A
32	2	1	2.0	$<$	5	8	2.83	A	A

Students 14 and 20, for example, had more cohesion with Mr. Hi's group than with A's, and they aligned with Mr. Hi's faction in attendance at meetings. Students 9 and 32, on the other hand, had more cohesion with A's group and aligned with his faction. Each of these four people had to make a choice to drop a tie with the leader whose faction they rejected. If we remove the line connecting 14 to 34 (A) because 14 chose to join T's faction, for example, we observe in Figure 14 that even this one edge-removal results in a smaller 4-connected cohesive block of six persons {1, 2, 3, 4, 8, 14}, out of the original 10 in block 5, nested within a larger 3-cohesive block. All six persons in this 4-block align with T, as predicted from cohesion. If we allocate the remaining nodes in the 3-block according to their cohesion with T versus A, person 20 is predicted to go with T and the remainder with A. Allocating those in the 2- and 1-components by the same procedure, only person 10's alignment is indeterminate, and 10 was one of the three not factionally aligned.

The results of this test of the global predictions of faction membership from our connectivity measures are summarized in Table 6 ($r = .969$, $p < .000001$). The columns indicate whether cohesion is greater, equal, or less with Mr. Hi (T) than with John (A). This correlation is nearly identical to Zachary's prediction using the Ford-Fulkerson maximum flow–

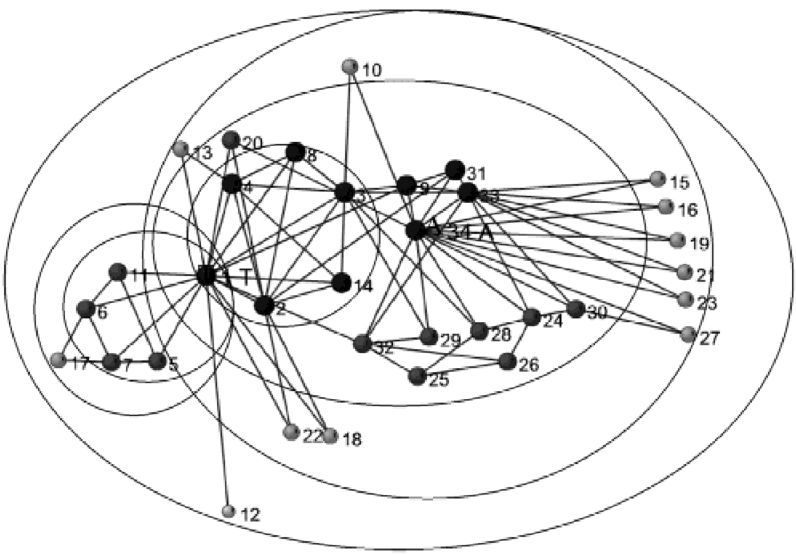


FIGURE 14. Nested cohesive sets by *k*-connectivity after person 14 affiliates with T.

TABLE 6
Predictions of Faction Membership from Structural Cohesion ($r = .969$)

Cohesion Faction	Mr. Hi (T)	Equal for T and A	John (A)	Members by ID number
Mr. Hi's (T)	15			1–8,11–14,18,20,22
None	1	1	1	17, 10, 19
John's (A)			16	9,15,16,21,23–34

minimum cut algorithm on weighted edges, as shown in Table 7 ($r = .955$). Zachary's prediction, however, contains an unwarranted postulate of some kind of network "agency" that "finds" an optimal edge-cut without any explanation as to mechanism.

C. Evaluation of Results

Alignment of factions in the karate club is predicted by structural and path cohesion as well as by Zachary's adhesive weighted edge-cuts. The results shows the contribution of conditional density, in addition to cohesion measured by connectivity, to identifying localized high-density subgroups within cohesive blocks. The high-density subgroup (set 6) that was the core of T's support group was a very compact group with minimum distances among members, which may have contributed to leader T's retention of so many followers in the breakup of the club. Zachary had the right result as well, but possibly for the wrong reasons.[38] Although both models make near-perfect predictions, the cohesion argument has advantages over the capacitated flow and possibly other arguments in the karate study on the grounds of parsimony (use of unweighted over weighted edges) and a process model of agency as the mechanism involved in segmentation.[39]

[38] Another reason that Zachary's capacitated flow argument does so well is that flow, in the unweighted case, is highly correlated with number of independent paths.

[39] The fact that Mr. Hi is the cutnode in a bifurcated network might help to explain—in sociological terms—why he is the instigator of the dispute in the first place: He has a set of at least five potential students who were never integrated into the larger cohesive block containing the administrator (#34), and for Mr. Hi it was clear from the beginning that they would follow his leadership. He was also a strong figure for many of his other adherents.

TABLE 7
Predictions of Faction Membership from Zachary's Minimum Weighted Edge-Cut ($r = .955$)

Edge-Cut Faction	Mr. Hi (T)	John (A)	Members by ID number
Mr. Hi's (T)	15		1–8,11–14,18,20,22
None	1	2	17, 10, 19
John's (A)		16	9,15,16,21,23–34

V. TESTING PREDICTIVENESS OF COHESION MEASURES ON A LARGER SCALE

The evidence of the karate data is useful in evaluating connectivity as a measure of ohesion. The aggregate cohesion measure makes correct predictions about the consequences of cohesion for individual behavior and the emergence or division of social groups, but it does so on a relatively small scale in which different measures and approaches to cohesion might give similar results. Everett (1996) analyzed the overlap of cliques in Zachary's data, and got similar results concerning cohesive subsets that predicted faction membership. Large high-connectivity groups will not in general, however, be constructed out of intersecting cliques. We do not presume that connectivity is the only measure of cohesion, only that it is a fundamental component of interpersonal cohesion in social groups, large or small.

In our computation of factional groups for the karate club, we used the measure of node-flow (number of node-independent paths) to compute pairwise cohesion (White and Newman 2001). Pairwise cohesion is computed separately for each pair of nodes in a group, whereas connectivity is the minimum of these values over all pairs. Pairwise flow, which measures the number of edge-independent paths, is always the same or greater than node-flow. Johansen and White (forthcoming) successfully used the maximum flow measure, equivalent to an unweighted version of Zachary's method, in predicting large-scale political factions in a nomadic society. The use of pairwise measures of cohesive strength (both flow and node-flow) may be widely useful in studying patterns of social cohesion and adhesion, but it is still an open research question as to which measures make better predictions, and why.

Moody and White (2000) characterize the respective graphs of socially cohesive friendship groups in 12 American high schools, and again of 57 financially cohesive business groups, by computing k-components. They find that embeddedness in k-connected groups is a strong predictor of school attachment and is the only predictor of attachment among diverse network variables that replicates significantly across all schools. Using data from a study of business unity by Mizruchi (1992), they also show that cohesion or level of connectivity applied to the network of corporate interlocks among 57 firms, controlling for other network measures, predicts similarity in business behaviors. They illustrate the fundamental importance of connectivity and its hierarchical embedding,[40] and argue for a wide range of applications in sociology. Moody and White's (2000) algorithm successively removes sets of nodes with the lowest connectivity. By combining several algorithms of low complexity, the algorithm makes connectivity computations feasible for relatively large graphs. Studies using connectivity to measure social cohesion, such as surmised but not actually employed by Grannis (1998), however, are still quite rare.[41]

Brudner and White (1997) and White et al. (2001), for example, identified sociologically important cohesive blocks in two large (n = 2332 and 1458 respectively), sparse networks using the concept of cohesion measured by connectivity. The first of these studies showed that membership in a cohesive block, defined by marital ties among households in an Austrian farming village, was correlated with stratified class membership, defined by single-heir succession to ownership of the productive resources of farmsteads and farmlands. In the second study, they found that the cohesive block defined by marital ties of Mexican villagers was restricted to a bicomponent that included families with several generations of residence and excluded spatial outliers and recent immigrants and families in adjacent villages. The cohesive block defined by *compadrazgo*,[42] on the other hand, crosscut this village nucleus and integrated spatial outliers and recent immigrants. In contrast to the first study, the Mexican case established a network basis for the observed cross-village egalitarian class structure.

[40] In both their analyses, they define two measures, one the highest k-connected subgraph to which each node belongs, and the other a measure of cohesive embeddedness, discussed above. The two measures will typically be highly correlated.

[41] NSF grant #BCS-9978282, "Longitudinal Network Studies and Predictive Cohesion Theory," Principal Investigator Douglas R. White and consultant Frank Harary, is focused on comparative studies of this type.

[42] Ritual kinship established between parents and godparents.

To illustrate how structural and path cohesion might be combined with a conditional distance approach, our example of cohesiveness structures in the Internet might be extended to study whether cohesive blocks of Web sites that exceed a certain conditional density (or effective diameter) may define the user communities and content or functional clusters on the Web. Like density and distance, diameter within cohesive sets—what we would call "conditional diameter"—is one of the other predictive measures that could be developed within a methodology for analysis of cohesive blocks, although we will not do so here. As a further example, "invisible colleges" in intellectual and citation networks are likely to be predicted both from cohesive blocks and conditional densities or diameters.

VI. SUMMARY AND CONCLUSION

In arguing for a distributed and scalable conception and measurement of cohesion, our purpose is to provide a theoretical foundation for asking some empirical questions about social cohesion that logically lie at the heart of sociology and social anthropology. It is clear that cohesion is an important concept fundamental to defining social groups and their boundaries as emergent phenomena. Work carried out by Watts (1999a,1999b) on the small world problem tied network structure to an important global characteristic of networks, their "connectedness," but like Wasserman and Faust (1994:115–17) he did not utilize the graph theoretic measurement of connectivity.[43] Small-scale "social psychological" cohesion based on the model of cliques and attachment-to-group makes a difference because social psychological cohesion affects the strength of group norms, how much individuals are willing to sacrifice for the groups, and so forth. But in a large network what difference does subgroup cohesion at a macro-level make when defined by *k*-component connectivity?[44] This is an empirical problem of considerable significance. Suppose we have two large groups both equal in the density of choices, both connected and similar in other specific structural features as well, but one being more cohesive

[43] Graph theoretic terminology has the advantage of distinguishing simple connectedness from connectivity, with its higher-order properties of structural and path cohesion. Watts uses the two terms interchangeably, thereby ignoring one of the major contributions of graph theory to the study of group cohesion. Wasserman and Faust are aware of the graph theoretic concept of connectivity as an indicator of cohesion but do not discuss methods or algorithms to apply to social networks to compute levels of connectivity.

[44] Another common use of the term connectivity is for the distribution of number of ties for each node, synonymous with the graph theoretical concept of degree.

than the other in the sense defined here. What would be the concomitant differences in the behavior of the groups and their members?

This paper goes back to a fundamental problem for network theory to establish a solid theoretic and measurement basis by which macrolevel questions of this sort can be examined in large-scale social networks. It is the foundational paper for a series of empirical studies addressing such issues. A subsequent paper in this series (Moody and White 2000) contains literature reviews and empirical analyses of case studies that examine the consequences of cohesion defined as *k*-component connectivities in large-scale social groups.

Within a social group, high connectivity plus a modest additional density of randomly distributed ties that reduce average path length within the group (Watts 1999b) is capable of generating large-scale group cohesion, as we showed in Section I.D. Our hypotheses (I.B) suggested how such cohesion might affect coordinated social action, social homogeneity, the emergence of group norms at the macro level and, in sufficiently compact groups (I.C), the emergence of interpersonal trust. Watts (1999a, 1999b) has alerted sociologists to reconsider the problem of "distributed social phenomena" in terms of his models and parameters of "small worlds"—large networks with local clustering of ties but relatively low average distance between members. Connectivity plus conditional density is another model of distributed social phenomena that defines a measure of cohesion not in terms of local clustering but of node-independent paths and invulnerability to disconnection by node removal.[45] The identification of bounded connectivity subgroups in a network is an ideal means of finding the boundaries of cohesive worlds ("small" in the sense of low average path lengths but not necessarily locally clustered) and measuring the degree of cohesiveness in each of their embedded subgroups.[46]

[45] A typical critique of connectivity-based measures of cohesion might run like this: "a cycle of 1000 people (connectivity 2) running from the United States to China and back does not constitute a cohesive group." Surely not, but a group of 1000 people with connectivity 5 (or higher, each a higher embedded level of cohesion), conditional density of 10 percent, and first and second shortest average path lengths of 3.5 and 4.0 (a "small world" as defined by Watts 1999a, b) is a large-scale group with considerable cohesiveness.

[46] Future research can combine the "small world" approach that takes as key variables the average path length of the first and second shortest independent paths between pairs of nodes (Watts 1999a, b) with our connectivity approach. When the average path lengths of the first and second shortest independent paths in a network are short, the logical implication is that the average cycle length between any two pairs of nodes is also relatively short (approximated by the sum of the two shortest independent path-length averages), and thus measuring the average cycle length of

A potential bias in favor of defining social cohesion as existing "only" in face-to-face groups may well have accumulated from the past century of small-group research, reinforced by "common sense" about primary groups. We have shown step by step, within both an intuitive and a graph theoretic framework, giving concrete hypotheses and sample results, the limitations of such a bias. Using the data of social networks we examined cohesion as the network component of a more inclusive concept of solidarity, which includes individual psychological attachments to a group, but does not stop at the boundaries of primary groups. We offered an alternative methodological perspective and detailed hypotheses as to how network cohesion may exist in large-scale groups as easily as in small.

Our precise and scalable method for measuring cohesion in networks and subgroups of any size has the advantage of detecting boundaries of subgroups, finding hierarchies of embedded subgroups, and measuring cohesiveness at each level of embedding. No other method of measuring cohesiveness has these advantages (most give an overwhelming welter of overlapping subgroups that introduce a second intervening level of analysis and interpretation), and our review of graph theoretic concepts (see also White 1998) shows no other method to possess equal validity in terms of the construct criteria that cohesive groups are "resistant to breaking apart" (see our definition 1.1.2) and are weaker or stronger in proportion to the "multiplicity of bonds" that hold them together (1.2.2). When combined with conditional density, the connectivity-based measure of cohesion has measurement validity in that our measure of cohesion increases with each additional edge added to a graph of fixed size. In the karate club example we showed how cohesive blocks and their conditional density contribute to a process of group division.

A fundamental intuition involved in the concept of social cohesion, we argue, is consistent with the idea that the greater the minimum number of actors whose removal disconnects a group, the greater the cohesion. Equivalently, as demonstrated by Menger's Theorem (1927), the greater the number of multiple independent paths, the higher the cohesion. The level of cohesion is higher when members of a group are connected as opposed to disconnected, and further, when the group and its

biconnectivity in the network. Similarly, the sum of the averages for first, second, and third shortest independent paths give an approximation of the average cycle length of triconnectivity, and so forth. In the case of triconnectivity, there are two independent "shortest cycles" between pairs of nodes. The relationship of average path length in *k*-connected structures needs to be investigated both in simulations and large-scale empirical network studies.

actors are not only connected but also have redundancies in their interconnections. Overlapping circles of friends increase social cohesion, for example, although our operationalization of the cohesion is not the same as the "intersecting social circles" concept of Simmel.[47] The higher the redundancies of independent connections between pairs of nodes, the higher the cohesion, and the more social circles in which any pair of persons is contained.[48]

To give a brief synopsis of the arguments behind the methodology presented here, given that the cohesiveness of a group is greater when there are higher redundancies of interconnections by multiple independent paths, the cohesion of a graph or subgraph is measured by its connectivity and, on a finer scale, by surplus density conditional on connectivity. Measurements of social cohesion by connectivity and conditional density are constructed by the following two sets of definitions derived from graph theory:

1. The *connectivity* $\kappa(G)$ of a graph G is the smallest number k of nodes whose removal disconnects any component of G or reduces the order of any component to a single node. A graph G is *k-connected* if $\kappa(G) \geq k$. A *k-component* is a maximal k-connected subgraph of G. A graph G is *k-cohesive*, to coin a new sociological term, if $\kappa(G) = k$. Hence, for each value of k, the k-components of a social network represented by a graph G define empirical social groups with a corresponding level k of cohesion. Subgroups with higher levels of cohesion are embedded in those with lower cohesion since the k-components of a graph form hierarchies by inclusion. According to Menger's Theorem, a graph is k-connected if and only if every pair of nodes is connected by at least k independent paths. The redundancy of multiple independent paths connecting actors is fundamental to measuring group cohesion as distinct from the proximities of actors in a network. The *structural cohesion* of a group is thus the minimum num-

[47] In Simmel's (1908) conception, zones around each ego or ego-memberships in groups simply overlap or intersect to form extensive connected networks (cf. Blau 1964; Kadushin 1966), but without necessarily forming higher-order cycles or connectivity sets (see footnote 25 regarding Alba and Kadushin's unsuccessful attempt to operationalize the higher-order cycles concept of cohesion).

[48] The widely used "social circles" approach to large-scale cohesion as webs of overlapping cliques (Alba 1972, 1973, 1982; Alba and Moore 1978) has the same defect as Freeman's (1996) intersecting cliques: pairs of nodes connected at some distance by multiple independent paths are not necessarily detected as part of the same cohesive subset.

ber *k* of its actors whose removal would not allow the group to remain connected or would reduce the group to but a single member. Hence a *k*-cohesive block is not only *k*-connected but every pair of actors is connected by at least *k* node-independent paths.

2. *Conditional density* measures cohesion on a finer scale, that of surplus density beyond that implied by connectivity. For each of the *k*-components of a graph, these two measures may be combined into an aggregate measure of social cohesion, suitable for both small- and large-scale network studies.

As a distributive phenomenon with emergent properties—such as might define the boundaries of social groups—connectivity is of crucial importance to the study of social networks. Many types of large-scale cohesive sets not detectable by other measures are identifiable from *k*-component connectivities. Correlations between hierarchical embeddedness in cohesive blocks and potential effects of cohesion—such as school attachment, or similarities in business behaviors, as in the studies of Moody and White (2000)—underscore the conceptual and substantive importance of connectivity as the primary measure of cohesiveness. In the study of social networks, both large and small, node connectivities and conditional densities are fundamental measurement concepts for social cohesion. Yet, one of the preconceptions about cohesion that is most resistant to change is the idea that in social networks, social interaction has only proximal effects, and that indirect effects quickly decay as we move from direct effects to effects along paths of distance 2 or 3, beyond which indirect effects are regarded as minimal.

It is worth stressing once more that what this bias in preconceptions of social cohesion omits are the two fundamental properties of the redundancies created by multiple independent pathways and multiple-node cut sets. First, independent pathways are convergent in their indirect effects, even at a distance. Independent paths between every pair of nodes in a cohesive block defined by connectivity at level *k* (which necessarily equals the minimum number of such paths) may more than compensate for the decay of effects of cohesive interaction along long paths. Studies of large-scale social diffusion, for example, typically rest upon and demonstrate the fact that long paths matter. What connectivity provides in terms of transmission effects within the internal networks of cohesive groups is the possibility for repetition along multiple independent pathways of rumor, information, material item, and influence transmission. Second, multiple independent pathways (equinumerous to minimum cuts) necessarily imply stronger bonding between pairs of nodes, regardless of distance decay. It is *k* times as hard to break apart a network tying nodes together by *k* node-independent pathways than it is to break apart a single chain that connects them.

The effects of multiple bonding and redundancy or repetition along convergent independent pathways are crucial in the formation of social coherence, social norms, sanctions and solidarities, and the emergence of socially or culturally homogeneous groups, and thus should be of focal interest to the study of social cohesion, including cohesion on a very large scale.

REFERENCES

Alba, Richard D. 1972. “COMPLT—A Program for Analyzing Sociometric Data and Clustering Similarity Matrices.” *Behavioral Science* 17:566.

———. 1973. “A Graph Theoretic Definition of a Sociometric Clique.” *Journal of Mathematical Sociology* 3:113–26.

———. 1982. “Taking Stock of Network Analysis: A Decade’s Results.” Pp. 39–74 in *Research in the Sociology of Organizations* 1, edited by Samuel B. Bacharach. Greenwich, CN: JAI Press.

Alba, Richard. D., and Charles Kadushin. 1976. “The Intersection of Social Circles: A New Measure of Proximity in Networks.” *Sociological Methods and Research* 5:77–102.

Alba, Richard D., and Gwen Moore. 1978. “Elite Social Circles.” *Sociological Methods and Research* 7:167–88.

Batagelj, Vladimir and Andrej Mrvar. 1997. “Networks/Pajek: A Program for Large Networks Analysis.” University of Ljubljana, Slovenia. http://vlado.fmf.uni-lj.si/pub/networks/pajek/.

———. 1998. “Pajek: A Program for Large Networks Analysis.” *Connections* 21:47–57.

Blau, Peter. 1964. *Crosscutting Social Circles*. Orlando, FL: Academic Press.

Borgatti, Stephen P., Martin G. Everett, and Paul Shirley. 1990. “LS Sets, Lambda Sets and Other Cohesive Sets.” *Social Networks* 12:337–57.

Brudner, Lilyan A., and Douglas R. White. 1997. “Class, Property, and Structural Endogamy: Visualizing Networked Histories.” *Theory and Society* 26:161–208.

Carron, Albert V. 1982. “Cohesiveness in Sport Groups: Implications and Considerations.” *Journal of Sport Psychology* 4:123–38.

Dirac, Gabriel A. 1960. “Généralisations du Théorème de Menger.” *Comptes Rendus de Académie des Sciences. Paris* 250:4252–53/405–09.

Doreian, Patrick, and Katherine L. Woodward. 1994. “Defining and Locating Cores and Boundaries of Social Networks.” *Social Networks* 16:267–93.

Even, Shimon. 1979. *Graph Algorithms*. Potomac, MD: Computer Science Press.

Everett, Martin G. 1982a. “Graph Theoretic Blocking of k-plexes and k-cutpoints.” *Journal of Mathematical Sociology* 9:75–84.

———. 1982b. "A Graph Theoretic Blocking Algorithm for Social Networks." *Social Networks* 5:323–46.

———. 1983. "An Extension of EBLOC to Valued Graphs." *Social Networks* 5:395–402.

———. 1984. "An Analysis of Cyclically Dense Data using EBLOC." *Social Networks* 6:97–102.

———. 1996. "Cohesive Subgroups." http://www.analytictech.com/networks/chapter_4_cohesive_subgroups.htm.

Fararo, Thomas, and Patrick Doreian. 1998. "The Theory of Solidarity: An Agenda of Problems." Pp. 1–31 in *The Problem of Solidarity: Theories and Models*, edited by Patrick Doreian and Thomas Fararo. Amsterdam: Gordon and Breach.

Festinger, Leon, Stanley Schachter, and Kurt Back. 1950. *Social Pressures in Informal Groups*. New York: Harper.

Ford, Lester R., and David R. Fulkerson. 1956. "Maximal Flow Through a Network." *Canadian Journal of Mathematics* 8: 399–404.

Freeman, Linton C. 1992. "The Sociological Concept of 'Group': An Empirical Test of Two Models." *American Journal of Sociology* 98:55–79.

———. 1996. "Cliques, Galois Lattices, and the Structure of Human Social Groups." *Social Networks* 18:173–87.

French, John R. P., Jr. 1941. "The Disruption and Cohesion of Groups." *Journal of Abnormal and Social Psychology* 36:361–77.

Gibbons, Alan. 1985. *Algorithmic Graph Theory*. Cambridge, England: Cambridge University Press.

Granovetter, Mark S. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78:1360–80.

———. 1983. "The Strength of Weak Ties: A Network Theory Revised." *Sociological Theory* 1:203–33.

Grannis, Richard. 1998. "The Importance of Trivial Streets: Residential Streets and Residential Segregation." *American Journal of Sociology* 103:1530–64.

Gross, Neal, and William E. Martin. 1952. "On Group Cohesiveness." *American Journal of Sociology* 52:546–54.

Harary, Frank. 1953. "On the Notion of Balance of a Signed Graph." *Michigan Mathematics Journal* 2:143–46

———. 1969. *Graph Theory*. Reading, MA: Addison-Wesley.

———. 1983. "Conditional Connectivity." *Networks* 13:347–57.

Harary, Frank, and Dorwin Cartwright. 1961. "The Number of Lines in a Digraph of Each Connectedness Category." *SIAM Review* 3:309–14.

Harary, Frank, and Yuji Kodama. 1964. "On the Genus of an n-connected Graph." *Fundamenta Mathematicae* 54:7–13.

Harary, Frank, Robert Norman, and Dorwin Cartwright. 1965. *Structural Models: An Introduction to the Theory of Directed Graphs*. New York: Wiley.

Johansen, Ulla C., and Douglas R. White. Forthcoming. "Collaborative Long-Term Ethnography and Longitudinal Social Analysis of a Nomadic Clan in Southeastern Turkey." In *Chronicling Cultures: Long-Term Field Research in Anthropology*, edited by Robert V. Kemper and Anya Royce. Walnut Creek, CA: Altamira Press.

Kadushin, Charles. 1966. "The Friends and Supporters of Psychotherapy: On Social Circles in Urban Life." *American Sociological Review* 31:786–802.

Kolchin, Valentin F. 1999. *Random Graphs*. Cambridge, England: Cambridge University Press.

Laumann, Edward O., Peter V. Marsden, and David Prensky. 1989. "The Boundary Specification Problem in Network Analysis." Pp. 61–87 in *Research Methods in Social Network Analysis,* edited by Linton C. Freeman, Douglas R. White and A. Kimball Romney. Fairfax, VA: George Mason University Press.

Libo, Lester M. 1953. *Measuring Group Cohesiveness*. Ann Arbor: Research Center for Group Dynamics, Institute for Social Research, University of Michigan.

Markovsky, Barry. 1998. "Social Network Conceptions of Solidarity." Pp. 343–72 in *The Problem of Solidarity: Theories and Models,* edited by Patrick Doreian and Thomas Fararo. Amsterdam: Gordon and Breach.

Markovsky, Barry, and Mark V. Chafee. 1995. "Social Identification and Solidarity: A Reformulation." Pp. 249–70 in *Advances in Group Processes,* vol. 12, edited by Barry Markovsky, Jodi O'Brien and Karen Heimer. Greenwich, CT: JAI Press.

Markovsky, Barry, and Edward J. Lawler. 1994. "A New Theory of Social Solidarity." Pp. 113–37 in *Advances in Group Processes,* vol. 11, edited by Barry Markovsky, Jodi O'Brien and Karen Heimer. Greenwich, CT: JAI Press.

Menger, Karl. 1927. "Zur Allgemeinen Kurventheorie." *Fundamenta Mathematicae* 10:96–115.

Mizruchi, Mark S. 1992. *The American Corporate Network, 1904–1974*. Beverly Hills, CA: Sage Publications.

Moody, James, and Douglas R. White. 2000. "Social Cohesion and Connectivity: A Hierarchical Conception of Social Groups." *Santa Fe Institute Working Paper* 00-08-049, submitted to the *American Journal of Sociology*.

Moreno, Jacob L., and Helen H. Jennings. 1937. "Statistics of Social Configurations." *Sociometry* 1:342–74.

Morowitz, Harold. 2000. "Emergence and the Law." *Complexity* 5(4):11–12.

Mudrack, Peter E. 1989. "Defining Group Cohesiveness: A Legacy of Confusion?" *Small Group Behavior* 20:37–49.

Palmer, Edgar N. 1985. *Graphical Evolution: An Introduction to the Theory of Random Graphs*. New York: Wiley.

Powell, Walter W., Douglas R. White, Kenneth W. Koput, and Jason Owen-Smith. 2001. "Evolution of a Science-Based Industry: Dynamic Analyses and Network Visualization of Biotechnology." Working paper, Santa Fe Institute.

Scott, John. 2000. *Social Network Analysis: A Handbook.* 2d ed. London: Sage.

Seidman, Stephen B. 1983. "Internal Cohesion and LS Sets in Graphs." *Social Networks* 5:97–107.

Seidman, Stephen B., and Brian L. Foster. 1978. "A Graph-theoretic Generalization of the Clique Concept." *Journal of Mathematical Sociology* 6:139–54.

Simmel, Georg. 1908. *Soziologie: Untersuchungen uber die Formen der Vergesellschaftung*. 2d ed. Munich, Germany: Duncker and Humbolt. Partly translated [1950] in *The Sociology of Georg Simmel*, edited by K. H. Wolf. New York: Free Press.

Skvoretz, John. 1998. "Solidarity, Social Structure, and Social Control." Pp. 373–402 in *The Problem of Solidarity: Theories and Models,* edited by Patrick Doreian and Thomas Fararo. Amsterdam: Gordon and Breach.

Tutte, William T. 1966. *Connectivity in Graphs*. Toronto: University of Toronto Press.

Wasserman, Stanley, and Katherine Faust. 1994. *Social Network Analysis: Methods and Applications*. Cambridge, England: Cambridge University Press.

Watts, Duncan J. 1999a. *Small Worlds: The Dynamics of Networks Between Order and Randomness*. Princeton Studies in Complexity. Princeton: Princeton University Press.

———. 1999b. “Networks, Dynamics, and the Small-World Phenomenon.” *American Journal of Sociology* 105:493–527.

White, Douglas R. 1998. “Concepts of Cohesion, Old and New: Which Are Valid, Which Are Not?” Working paper, Department of Anthropology, University of California, Irvine.

———. 1999. “Longitudinal Network Studies and Predictive Cohesion Theory.” NSF grant #BCS-9978282, http://eclectic.ss.uci.edu/~drwhite/nsf/nsf.htm

White, Douglas R., and Mark Newman. 2001. “Fast Approximation Algorithms for Finding Node-independent Paths.” Working paper, Santa Fe Institute.

White, Douglas R., M. Schnegg, L. A. Brudner, and H. Nutini. Forthcoming. “Conectividad Múltiple y sus Fronteras de Integración: Parentesco y Compadrazgo en Tlaxcala Rural.” In *Redes Sociales: Teoría y Aplicaciones*, edited by Jorge Gil and Samuel Schmidt. Mexico City: UNAM Press.

Whitney, Hassler. 1932. “Congruent Graphs and the Connectivity of Graphs.” *American Journal of Mathematics* 54:150–68.

Zachary, Wayne W. 1975. *The Cybernetics of Conflict in a Small Group: An Information Flow Model*. M.A. thesis, Department of Anthropology, Temple University.

———. 1977. “An Information Flow Model for Conflict and Fission in Small Groups.” *Journal of Anthropological Research* 33:452–73.