Structural Cohesion and Embeddedness: A hierarchical conception of social groups.*

James Moody
Department of Sociology
The Ohio State University

&

Douglas R. White
Department of Anthropology
University of California-Irvine

5/9/2001

Word Count (including text, figures, tables and bibliography): aprox. 14,000

* This research uses data from the Add Health project, a program project designed by J. Richard Udry (PI) and Peter Bearman, and funded by grant P01-HD31921 from the National Institute of Child Health and Human Development to the Carolina Population Center, University of North Carolina at Chapel Hill, with cooperative funding participation by the National Cancer Institute; the National Institute of Alcohol Abuse and Alcoholism; the National Institute on Deafness and Other Communication Disorders; the National Institute on Drug Abuse; the National Institute of General Medical Sciences; the National Institute of Mental Health; the National Institute of Nursing Research; the Office of AIDS Research, NIH; the Office of Behavior and Social Science Research, NIH; the Office of the Director, NIH; the Office of Research on Women's Health, NIH; the Office of Population Affairs, DHHS; the National Center for Health Statistics, Centers for Disease Control and Prevention, DHHS; the Office of Minority Health, Centers for Disease Control and Prevention, DHHS; the Office of Minority Health, Office of Public Health and Science, DHHS; the Office of the Assistant Secretary for Planning and Evaluation, DHHS; and the National Science Foundation. Persons interested in obtaining data files from The National Longitudinal Study of Adolescent Health should contact Jo Jones, Carolina Population Center, 123 West Franklin Street, Chapel Hill, NC 27516-3997 (email: jo_jones@unc.edu). Research leading to new applications of the concept of connectivity to large scale social networks was funded by NSF grants SBR-9310033 and BCS-9978282 to Douglas White, the first entitled "Social Transmission and Exchange: Cooperative Research at UCI, UNI Cologne, CNRS Paris 1993-95," and the second "Longitudinal Networks Studies and Predictive Social Cohesion Theory, 1999-2002." The first grant was matched by an Alexander von Humboldt Transatlantic Cooperation Award to Thomas Schweizer, University of Cologne, who provided productive commentary on the applications of connectivity and network methods to large-scale ethnographic projects. Thanks are due to Alexis Ferrand who provided a venue for the first presentation of some of these ideas at the Université de Lille, Institut de Sociologie, in 1998, and to Frank Harary who provided dense critiques and suggestions for the formal graph theoretic terminology. Thanks also to Mark Mizruchi for making his data on interlocking directorates available, to Scott Provan, David R. Karger, Mechthild Stoer and Sandeep Sen for help with the algorithmic aspects of this paper, and to Peter Bearman, Lisa Keister, Bob Farris, David Jacobs and three anonymous reviewers for comments on earlier drafts. In making final revisions, we are indebted to the commentary of the following members of the Santa Fe Institute Working Group on Co-Evolution of Markets and the State: John Padgett, Walter Powell, David Stark, Sander van der Leeuw, and Sanjay Jain.

Structural Cohesion and Embeddedness:
A hierarchical conception of social groups

Abstract

While questions about social cohesion lie at the core of our discipline, no clear definition of cohesion exists. We present a definition of structural cohesion based on network connectivity that leads to an operationalization of a dimension of social embeddedness. Structural cohesion is defined as *the minimum number of actors who, if removed from a group, would disconnect the group*. This definition leads to hierarchically nested groups, where highly cohesive groups are embedded within less cohesive groups. An algorithm developed and implemented (by Authors) identifies these nested groups by levels of structural cohesion, and thus measures the maximum levels of structural cohesion possessed by individuals as members of structurally cohesive sub-groups. We discuss the theoretical implications of this definition and demonstrate the empirical applicability of our conception of nestedness by testing the predicted correlates of our cohesion measure within high school friendship and interlocking directorate networks.

"...social solidarity is a wholly moral phenomenon which by itself is not amenable to exact observation and especially not to measurement."
(Durkheim, (1893 [1984], p.24)

"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 of three, a group continues to exist even in case one of the members drops out."
(Simmel (1908 [1950], p. 123)

Introduction

Questions surrounding social solidarity – understanding how social collectivities are bound together – are foundational for sociologists and have engaged researchers continuously since Durkheim. Analytically, solidarity can be partitioned into an *ideational* component, referring to the psychological identification of members within a collectivity and a *relational* component, referring to the connections among members of the collectivity. Durkheim identified the connection between these components in *The Division of Labor* when he linked changes in the common consciousness to the transition from mechanical to organic societies. In what follows, we identify an important feature of the relational dimension of social solidarity that scales across collectivities of any size. For clarity and theoretical consistency, we refer to this aspect of the relational dimension of social solidarity as *structural cohesion*, which we show rests on the pattern of multiple connections within the group. This conception of structural cohesion simultaneously defines a group property that depends on relations within the group, a positional property that situates the group relative to others in the population, and individual properties of memberships in structurally cohesive groups. We argue that this particular pattern of relations has deep implications for general social organization that is relevant across many sub-fields within sociology. The methodological result of this conception of structural cohesion is a technique that we call *cohesive blocking*, which we demonstrate on highschool friendship and interlocking directorate networks.

Theoretical definitions of social cohesion (and corresponding measures of "cohesive groups") are not new. Conflation of the ideational and relational dimensions of solidarity, however, defining both as "cohesion," is not uncommon. Lack of consensus has lead to a wide variety of empirical measurements ranging from individual feelings (Bollen and Hoyle 1990) to global features of relational structures (Freeman 1992; Friedkin 1984), and multiple stages in-between (Carron 1982; McPherson and Smith-Lovin 1986). The resulting "legacy of confusion" surrounding cohesion (Mudrack, 1989) hampers our ability to systematically extend theoretical treatments of social cohesion beyond programmatic statements or intuitive notions. Instead, a theoretical divide has developed between researchers employing widely differing empirical measures on small groups and theorists interested in large networks or general properties of social organization. At least part of this theoretical divide likely results from scale differences. Work on social cohesion and cohesive groups has mainly focused on small "primary groups": often with fewer than a dozen and rarely more than 100 people.[1] Theorists interested in how societies are organized, how redundant links help information travel long distances, or how class structures are maintained, seemingly have little use for work based in such limited settings. In contrast, we argue that structural cohesion, as one element of the relational dimension of social solidarity, defines a fundamental dimension of social organization that is scale independent.

The confusion surrounding cohesion and solidarity is primarily a theoretical problem. If we wish to build meaningfully on prior empirical work, theorists need to exactly specify the types of patterns that generate structural cohesion and theoretically link these patterns to substantive outcomes through operative mechanisms, such as the flow of communication or resources. While often inconsistent and vague, the common intuition surrounding previous notions of structural

[1] Due to the long history of small, face-to-face research on "groups", we would prefer to avoid the use of this term altogether, in favor of broader terms such as "collectivity" or "sub-structure" that carry much less theoretical baggage. Such a substitution, however, results in decidedly awkward writing. We thus maintain the use of "group" at times, but remind readers that our conception is not limited to the small face-to-face primary group structures commonly referred to by the term.

cohesion rests on the robustness of the collectivity to disruption. "Cohesive groups" are thought to be those that are held together well and thus difficult to break apart. Extending Simmel's insight on the distinction between dyads and triads, we argue that a group so fragile that the removal of a single person would destroy the group is not very cohesive, while a group that can withstand the loss of many members is much more cohesive. This extension implies a differentiation between two poles of social organization, with corresponding theoretical implications for the character of the group. On the one hand, collectivities can be united through connections that are dependent on individual actors, which we label "weakly structurally cohesive." On the other hand, collectivities can be united through a diffuse pattern of relations that weave members together through multiple independent connections, which are more "strongly" structurally cohesive (see also White and Harary 2001) and for which we will develop a precise quantitative measure.[2]

Identifying cohesive structures is only one part of analyzing groups, and a more informative approach simultaneously tells us how such groups relate to each other. Our conception of structural cohesion necessarily entails a positional analysis of the resulting groups with respect to their *nesting* in the population at large. Theoretically, the concept of *nestedness* captures one dimension of Granovetter's concept of social embeddedness (Granovetter 1985; Uzzi 1996). Like "solidarity", "embeddedness" is a multidimensional construct relating generally to the importance of social networks for actors, indicating that actors integrated within multiplex social networks face different sets of resources and constraints than those who are not embedded in such relations. While widely used in sociology, embeddedness has rarely been given an explicit operational definition. When contrasted with individualistic theoretical frameworks that deny the importance of relations for behavior, this level of theoretical ambiguity is profitable. Allowing embeddedness to remain at the level of orienting statement, however, significantly lowers our ability to build on previous work.

---

[2] The strong/weak distinction will not remain categorical, as with strong and vs. weak components in a graph, nor imprecise, as between "weak" and "strong" ties (Granovetter 1973).

While we do not claim to capture the full range of either "cohesion" or "embeddedness," cohesive blocking provides an exact analytic operationalization of a dimension of each. After reviewing previous theoretical work on social solidarity, we provide a formal definition of structural cohesion that can be directly operationalized as network connectivity (Harary 1969). We then demonstrate the value added of our conception by applying it in two widely differing settings: high school friendship networks and interlocking directorates. With this work, we extend insights on social cohesion to collectivities ranging in size from small groups to many thousands of people, which applies equally well across multiple sociological domains. The qualitative relational feature, as Simmel pointed out, is whether a group depends on particular individuals for its connectedness. The relevant quantitative measure is the number of individuals whose stability is required to keep the group connected.

Background and Theory

Research on social cohesion has been plagued with contradictory, vague and difficult to operationalize definitions (for reviews, see Doreian and Fararo 1998; Mizruchi 1992; Mudrack 1989). While previous definitions of cohesion seem to share a common intuitive core resting on how well a group is "held together," the conceptualization is often murky. What does it mean, for example, that cohesive groups should display “connectedness” (O'Reilly and Roberts 1977)? Or that cohesion is defined as a "field of forces that act on members to remain in the group" (Festinger, Schachter, and Back 1950) or "the resistance of a group to disruptive forces" (Gross and Martin 1952)? Dictionary definitions of cohesion rest on similar ambiguities, such as “[t]he action or condition of cohering; cleaving or sticking together” (OED, 2000).

In his review of the cohesion literature, Mudrack (1989:38) identifies the focus on individuals as a significant gap between researchers' conceptions and their operationalizations of cohesion. By focusing on individuals, social psychologists and small group researchers often confound outcomes and correlates of cohesion with cohesion itself, commingling the relational and ideational components of social solidarity. To be analytically useful, then, we must differentiate the togetherness of a group from the *sense* of togetherness that people express. Using subjective factors alone, such as a "sense of we-ness" (Owen 1985), "attraction-to-group" (Libo 1953) or the ability of the group to attract and retain members that may result *from* relational cohesion (or vice-versa) unnecessarily limits our ability to ask questions about how relational cohesion affects (and is affected by) social psychological factors.[3]

The ability to directly operationalize structural cohesion through social relations is one of the primary strengths of a relational conception of social cohesion. The "forces" and "bonds" that hold the group together are the observed relations among members, and cohesion is an emergent property of the relational pattern. While intuitively appealing, the practice of identifying relational cohesion has been limited. Unfortunately, much of the previous work on relational cohesion has focused on very small groups, and thus used cliques —collections of people with direct relations to every other person in the group — and their overlaps as the archetypes of structural cohesion.[4] Since the interpersonal investments required to form cliques out of dyadic relations rises exponentially with the number of people, this conceptual frame necessarily limits our ability to explore structural cohesion in large social settings.[5] This is unfortunate, since the use of

---

[3] Definitions that rest primarily on the ability to attract or retain members are theoretically weak. Consider a situation such as the stock market in the late 1990s. The large profits to be gained in the market *attracted* many and the continuing profits *retained* those same actors. Certainly all investors do not constitute a cohesive group, and if they were to form meaningful relations, the Securities and Exchange Commission would cite them for conflict of interest. On the other hand, examples of clear social units, such as dysfunctional families, where people are strongly linked (through blood) but unhappy to be members provide an example of structural cohesion where the psychological sense of "attraction" to the group might be quite low.

[4]The reason for the small group bias in relational work is likely largely a result of methodological limitations, which are thankfully starting to become less severe.

[5] The analysis of overlaps among cliques (Freeman 1996) to study cohesion suffers from two defects. First, pairs of overlapping cliques may have only the weakest cohesion between them if they share only a single member. By the same token, pairs of cliques with larger overlaps, say of size k, can have only level k of cohesion, of the number of actors in the overlap. Even more serious, however, is that even subgroups where everyone is a friend or a friend's friend may be excluded from clique- and clique-overlap analyses, as with four friends connected in a circle who do not constitute a clique. More generally, such analyses will often miss a great deal of the structure of social cohesion.

cliques constrains our ability to extend insights initially developed for small groups to social settings at a much wider scale. Much of the early work in social networks (Pool & Kochen 1978; Rapoport and Horvath 1961; White, et al. 1976), and its' continued theoretical underpinnings (Emirbayer 1997), suggests such an extension, which depends on identifying features that do not necessarily limit group size.

From the proceeding, a preliminary and intuitive definition of structural cohesion follows:

Def. 1.1. A collectivity is structurally cohesive to the extent that the social relations of its members hold it together.

There are five important features of this preliminary definition. First, it focuses on what appears constant in intuitive notions of cohesion: a property describing how a collection of actors is united. Second, it is expressed as a property of the group. Individuals may be embedded more or less strongly within a cohesive group, but if the maximum-sized group at a given level of strength is uniquely well defined, then the individuals and the group have a unique level of cohesion. Third, this conception is continuous.[6] Some groups will be weakly cohesive (not held together well) while others will be strongly cohesive. Fourth, structural cohesion rests on observable social relations among actors. Finally, network size, as we shall see, is irrelevant.

What, then, are the relational features that hold collectivities together? Clearly, a collection of individuals with no relations among themselves is not cohesive. If we imagine relations forming among a collection of isolates,[7] we could identify a point where each person in the group is connected to at least one other person in such a way that we could trace only a single path from each to the other. Thus, a weak form of structural cohesion starts to emerge as these islands be-

---

[6] See White and Harary (2001) for further refinement of this aspect of our definition.

[7] Actors without any relations to others in the population.

come connected through new relations.[8] This intuition is captured well by Markovsky and Lawler (1994) when they identify "reachability" as an essential feature of cohesion. Additionally, however, as new relations form among previously *connected* pairs, we can trace *multiple paths* through the group. Intuitively, the ability of the group to "hold together" increases with the number of independent ways that group members are linked.

That cohesion seems to increase as we add relations among pairs has mislead many researchers to focus on the *volume* (or density) of relations within and between groups as their defining characteristic (Alba 1973, Fershtman 1997; Frank 1995; Richards 1995). There are two problems with using relational volume to capture structural cohesion in a collective. First, consider again our group with one traceable path among all members. We can imagine moving a single relation from one pair to another. In so doing, the ability to trace a path between actors may be lost, but the number of relations remains the same. That is, if volume does not change (i.e., the number of ties remains constant when one rearranges ties in a network) but reachability does, then volume alone cannot account for structural cohesion.

Second, the initial (and weakest) moment of structural cohesion occurs when we can trace only a single path from each actor in the network to every other actor in the network (technically, this type of graph is known as a *tree*). Imagine further that our ability to trace a chain from any one person to another always passes through a single person (technically, this graph is known as a *star*). This might occur, for example, if all relations revolved around a charismatic leader. In the latter case, each person might have ties to the leader, and be connected only through the leader to every other member of the group. While connected, such groups (whose internal relations are trees or stars) are notoriously fragile. As Weber (1978: 1114) pointed out, the loss of a charismatic leader will destroy a group whose structure is based on an all-to-one re-

[8] See Hage and Harary (1996) for a discussion of this process among islands in Oceania. We recognize that social groups can form from the dissolution of past groups; the above discussion is useful only in understanding the essential character of structural cohesion.

lational pattern. Thus, increasing relational volume but focusing it through a single individual does not necessarily increase the ability of the group to hold together.

Yet, seemingly robust groups, such as cults and terrorist networks, do maintain an all-to-one relational structure. To distinguish collectivities that depend on single actors[9] from those that are multiply interconnected, we analytically distinguish *weak* from *strong* structural cohesion. While groups with an all-to-one relational organization *may* be stable and robust to some forms of disruption, this stability often results from the extraordinary efforts needed to maintain an inherently weak relational structure. Consider a terrorist network, for example, that might have a star (spoke-and-hub) configuration. For security, each spoke cell knows nothing of the other cells. Since any randomly captured member is unlikely to be the unique central node, such a person cannot know enough of the structure to put the entire group at risk, and the network as a whole can be maintained in the face of a concerted effort to destroy it. If the hub is identified, however, the organization as a whole will be destroyed. Thus, the stability of the structure depends on the ability to keep the hub hidden. The lengths to which such groups go to keep the hub hidden attests to their fundamental structural weakness.

Markovsky and Lawler (1994, Markovsky 1998) make a similar point when they argue that a uniform distribution of ties is needed to prevent a network from splitting into multiple subgroups.

> "… the organization of [cohesive] group ties should be distributed throughout the group in a relatively uniform manner. This implies the absence of any substructures that might be vulnerable, such as via a small number of 'cut-points' to calving away from the rest of the structure." (Markovsky, 1998 p. 245).

---

[9] Or a series of single actors, each of which is a separator of the graph, which relates to graph centralization. The distinction between the number of such central actors and the structure of centrality is a further dimension that distinguishes networks and has important implications for the quality of life within such collectivities. Note that high structural cohesion typically diminishes the total centralization of a network, but low structural cohesion, such as a single cycle, does not necessarily increase centralization. The two concepts measure very different things. We focus below on the importance of unilateral action for disruption or controlling the group or group's resources as the defining substantive feature of weak vs. strong structural cohesion (keeping in mind that these are poles along a continuum).

Such vulnerable substructures form when network relations are focused through a small number of actors. If pairs of actors are linked to each other through multiple others, the structure as a whole is less vulnerable to this type of split.

The substantive character of groups that are vulnerable to unilateral action differ significantly from those expected of groups with multiple independent connections, and often in ways that do not connote group solidarity. First, the group as a whole is vulnerable to the will and activities of those who can destroy the group by leaving. Dramatic instances of this characteristic are evident in cults, but can be seen as well in any group where a single individual's actions can determine the social standing of the collectivity. Second, actors that can disconnect the group are also actors that can control the flow of resources in the network. As has long been known from Network Exchange Theory, networks with structural features leading to control of resource flow generate power inequality (Willer 1999)). Third (as implied by Markovsky and Lawler), weakly cohesive organization promotes segmentation and fractionalization into subsets of members who remain cohesive after removal of a cut-point and are only minimally connected to the rest of group. Such subgroups may be more likely to develop norms and values that are distinct, leading to schisms and factions.

In contrast to weak structurally cohesive groups, however, collectivities that do not depend on individual actors are less easily segmented. The presence of multiple paths, passing through different actors, implies that if any one actor is removed, alternative linkages among members still exist to maintain social solidarity. Information and resources can flow through multiple paths, making minority control of resources within the group difficult. As such, the inequality of power implicit in weakly cohesive structures is not so pronounced in stronger ones.

Given the above, we amend our preliminary definition of structural cohesion to make explicit the importance of multiple connectivity.

> *Def. 1.2. A group is strongly (as opposed to weakly) structurally cohesive to the extent that multiple independent social relations among all pairs of members of the group hold it together.*

This definition provides a metric for our structural cohesion continuum and captures Simmel's concept of the uniquely social nature of groups in his discussion of dyads and triads (1950:135). In a dyad, the existence of the group rests entirely in the actions of each member: Recall that either member, acting unilaterally, could destroy the dyad by leaving. Much of the power and intimacy associated with purely dyadic relations rests on this fact. However, once we have an association of three, a weakly cohesive group will remain even if one of the members leaves. In triads, the social unit is no longer dependent upon a single individual, and thus the actions of the social unit take on new (and uniquely social) characteristics. The key insight captured by this definition is the supra-individual status of the social unit, which is not limited to small groups, but can scale to relations among groups of any size. Groups of any size that depend on connections through a single actor are at one pole of the measure given in Def. 1.2, weakly cohesive, while those that rest on connections through two actors are stronger, and those depending on connections through many actors are stronger yet. The strongest cohesive groups are those (cliques) in which every person is directly connected to every other person, though this level of cohesion is rarely observed except in small of primary groups.[10] For large groups, there is usually a threshold of diminishing returns to higher connectivity.

A Formal Specification of Structural Cohesion

To formalize our conception of structural cohesion, and thereby develop a method for identifying cohesive groups in a network, we need a language capable of accurately expressing relational patterns in a group. The language of graph theory provides this analytical purchase. Formally, we refer to a graph **G**(**V**,**E**), where the vertices, **V**, represent our set of n actors and the edges, **E**, represent the m relations among actors defined as a set of pairs (v_i, v_j). Actor *i* is *adja-*

[10] It is important that our measure of structural cohesion has its maximum pole with the fully connected clique, linking us to previous conceptions of network cohesion. Our contribution, methodologically, is to show which feature of the fully connected clique can be generalized to larger collectivities to account for social structural cohesion.

cent to actor *j* if $(v_i, v_j) \in$ **E.**[11] Structural cohesion depends on how pairs of actors can be linked through chains of relations, or *paths*. A path in the network is defined as an alternating sequence of distinct nodes (vertices) and edges, beginning and ending with nodes, in which each edge is incident with its preceding and following nodes. We say that actor *i* can *reach* actor *j* if there is a path in the graph starting with *i* and ending with *j*. Two paths from *i* to *j* are *(node-)independent* if they have only nodes *i* and *j* in common.[12] If there is a path linking every pair of actors in the group then the graph is *connected*. In general, a set is *maximal* with respect to a given property if it has the property but no proper superset does. A *component* of a graph **G** is a maximal connected subgraph of **G**. If there are two or more (node-independent) paths connecting every pair of nodes in the graph, the graph is biconnected, and in general, if there are at least *k* (node-independent) paths connecting every pair, the graph is *k-connected* and called a *k-component*. Components are the maximal sets of actors in which each actor is reachable from every other. Likewise, in any component of a graph, the paths that link two non-adjacent vertices must pass through a given subset of other nodes. These nodes, if removed, would disconnect the two actors. Any such set of nodes, **S**, is called an (*i*,*j*) *cut-set* if every path connecting *i* and *j* passes through at least one node of **S**. If there is only one node in **S**, it is called a cut-node. Let **N**(*i*,*j*) be the smallest size of an (*i*,*j*) cut-set, called the *(node-) connectivity, k,* of **G**. One of the deepest theorems about graphs in general is that any graph with node connectivity *k* is at most *k*-connected, and any graph that is precisely k-connected has node connectivity k. The intuitive cut-set "resistance to being pulled apart" criterion of cohesion and the multiple independent paths "held to-

[11] For the purposes of connectivity, we assume that actors do not relate to themselves and thus $(v_i, v_i) \notin$ **E**.

[12] Similarly, two paths from *i* to *j* are *edge-independent* if they have no edges in common (Harary 1969). White and Harary (2001) examine the weaker adhesive bonds that result from edge-connectivity.

gether" criterion of cohesion are thus formally equivalent[13] (see Harary 1969 for Menger's proof).[14]

In general, the maximal sub-graphs of **G** at various levels of connectivity *k* are called the *connectivity sets* or *k-components* of the graph. In common terminology a 2- or biconnected component is called a *bicomponent* and a 3-connected component a *tricomponent*; these in turn are called *giant bicomponents, tricomponents,* etc., if they contain more than 50% of the nodes if a graph. Figure 1 presents examples of networks with differing levels of structural cohesion. Note that in each of these three groups the *number* of relations is held constant, but the edges are *arranged* such that total structural cohesion increases from left to right.

---

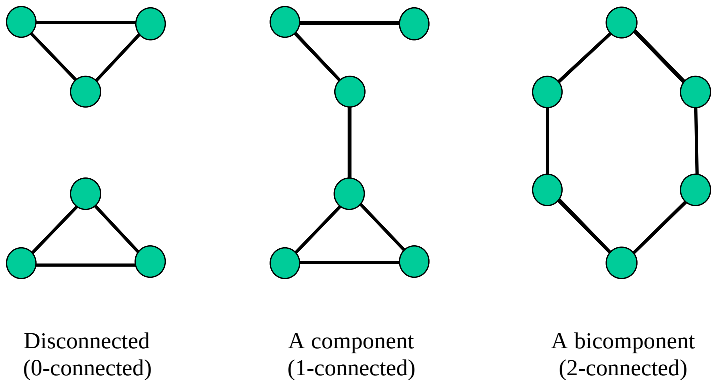
[13] It is easy to show that for any u,v pair of individuals, N_p, the maximum number of node-independent paths between them in the group is no greater than N_s, the minimum number of other individuals or direct u,v ties whose removal is needed to disconnect them within a group, but it requires a deep proof to show that $N_s = N_p$.

[14] Harary et al. (1965:25) were the first to note that "in general, if the sociometric structure of a group forms a block [in this context: is 2-connected], then the group is very cohesive." Computation of bicomponents (but not higher k-connected subgroups) is implemented in Pajek and UCINET. Harary's (1969) definition of levels of connectivity is cited by Wasserman and Faust (1994:115-117) as one way to conceptualize the cohesion of a graph, but they do not apply it to finding cohesive subgroups. Nor do they discuss the relevance of Menger's theorem to connectivity as a measure of cohesion. White (1998) is the first to define structurally cohesive subgroups in terms of k-connectivity and to introduce the relevance of Menger's theorem to measurement of structural cohesion, insofar as we are aware, and shows that groups formed through other procedures based on generalizing the clique concept (k-plex, n-clique etc.) are not necessarily structurally cohesive.

In a companion paper, White and Harary (2001) formalize the definition of structural cohesion and the critiques of alternate measures of cohesive subgroups and then go on to discuss the relation between connectivity and density. They also develop second but weaker dimension upon which such groups could be arranged that relates to *edge-connectivity*, measured by the minimum number of edges that must be removed in a connected group that will result in its disconnection. It can be shown that a graph of any level of edge-connectivity may still be separable by removal of a single actor, which means that the unilateral power of actors can be high even when there are many relations connecting people (White and Harary 2001). Hence edge-connectivity is a much weaker measure (which they call adhesion) than node-connectivity (structural cohesion). The key to structural cohesion thus rests in how the pattern of relations makes unilateral action impossible.

The unique methodological contributions of the present paper are to develop (1) the algorithm for computing k-connected subgroups, and (2) the measure of nestedness.

Figure 1. Examples of Connectivity Levels



Based on the intuitive notions captured in definition 1.2 and the formal graph properties presented above, we can now provide a final definition of structural cohesion.

> Def. 1.3a. *A group's structural cohesion is equal to the minimum number of actors who, if removed from the group, would disconnect the group.*

This definition can be operationalized directly with node connectivity. Note that for successive values of connectivity k, for any given network, there is a unique set of subgroups with this level of structural cohesion, each defined by a maximal set of actors with that level of connectivity. Because of the formal equality between the number of cut-nodes and the number of (node-) independent paths, the disconnect version of definition 1.3a can be restated *without any loss of meaning* in "held together" terms as:

> Def. 1.3b. *A group's structural cohesion is equal to the number of independent paths linking each pair of actors in the group.*

This pair of equivalent definitions of structural cohesion retains all five aspects of our original intuitive definition of structural cohesion, while providing an exact operationalization in graph connectivity that can be applied to groups of any size. Given that the graph is connected, the greater the dependence on a small number of actors, the weaker the structural cohesion, and the more the vulnerability of the group to the activity of fewer and fewer members. As the connectivity of a graph increases, vulnerability to unilateral action decreases, and the collectivity takes on the characteristics associated with strong structural cohesion. Based on Simmel's discussion of the dyad, we argue that a connectivity of 2 (as in a bicomponent) is the minimum level of distinction between the weakest and the stronger structurally cohesive groups. The latter, of course, are ranked by their k-connectedness.

For cohesion in very large groups, the key insight contained in these definitions is that the number of ties at the level of individuals need only be above a manageable minimal threshold (say of degree 7-10) in order for the minimum numbers relevant to the structural cohesion of the group as a whole[15] to be at that threshold as well. That is, a connectivity of 10 for a very large group – which is a manageable number of ties at the level of every individual in the group – is a high level of structural cohesion for the group as a whole, even if the group numbers in the thousands.

This conception of cohesion differs markedly from other approaches to identifying "cohesive groups" in social networks.[16] Group identification methods based on number of interaction partners (k-cores), minimum within group distance (N-cliques), or relative in-group density, *may* be structurally cohesive, but are not *necessarily* so. In every case, the method used to identify groups cannot distinguish multi-connected groups from groups vulnerable to action from (such as removal or disconnection of) a single individual. As such, any empirical application of these methods to a theoretical problem of *structural cohesion* risks ambiguous findings. Because such measures optimize a topological feature that only correlates with node connectivity, but optimizes a feature other than node connectivity, observed empirical results may be uniquely due to the particular features optimized by each technique instead of structural cohesion *per se*. By distinguishing structural cohesion from factors such as density or distance, we can isolate the relative importance of connec-

---

15 These are N_p, the maximum number of node-independent paths between them (Def. 1.3b), and N_s, the minimum number of direct ties or other individuals whose removal is needed to disconnect them within the group (Def. 1.3a), whose equivalence is proven by Menger's Theorem.

16 A detailed comparison of each alternative method to connectivity, expanding on those of White (1998) and White and Harary (2001), is available from the authors upon request.

tivity in social relations from these other factors. Distance between members, the number of common ties and so forth might affect outcomes of interest, but our ability to extend social theory in formal network terms depends on our ability to *unambiguously* attribute social mechanisms to such topological features. Connectivity provides researchers with the ability to disentangle the effects of structural cohesion from other topological features of the network.

Theoretical Implications of Network Connectivity

The fact that information, resources and risks flow through social relations is an important reason why networks matter for social life. If resources flow through networks then the structure of the network largely determines the distribution of resources, access to information, and allocation of power in a social system. Previous work on power exchange structures (Blau 1964; Cook et al. 1983; Lawler and Yoon 1993; Willer 1999) and structural holes (Burt 1992) all suggest that the ability to control resource flow is an important source of power.

A connectivity conception of the relational component of social (i.e., structural) cohesion provides a direct link between the distribution of resources and the distribution of connections through which resources flow. One of the defining properties of a *k*-component is that every pair of actors in the collectivity is connected by at least *k* independent paths. This implies for $k > 1$ that information and resources can be exchanged within the group without any individual (or ensemble up to k-1 individuals) being able to control the resources. Substantively, such networks are characterized by a reduction in the power provided by structural holes (Burt 1992; White et al. 1976): The ability of any person to monopolize resource flow, and thus have power within the setting, is limited as connectivity increases.

The industrial development of "just-in-time" inventory systems provides an interesting example. When viewed as a network of resource flows, the most efficient systems resemble spanning trees — and thus the production of the entire product depends on each element producing a required good at the right time. Under this structure, labor has accentuated power since strikes at one plant can disable the entire system. Strikes effectively remove the struck factory from the production network and disconnect the overall production process. In fact, recent trends toward "just-in-time" production processes are not new, but were used extensively early in the auto industry. It became clear, however, that this production structure gave labor power. To counter, management expanded the production network to include alternative sources (other factories and storehouses), building redundancy into the system (Schwartz, 2001).

The length of a path is often considered critical for the flow of goods through a network, as flow may degrade with relational distance. That is, the probability that a resource flows from node *i* to node *j* along path *p* is equal to the product of each dyadic transition probability along that path. When multiplied over long distances the efficacy of the information diminishes even if the pairwise transmission probability is high. For example, the probability that a message will arrive intact over a 6-step chain[17] when each dyadic transmission probability is 0.9 will be 0.53. The fragility of long-distance communication rests on the fact that at any step in the communication chain, one person's failure to pass the information will disrupt the flow. For a structurally cohesive group, however, expected information degradation decreases with each additional *independent* path in the network. The comparable probability of a 6-step communication arriving given two independent paths is 0.78.[18] As the number of independent paths increases, the likelihood of the information transmission increases.[19]

---

[17] The purported average acquaintance distance among all people in the United States (Milgram 1969).

[18] We calculate this as the product of the dyadic probabilities for each path, minus the probability of transmission through both paths. Thus, for two paths the formula is $2(p_{ij})^{d} - (p_{ij})^{2d}$, where d is the distance.

[19] It is possible to limit connectivity by path distance, under circumstances where theory suggests long-distance flow is not possible. Define a longest chain as a path of maximal length d between endnodes *u* and *v* such that no node between *u* and *v* has degree > 2. It is easy to prove that there exists a unique set of longest chains for each length d in any graph, and to eliminate all such chains for d greater than some cutoff. To limit connectivity sets by path distance, we define a cutoff d and remove all longest chains of length > d for a graph **G**, resulting in $\mathbf{G}^{d}$. For subgraphs of $\mathbf{G}^{d}$ with connectivity *k* there will then be *k* or more paths of length *d* or less between every pair of nodes. We caution against this approach, under most circumstances, since connectivity structures can boost the transmission signals of flow that researchers might otherwise think has little chance of traveling long distances.

When the flow is not subject to degradation, but only to interruption, increasing connectivity will provide faster and more reliable transmission throughout the network.[20] In a high-connectivity network, even if many people stop transmission (effectively removing themselves from the network), alternate paths provide an opportunity for spread. If a virus is flowing through the network, multiple transmission routes make prevention much more difficult, increasing the hazard of the network for disease distribution. For information, high connectivity increases reliability, since multiple exposures allow individuals to combine similar information from multiple sources. Similar examples can be found with respect to the flow of normative and other cultural goods. A theory of norms based on socialization and transmission would suggest that the higher the level of cohesion, the greater the consistency in normative behavior.

The importance of network connectivity rests on how subgroups within the population are related. It is always possible for two k-cohesive groups to overlap by k-1 nodes. However, we often observe distinct, non-overlapping ($k+j$)-cohesive subgroups empirically within a larger k-connected population. This structure has important implications for the carrying capacity of a network over long distances. Local pockets of high connectivity can act as amplifying substations for information (or resource, or viral) flow that comes into the more highly connected group, or boosting a signal's strength[21], and sending it back out into the wider population. The observed patterns typical in small world graphs (Milgram 1969; Watts 1999; Watts and Strogatz 1998) are a natural result of local relational action nested within a larger network setting. Thus, processes based on the formal properties of connectivity may account for many of the observed substantive features of small world networks.

---

20 Computer viruses are an excellent example of such flows, as recent outbreaks such as "Melissa" and the "Love Bug" show.

21 Signal amplification might depend on averaging or combining degraded copies of the same signal or message so as to filter noise, thus increasing reliability.

Since connectivity sets can overlap, group members can belong to multiple groups. While observed overlaps at high levels of connectivity may be rare, any observed overlaps are likely substantively significant.[22] If an individual belongs to more than one maximal *k*-cohesive group, that individual is part of a unique subset of *k*-1 individuals whose removal will disconnect the two groups. Members of such bridging sets may be *structurally equivalent* with respect to the larger cohesive sets that they bridge. As such, a *positional* and *relational* structure comes out of the analysis of cohesive groups, groups that are much larger, fewer, and easier to distinguish than by our traditional notions of sociological cliques, providing some of the same theoretical purchase blockmodels were designed to provide (Burt 1990; Lorrain and White 1971; White et al. 1976), but focusing on subgraphs that may overlap rather than partitions of nodes.[23]

This relational conception of cohesion provides sociology with a useful tool for understanding processes related to the formation of social classes, ethnicity, and social institutions. While a longstanding promise of network research (Rapoport and Horvath 1961; Emirbayer 1997; White et al. 1976), the conceptual tools needed to identify the empirical traces of such processes have been sorely lacking. In contrast, Brudner and White (1997) and White et al. (in press) identified sociologically important structurally cohesive sets in two large (n = 2332 & 1458 respectively) and sparse networks. The first of these studies showed that membership in a structurally cohesive group, defined by marital ties among households in an Austrian farming village, was

---

[22] Some researchers consider overlapping subgroups too empirically vexing to provide useful analysis (see also footnote 5). It is important to point out that (1) k-components are strictly limited in the size of such overlaps, making the substantive number of such intermediate positions small -- especially compared to cliques, (2) that each such position, because of its known relation to the potential flow paths and cycle structure of the network, can be theoretically articulated in ways that are impossible for clique overlaps, and (3) even when they are empirically difficult to handle, may well be an accurate description of relationship patterns. Arguments that relative density groups (c.f. Frank, 1995) solve this problem by assigning each actor to their preferred group (based on number of nominations) fail to account for people who have ties across many sub-groups, such that the total number of ties to people in other groups is higher than the number of ties to people in the group they have been assigned to.

[23] It is for this reason that cohesive blocking cannot in general be subsumed as a form of blockmodeling: cohesive blocks may overlap, and do not form partitions. Overlaps are crucial to cohesive structures.

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 structurally cohesive group defined by marital ties of Mexican villagers was restricted to a core that included families with several generations of residence and excluded recent immigrants and families in adjacent villages. The structurally cohesive group defined by *compadrazgo*,[24] on the other hand, crosscut this village nucleus and integrated recent immigrants. In contrast to the first study, the Mexican case established a network basis for the observed cross-village egalitarian class structure.

Linking structurally cohesive subgroup membership to institutions that provide formal access to power suggests a new approach to the study of social stratification and the state. White et al. (1999), for example, identify an informally organized “invisible state” created by the intersections of structurally cohesive groups across multiple administrative levels. They show that those who share administrative offices during overlapping time spans build dense clique-like social ties within a political nucleus while maintaining sparse locally tree-like ties with structurally cohesive groups (globally multiconnected) in the larger region and community. The locally dense and the globally sparse multiconnected ties act as different kinds of amplifiers for the feedback relations between larger cohesive groups and their government representatives.

Multiple connectivity has been shown important for the dynamic development of new organizational forms. Using some of the conceptual tools developed here, Powell, White and Koput (n.d.) showed that multiconnectivity played a critical role in the formation of the networked organizational form that facilitated communication, coordination and coopoeration in the biotechnical industry (see Powell 1990, 1996, Power et al. 1996). Over time, multiconnectivity changed dramatically. While simple connectivity (the size of the giant component) rose linearly, multiconnectivity showed a dramatic non-linear increase, followed by relative stability. This in-

[24] Ritual kinship established between parents and godparents.

crease rested on a hierarchy of nested hubs of increasing k-connectedness (rising to a connectivity of 8 or more for hundreds of firms). Interestingly, this phase shift corresponded to a period of maximum financial investment and venture capital excitement. Attesting to the substantive importance of multiconnectivity, firms that were higher in connectivity were, by all indicators, more successful.

Additionally, Powell, White and Koput's (n.d.) ability to link structural dynamics to changes in actor-level behavior suggests that multiple connectivity may be important in a broad class of relational structures. The empirical biotechnical data fit a degree-based power-law model (see Newman et al. 2000 for methodology) for popularity bias in terms of previous ties almost perfectly. The prediction from micro-behavior of choice of partners according to a popularity bias (resulting in power-law distributions of nodal degree) to a topological outcome of hierarchically nested connectivity hubs, as seen in the evolution of the biotechnical network, implies that this type of connectivity structure may be important in many other types of networks that have actor-degree power-law distributions. These include the scientific collaboration networks studied by Newman (2001), the internet (Barabási and Albert 1999, Albert et al. 1999), the electricity grid (Watts 1999), movie actor and director co-participation (Newman, Strogatz and Watts 2000), boards of directors (Newman, Strogatz and Watts 2000), and disease diffusion (Anderson and May 1991; Klovdahl 1985). These studies have examined simple connectivity but none as yet have examined the cohesivenes structures of k-connectivity.[25]

Social network researchers have traditionally focused on small, highly connected groups. Identifying connectivity as a central element of cohesion frees us from focusing on these small groups by identifying patterns through which influence or information can travel long distances. This focus on larger group connectivity might be especially important in epidemiological net-

[25] It remains to be researched as to the very different implications of connectivity for networks with power-law distributions of degree, such as websites, scientific article collaborations, biotech firm collaborations, movie actors and director collaborations, on the one hand, and for network with exponential distributions, such as kinship and marriage linkage, close friendships, etc.

works and explanations of norm formation and maintenance. The rise of electronic communication and distributed information systems suggest that distance will become less salient as information can travel through channels that are remarkably robust to degradation. By extending our vision of cohesion from small local groups to large, extended relations, we are able to capture essential elements of large-scale social organization that have only been hinted at by previous social network research, providing an empirical tool for understanding realistically sized lived communities.

Hierarchical Nesting of Connectivity Sets

A group's structural cohesion depends on the minimum number of people who unite the group. This continuous conception of cohesion implies two ways that groups can be positioned relative to each other within a population. On the one hand, we might have groups that "calve away" from the rest of the population such as those discussed by Markovsky and Lawler (1994). In such cases, cohesive groups rest "side-by-side" in the social structure, one distinct from the other. This is the kind of description researchers have for 'primary' social groups (Cooley 1912;), which we expect to find at very high levels of structural cohesion and low network distances between members. Alternatively, structurally cohesive groups could be related as Russian dolls – with increasingly cohesive groups nested inside each other. The most common such example would be a group with a highly cohesive core, surrounded by a somewhat less cohesive periphery, as has been described in areas as widely ranging as cults (Lofland and Stark 1965) and trade among nations (Smith and White 1992). A common structural pattern for large systems might be that of hierarchical nesting at low levels of connectivity and non-overlaping groups at high connectivity. The substantive high-connectivity interactions within separate cores are likely to prove important, but the low-connectivity interactions between them may be important for quite different reasons.[26]

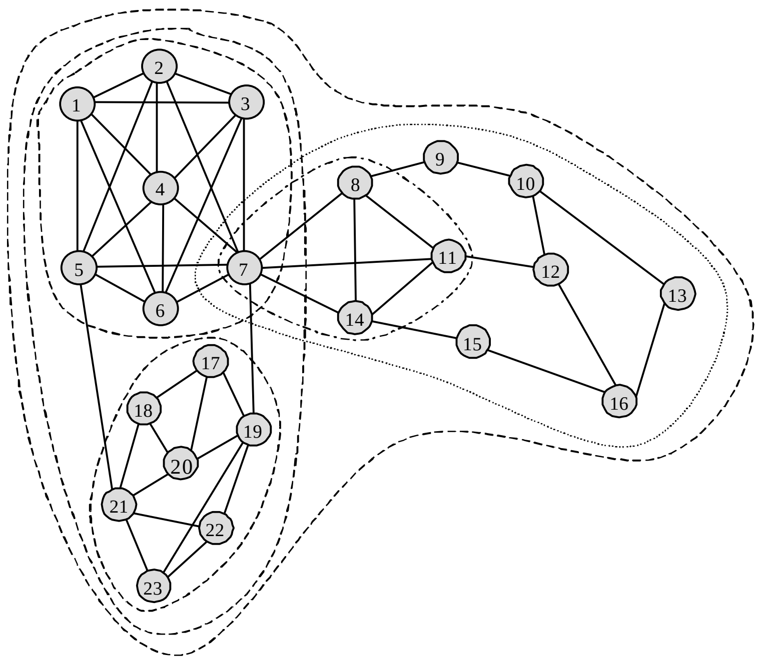
This nested conception of cohesion reflects a unique linkage between structural cohesion and social embeddedness (Granovetter 1985). The general concept of embeddedness, originally introduced in the early 1950s (see Portes and Sensenbrenner 1993 for a discussion) and reintroduced by Granovetter (1985), has had a significant influence within current sociological research and theory.[27] While used most often in research on economic activity (Baum and Oliver 1992; Portes and Sensenbrenner 1993; Uzzi 1996; 1999) or stratification (Brinton 1988), embeddedness has been used to describe social support (Pescosolido 1992), processes in health and health policy (Healy 1999; Ruef 1999), family demography (Astone et al. 1999) and the analysis of criminal networks (Baker and Faulkner 1993; McCarthy et al. 1998). While embeddedness often mixes murkily with the even broader concept of social capital, most treatments of embeddedness refer to some notion of an actor's *relative involvement depth* in social relations. If cohesive groups are nested within one another, then each successive group is more deeply embedded within the network. As such, one aspect of embeddedness — the depth of involvement in a relational structure — is captured by the extent to which a group is nested within the relational structure. To gain an intuitive sense for how cohesive blocking necessarily generates both a description of the group and a relative positional standing for the group, consider the example given in figure 2.

---

[26] Not all large systems have multiple or differentiated hubs or cores, however, as evidenced by the collaborative network of the biotechnical industry. Powell, White and Koput (n.d.) provide a theory for one set of conditions under which the core of a large network will be integrated.

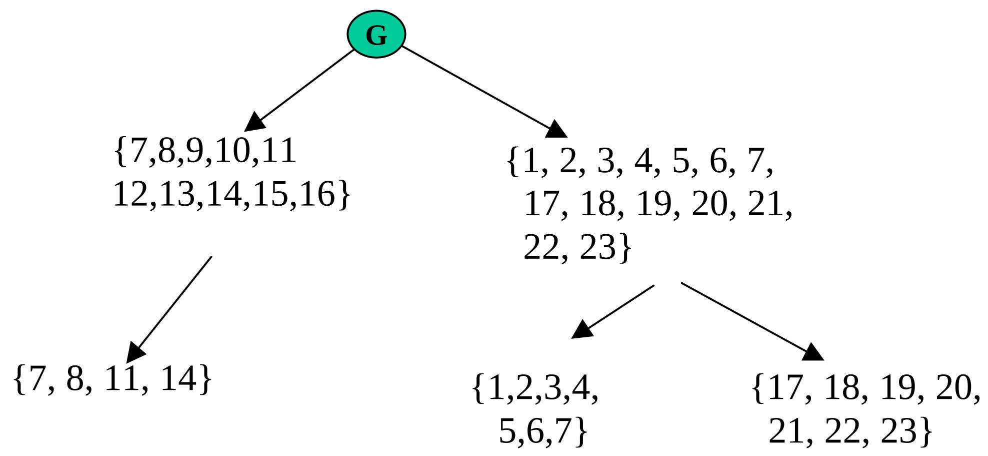
[27] At last count (March, 2001), the social citation index lists close to 1200 citations to Granovetter’s 1985 article on social embeddedness.

Figure 2. Nested Connectivity Sets



This network has a single component inclusive of all actors. Embedded within this network are two biconnected components: nodes {1 –7, 17-23} and {7-16}, with node {7} involved in both. Within the first bicomponent, however, members {1-7} form a 5-component and members {17-23} form a 3-component. Similarly, nodes {7,8,11, and 14} form another 3-component (a four-person clique) within the second bicomponent, while the remainder of the group contains no sets more strongly connected than the bicomponent. Thus, the group structure of this network is hierarchically ordered, with two large overlapping bicomponents, and internal to each one or two more cohesive subgroups, which is represented graphically in figure 3. This nesting of groups within the overall structure reflects a differential level of social embeddedness for each k-component within the *overall* network.

Figure 3. Cohesive blocking for the network in Figure 2



How does nestedness relate to other common network measures? To the extent that our measure captures the general location of actors and differentiates prominent actors, an actor's nestedness level can be thought of as a type of centrality (Freeman 1977; Wasserman and Faust 1994; see Harary et al, 1965 for a previous discussion along these lines). However, depth in the network is the outcome shared within a group whose lower size limit grows with connectivity, and as such is not attributable to an individual alone. Second, because connectivity is related to degree (each member of a *k*-component must have at least *k* relations), nestedness is necessarily correlated with degree. Third, as being on a shortest path between two actors tends to correlate with degree, betweenness centrality and nestedness will correlate as well. As we show in the empirical examples below, however, connectivity level is not at all *equivalent* to any of these indices of centrality, either singly or in combination, and measures something very different.

As a dimension of network embeddedness, nestedness has some desirable properties. Many of the previous discussions of embeddedness focus on an actor's multiplex involvement with other actors, highlighting that such actors are not free to form relations with others at will.[28]

---

[28] There is often a problem here analogous to definitions of "cohesion" that combine relational and psychological dimensions of solidarity. We would prefer a concept of embeddedness that lets us test whether a particular pattern of relations constrains decision making and actions, instead of defining embeddedness in terms of the resulting constraint.

Intuitively, embedded actors are tightly bound to others and deeply involved in the social network. Since our intuitive understanding of embeddedness refers to how difficult it would be to extract a set of actors from the wider network, our conception of nestedness provides a direct link between this dimension of embeddedness and the measure we use.

From the properties of connectivity sets, it is clear that the analysis of structural cohesion provides not only a method for identifying groups, but a method for positioning groups relative to each other with respect to depth of involvement in the social network. Because this method provides the ability to both identify cohesive groups and identify the position of each group in the overall structure, we term this method *cohesive blocking*. It is important to note the flexibility of this approach. The concept of cohesion presented here provides a way of ordering groups within hierarchically nested trees, with traditional segmented groups occupying separate branches of the cohesion structure (recall figure 3), but also with some overlap between groups in different branches (e.g., node 7 in the figure). The ability to accommodate both nested and segmented structures within a common frame is a strength of our model. The method distinguishes topological features related to connectivity from volume or distance or other characteristics and in so doing allows researchers the opportunity to distinguish among their effects.

A Method for Cohesive Blocking

How might one identify nested cut-sets empirically? By combining well known algorithms from computer science (Even and Tarjan 1975, Ball and Provan 1983, Kanevsky 1993, 1990), we can identify cut-sets in a network as follows:

Table 1. Cohesive Blocking Procedure for Identifying Connectivity Sets in a Graph

1. Identify the connectivity, *k*, of the input graph.
2. Identify all *k*-cut-sets at the current level of connectivity.
3. Generate new graph components based on the removal of these cut-sets (nodes in the cut-set belong to both sides of the induced cut.)
4. If the graph is neither complete nor trivial, return to (1), else end.

Details of the algorithm (Authors) are found in the Appendix.

Because cut-sets induce nested groups, one can describe connectivity sets not only by the resulting connectivity of the induced subgroup, but also by the level at which the group appears. That is, each group was uncovered only after a search process that broke the network at its most vulnerable points (identified the minimum cut-sets). Subgroups that appear at lower depths in the hierarchy are more deeply nested within the overall network, providing a direct measure of embeddedness within a collectivity.

Sets of actors who remain connected to each other through multiple cuts are insulated from perturbations at the edge of the social structure. They are deeply connected, and re-connected, through the multiple independent paths that define their cohesive group. This deep connectivity nicely captures the intuitive sense of being involved in relations that are, in direct contrast to "arms-length" relations, embedded in a social network (Uzzi, 1996). Nodes can also belong to multiple small-worlds as defined by cohesive blocks and their potential for overlap. We define an actor's nestedness in a social network as *the deepest cut-set level within which a given actor resides*. An actor's deepest point in a given network follows from the number of cuts needed to reach a point where no further cuts can be made.

Two Empirical Examples: High Schools and Interlocking Directorates

To demonstrate the importance of cohesive blocking for behavior in empirical settings we use data from two different types of networks. First, we use data on friendships among high school students taken from the National Longitudinal Survey of Adolescent Health (*Add Health*). This example illustrates how cohesive groups can be identified in large settings based on friendship, one of the most commonly studied network relations. The second example uses data on the interlocking directorate networks of 57 large firms in the United States (Mizruchi 1992). Since business solidarity has been an important topic of research on interlocks, we apply our method to this network and show how our conception of cohesion relates to political activity similarity. Of course, there is not space here to treat the subtle theoretical issues surrounding each of these substantive areas. Instead, the analyses below are designed to highlight how our concept of cohesion can add to empirical research in widely differing research settings.

Structural Cohesion in Adolescent Friendship Networks

Add Health is a school-based study of adolescents in grades 7-12. A stratified nationally representative sample of all public and private high schools (defined as schools with an 11th grade) in the United States with a minimum enrollment of 30 students was drawn from the Quality Education Database (QED) in April, 1994 (Bearman et al. 1996). Network data were collected by providing each student with a copy of the roster of all students for their school. Students identified up to five male and five female (10 total) friends from this roster.[29] For this paper, we use data on over 4000 students taken from a dozen schools with between 200 and 500 students (mean = 349) providing a diverse collection of public (83%) and private schools from across the United States.[30]

Nestedness and School Attachment

---

[29] The maximum number of nominations allowed was 10, but this restriction affected few students (3.1%). Mean out-degree is 4.15 with a standard deviation of 3.02. For purposes of identifying connectivity sets, we treat the graph as undirected, the algorithms needed for identifying connectivity can be modified to handle asymmetric ties. It was for directed graphs that Harary et al. (1965) developed their concept of cohesion as connectivity, although they offered no computational algorithms.

[30] This represents all schools in the dataset of this size. The selected size provides a nice balance between computational complexity and social complexity, as the schools are large enough to be socially differentiated and small enough for group identification to be carried out in a reasonable amount of time.

For each school, we employed the cohesive blocking procedure described in table 1 to identify all connectivity sets for each school friendship network. At the first level, we have the entire graph, which is usually unconnected (due to the presence of a small number of isolates). Most of the students in every school are contained within the giant bicomponent, and often within the largest tricomponent. As the procedure continues, smaller and more tightly connected groups are identified. At high levels of connectivity (k>5), identified subgroups rarely overlap. This implies a setting with multiple cores, differentially embedded in the overall school network. On the whole, friends in schools tend to be strongly connected, with a small number of students on the fringe of the social structure, the bulk of the students linked within a wide moderately cohesive (*k*=3 to 5) school network, and a small number of very tightly cohesive, deeply embedded groups (*k*=6 to 7), forming multiple cores within the wider school network.

When no further cuts can be made within a group, we have reached the end of the nesting process for that particular set of nodes. The level at which this cutting ceases describes the nestedness for each member of that group. An example of the nesting sequence for a school network of 401 nodes is given in figure 4. Here we see that most of the nodes are multiply connected at level 3, but a small 4-component occupies a separate stream of the structure. The highest levels of connectivity in this particular school are made up of a 6-component and a 7-component found at levels 55 and 56 respectively.

======
Figure 4 about here
======

Nestedness within the community should be reflected in a student's perception of his or her place in the school. The Add Health in-school survey asks students to report on how much they like their school, how close they feel to others in the school and how much they feel a part of the school.[31] Here, we use the mean of the three items as a measure of school attachment. If students correctly perceived their place in the overall school network, and if the hierarchical nesting captures this embeddedness, then there ought to be a significant positive relation between nestedness in the network and school attachment, net of any other factor that might be associated with school attachment.

Other variables that might affect a student's attachment to school include individual demographic and behavior characteristics and other features of the school friendship network. Since gender differences in school performance and school climate are well known (Stockard and Mayberry 1992), we would expect female students to have lower school attachment than male students. As students age we would expect the school to become a less salient focus of their activities, and grade in school is also controlled.[32] Students who perform well in school or who are involved in many extra-curricular activities should feel more comfortable in schools. Since students from small schools might feel more attached than students from large schools, we test a school level effect of size on the mean school attachment.

A significant feature of our approach is that we can differentiate the unique effects of network topological features that are often conflated with cohesion in standard network measures. First, the number of contacts a person has (degree centrality) reflects their level of involvement in a network. Substantively, we expect that those people with many friends in school are more likely to feel an integrated part of the school. Second, we might argue that having friends who are all friends with each other is an important feature of network involvement. As such, the density of one's personal (local) network is tested. Third, we might expect that those people who are most central in the network should have a greater sense of school attachment. Hence, between-

---

31 These are three items from the Perceived Cohesion Scale (Bollen and Hoyle 1990). The other three items used for Bollen and Hoyle’s scale were not included in the Add Health school survey.

32 Since school friendships tend to form within grade, controlling for grade in schools captures an important focal feature of the in-school network.

ness centrality is tested. Finally, it may be the case that the lived community of interest for any student is that set of students with whom they interact most often. We used NEGOPY to identify density based interaction groups within the school, and use the relative group density to measure this effect.[33] As such, membership in a relative density based friendship group could account for one's sense of attachment to school. If our conception of nestedness captures a unique dimension of network embeddedness, as our discussion above implies, then controlling for each of these features, we would expect to find a unique independent effect of nestedness on school attachment. We use a 2-level hierarchical linear model to test for these relations, with students nested in schools. The model was specified to allow coefficients to vary randomly across schools, with the school level intercept (substantively, mean school attachment) regressed on school size.

Table 2 presents the HLM coefficients for models of school attachment on the nestedness level, school activity, demographic, and other network factors. Model 1 presents a baseline model containing only attribute and school variables. As expected, females and students in higher grades tend to have lower school attachment, while students who are involved in many extracurricular activities or who get good grades feel more attached to the school. School size, while negative, is not statistically significant. In model 2, our measure of network nestedness is added to the model.[34] We see that there is a strong positive relation between nestedness and school attachment (note that the size of the standardized coefficient for nestedness is the largest in this model). Testing the difference in the deviance scores between model 1 and model 2 suggests that including nestedness improves the fit of the overall model. In models 3 - 6, we test the speci-

---

[33] Thanks to an anonymous reviewer for suggesting this specification, which uses fewer degrees of freedom than an alternative test which uses a dummy variable for each identified sub-group in the school. Tables with the alternative specification are available upon request and show no substantive difference in the nestedness effect.

[34] In addition to the nestedness level, we also tested a model using the largest *k*-connectivity value for each student. The results are very similar. Students involved in high-cohesion groups had higher levels of school attachment. All statistical significance patterns were the same for both models, and the relative values (as measured by the standardized coefficients) were similar.

fication including our measure and each of the four alternative network measures. In each case nestedness remains positive, significant and strong, while inclusion of the alternative measures adds little explanatory power (as seen by testing against model 2). In model 7, we include all potentially confounding network variables, and the relation between nestedness and attachment remains. The largest change in the coefficient for nestedness comes with the addition of degree, which is likely due to collinearity, as every member of a k-component must have degree $\geq$ k.

======
Table 2 about here
======

These findings suggest that individuals are differentially attached to the school as a whole, and thus the school differentially united, through structural cohesion. This finding holds net of school-level differences in school attachment, the number of friends people have, the interaction densities among their immediate friends or of their larger density-based interaction group, and their betweenness centrality level. That these other factors do not continue to contribute to school attachment implies a *unique* effects of structural cohesion, that would have been wrongly attributed to the other measures of network structure.

Table 2. HLM models of school attachment on network embeddedness.
(HLM Coefficients, Std. Errors, Standardized Coefficients[a])

	Mod 1	Mod 2	Mod 3	Mod 4	Mod 5	Mod 6	Mod 7
Intercept	3.85***	3.429***	3.361***	3.467***	3.354***	3.322***	3.379***
	(0.285)	(0.332)	(0.327)	(0.328)	(0.325)	(0.327)	(0.062)
School size[b]	-0.021	-0.050	-0.038	-0.064	-0.043	-0.020	-0.058
	(0.063)	(0.068)	(0.067)	(0.068)	(0.070)	(0.066)	(0.062)
	-0.015	-0.037	-0.027	-0.045	-0.030	-0.015	-0.041
Female (S2)	-0.189***	-0.148***	-0.155**	-0.149*	-0.153**	-0.149**	-0.161**
	(0.04)	(0.001)	(0.041)	(0.042)	(0.042)	(0.043)	(0.044)
	-0.10	-0.075	-0.079	-0.076	-0.078	-0.076	-0.082
Grade in school (S3)	-0.077***	-0.048*	-0.048*	-0.079*	-0.050*	-0.046	-0.051*
	(0.018)	(0.022)	(0.021)	(0.021)	(0.021)	(0.022)	(0.022)
	-0.097	-0.059	-0.060	-0.060	-0.062	-0.059	-0.063
Grade point Average (GPA)	0.132***	0.099***	0.095**	0.102**	0.104***	0.099**	0.099**
	(0.025)	(0.022)	(0.023)	(0.022)	(0.022)	(0.023)	(0.023)
	0.102	0.077	0.074	0.079	0.081	0.077	0.077
Extracurricular activities	0.115***	0.076***	0.078***	0.076***	0.075***	0.076***	0.076***
	(0.016)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.015)
	0.229	0.152	0.156	0.152	0.149	0.151	0.152
Nestedness		0.016***	0.016***	0.016***	0.011***	0.017***	0.011**
		(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.003)
		0.279	0.277	0.268	0.201	0.283	0.181
Local density			0.002				0.002
			(0.001)				(0.001)
			0.038				0.046
Betweenness centrality				0.030			-0.021
				(0.027)			(0.043)
				0.023			-0.016
Number of friends (degree)					0.021*		0.029
					(0.010)		(0.014)
					0.087		0.118
Relative Density Groups						0.0005	-0.001
						(0.002)	(0.002)
						0.006	-0.014
Deviance	9842.31	9577.57***	9578.29	9581.13	9571.06	9587.60	9574.31
N_1= 3606, N_2= 12							

* p ≤ .05, ** p ≤ .01, *** p ≤ .001
b) HLM V does not provide standardized coefficients. The reported standardized coefficients were calculate by replicating the model on a dataset where all variables were transformed to have a mean of zero and standard deviation of 1.
a) School level coefficient

Cohesion among Large American Businesses

A long-standing research tradition in inter firm networks has focused on the interlocking directorates of large firms (Mizruchi 1982; Mizruchi 1992; Palmer et al. 1986; Roy 1983; Roy and Bonacich 1988; Useem 1984). An important question in this literature, "at the core of the debate over the extent to which American society is democratic" (Mizruchi, 1992 p.32), is to what extent business in the U.S. is unified and if so, whether it is collusive. If businesses collude in the political sphere, then democracy is threatened. Yet, much of the literature has been vague in defining exactly what constitutes business unity, and thus empirical determination of the extent and effect of business unity (and possible collusion) is hard to identify.

Without treading on the issue of collusion *per se*, we approach the question of business unity as a problem of testing the predicted effects of structural cohesion. The extent to which a set of organizations in the business community is structurally cohesive ought to be a precursor or necessariy condition of its potential for unification through co-ordinated behavior. That is, since structural cohesion facilitates the flow of information and influence, coordinated action, and thus political activity, ought to be more similar among pairs of firms that are cohesively linked. Mizruchi (1992) makes this argument well, and highlights the importance of financial institutions for unifying business activity. He identifies the number of indirect interlocks between two firms as "...the number of banks and insurance companies that have direct interlocks with both manufacturing firms in the dyad" (p.108). Using data on large manufacturing firms, we identify the cohesive group structure based on indirect interlocks and relate this structure to similarities in political action.

The sample Mizruchi constructed consists of 57 of the largest manufacturing firms drawn from "the twenty major manufacturing industries in the U.S. Census Bureau's Standard Industrial Classification Scheme" in 1980 (Mizruchi 1992, p.91). In addition to data on directorship structure, he collected data on industry involvement, common stockholding, governmental regulations and political activity. The question of interest is whether the structure of relations among firms affects the similarity of their behavior. To explore whether firms that are similarly embedded also make similar political contributions, Mizruchi constructs a dyad-level political contribution similarity score.[35] He models this pair-level similarity as a function of geographic proximity, industry, financial interdependence, government regulations, and interlock structure.

Following Mizruchi, we focus on the pattern of structural cohesion based on interlocks created by banks and financial firms. Most firms in the network are involved in a strongly cohesive group, with 51 of the 57 firms members of the giant bicomponent. The nestedness structure consists of a single hierarchy that is 19 layers deep, and at the lowest level (at which no further minimum cuts can be made which would not isolate all nodes), 28 firms are members of a 14-connected component (the strongest k-component in the graph).

Does joint membership in more deeply nested subsets lead to greater political action similarity? To answer this question, we add an indicator for the deepest layer within which both firms in a dyad are nested. Thus, if firm *i* is a member of the 2nd layer but not the third, and firm *j* is a member of the 4th layer but not the fifth, the dyad is coded as being nested in the 2nd layer. As with the prior school example, we control for other network features. Table 3 presents the results of this model.

[35] The score is calculated as $S_{ij} = n_{ij} / \sqrt{n_i n_j}$, where S_{ij} = the similarity score, n_{ij} equals the number of common campaign contributions, and n_i and n_j equal the number of contributions firm *i* and *j* make respectively. The dyad level analysis is based on 1596 firm dyads.

Table 3. QAP Regression of political action similarity on dyad attributes
(Standardized Coefficients in parentheses)

Variable	Variable Description	1	2	3	4	5
Proximity	Headquarters located in same state	.017 (.043)	.013 (.032)	.013 (.034)	.015 (.039)	.015 (.039)
Same primary industry	Same primary two-digit industry	.012 (.024)	.017 (.034)	.017 (.036)	.016 (.034)	.016 (.034)
Common industry	Number of common two-digit industries	-.004 (-.008)	-.007 (-.015)	-.007 (-.015)	-.003 (-.006)	-.003 (-.006)
Market constraint	Interdependence based on transactions and concentration	.011+ (.098)	.009+ (.080)	.009+ (.082)	.009+ (.083)	.009+ (.083)
Common stockholders	Financial institutions that hold stock in both firms	.034* (.213)	.029* (.182)	.028* (.174)	.029* (.181)	.029* (.181)
Direct interlocks	Board of directors overlaps between firms	.021+ (.047)	.016 (.036)	.017+ (.037)	.018+ (.041)	.018+ (.041)
Regulated industries	Primary membership in regulated industry	.036+ (.115)	.034+ (.107)	.032 (.102)	.030 (.096)	.030 (.096)
Defense contracts	Common recipient of defense contracts	.084** (.170)	.083** (.166)	.082* (.165)	.082* (.166)	.082 (.165)
Indirect interlocks	Financial institutions with which firms interlock	.026** (.178)	.010** (.070)	.009 (.060)	.007 (.050)	.007 (.051)
Nestedness level	Level of embeddedness in the indirect interlock network		.004* (.201)	.005* (.257)	.004* (.203)	.004+ (.202)
Nestedness level	Level of embeddedness in the indirect interlock network		.004* (.201)	.005* (.257)	.004* (.203)	.004+ (.202)
Degree difference	Absolute difference in degree			.001 (.087)		-0.00 (-0.00)
Centrality difference	Absolute difference in betweenness centrality				.010+ (.111)	.010 (.111)
Constant		.171**	.156**	.137**	.144**	.144**
R-square		.195	.217	.222	.229	.228

+ p ≤.10, * p ≤.05, ** p ≤.01

In column one, we replicate the analysis presented in Mizruchi (1992), and in the remaining models we present findings with additional network indicators.[36] In the baseline model, we find that the more financial stockholders two firms have in common the greater the similarity of their political contributions. Additionally, indirect interlocks through financial institutions or jointly receiving defense contracts leads to similarity of political action. In model 2, we add the nestedness measure.[37] Net of the effects identified in model 1, we find a strong positive impact of cohesion within the indirect interlock network. As in the school networks, we test for the potentially confounding effects of degree and centrality.[38] No effect of network degree is evident, but betweenness centrality does evidence a moderate association with political similarity. When both variables are entered into the model, the statistical significance of nestedness drops slightly, but the magnitude of the effect remains constant. Based on the standardized coefficient values, nestedness has the strongest effect in each of the models 2-5.

The more deeply nested a given dyad is in the overall network structure, the more similar their political contributions. It is important to point out that this holds net of the adjacency of two firms, that is, the nestedness measure of structural cohesion has effects through indirect ties as a significant predictor of political similarity, in addition to the effect of direct adjacency created through a financial interlock.

Mizruchi identifies two potential explanations for the importance of financial interlocking on political behavior. Following Mintz and Schwartz (1985), banks and financial institutions may exercise control of firms by sitting on their boards. As such, two firms that share many such

---

36 Following Mizruchi (1992, p.121) we use the nonparametric quadratic assignment procedure (QAP) to assess the significance level of the regression coefficients. See Mizruchi for measurement details.

37 If instead of the joint nestedness level, we use the connectivity level (*k*) for the highest *k* both members are involved in, we find substantively similar results. The statistical significance of the connectivity level is slightly lower than the embeddedness level.

38 We cannot test for density-based subgroup effects, because NEGOPY assigns all members to the same group. This is a result of the high average degree within this network.

financial ties face many of the same influencing pressures and therefore behave similarly. A second argument, building on the debate surrounding structural equivalence and cohesion (Burt 1978), is that actors in similar network positions (i.e. with similar patterns of ties to similar third parties) ought to behave similarly. As in our argument for structural cohesion, Friedkin (1984) argues that influence travels through multiple steps, and thus has an effect beyond the direct link between two actors. His argument is supported and amplified by our finding that the multiple, independent paths which link pairs of structurally cohesive actors (as measured by joint nestedness) help transfer information among firms in a way that is able to coordinate politically similar activity.

Conclusion and Discussion

Social solidarity is a central concept in sociology. We have argued that solidarity can be analytically divided into two components, an ideational component and a relational component. We have defined structural cohesion as a measure of the relational component. We can further divide relational solidarity into weak and strong relational patterns. Understanding how collections of actors are linked together, how the interconnections among these actors change, and what influence involvement in cohesive groups plays in the lives of actors and organizations depends on a clear conceptualization of cohesion, as is provided by node connectivity.

When does cohesion start? Following authors such as Markovsky and Lawler, we argue that cohesion starts, weakly, when every actor can reach every other actor through at least a single path. The paths that link actors are the relational glue holding them together. We show that structural cohesion *scales* in that it is weakest when there is one path connecting actors, stronger when there are two, stronger yet with three, and finally when, for the n actors, there are almost as many (n-1) independent paths between them. As such, we have identified an essential dimension upon which structural cohesion rests. Namely, that for a group to be cohesive it cannot be easily separated. Thus, the essential substantive feature of a strongly cohesive group is that it has a status beyond any individual group member. We operationalize this conception of a social cohesion through the graph theoretic property of connectivity (Harary et al. 1965, White 1998), showing that structural cohesion increases with each additional independent path in a network.

When structural cohesion reaches its maximum of n-1 independent paths for n actors, every person is directly connected to every other in a clique: The distances between the actors shinks to the absolute minimum of 1. Yet the connectivity measure may be considered strong for groups that are very different than cliques when the group is large and the distances between members may be quite large. As distinct from many measures of cohesion, a connectivity of 10 in a group of 1000 may, in the case where links are costly to actors, be considered highly cohesive even if the average geodesic distance between members is quite large. A group with connectivity 10 is difficult to break apart by removal of sets of actors, and has at least ten independent paths holding every pair actors together regardless of the distance between them.

Many methods designed to identify cohesive groups, moreover, cannot distinguish weakly from strongly cohesive structures. Thus they cannot distinguish whether groups are at risk to disconnection by the unilateral action of their members. While such groups may have social significance, the operative feature of such groups may not be structural cohesion, but instead a unique result of social distance, number of partners, local density or centralization with strong leadership. By identifying a unique dimension upon which structural cohesion rests, we provide the ability to empirically separate the relative effect of *structural cohesion* from the other properties of networks that may be socially significant. Our empirical results in schools and corporate networks suggest that each of these dimensions may play a unique role in understanding behavior, and thus any method that conflates these effects is theoretically ambiguous.

Our conceptualization of structural cohesion simultaneously provides an operationalization of one dimension of network embeddedness. Cohesive sets in a network are nested, such that highly cohesive groups are nested within less cohesive groups. Since the process for identifying the nested connectivity sets is based on identifying the most fragile points in a network, those actors who are involved in the most highly connected portions of the network are often deeply insulated from perturbations in the overall network. Given the importance of the generalized concept of embeddedness in sociological work, a direct measure of this dimension of embeddedness is an important asset that will help provide clear empirical examinations of the relation of embeddedness to social outcomes.

Our presentation of structural cohesion has focused on the basic topological features of social cohesion, without respect to the particular features that might be relevant in any given case. We suspect that researchers may modify aspects of our conception of cohesion as theory dictates. Thus, in settings where flows that degrade quickly are of primary interest, one could account for the level of cohesion by incorporating a measure of the length of the paths or the strength of the ties. We caution, however, that much of the theoretical power of our conception of cohesion rests on the idea that multiple, indirect paths, routed through strongly cohesive subgroup, can magnify signals such that long distances can be united through social connections. Additionally, while we expect that cohesive groups will also be stable groups, we argue that this is an empirical question that can only be answered in particular settings. Finally, in a companion paper, White and Harary (2001) integrate network density and edge-connectivity with node-connectivity to provide a ratio level measure of cohesion that further distinguishes between levels of cohesion *within* given levels of *k*-connectivity.

In the present paper, we have applied our method to two very different groups in an effort to show how cohesion might be profitably used in very different empirical settings. Due to space limitations, subtle theoretical questions related to school attachment and political action have been ignored. Many potential questions, concerning interactions between embeddedness and type of actor for example, would be profitable to explore. The settings of study presented here are clearly only a limited subset of the types of relations through which cohesion effects might be important, and further research is required to understand how type and strength of relations affect the importance of cohesive structures for substantive outcome. Clearly, one could identify many more such applications in areas such as personal friendship networks (Cohen 1979-1980; Fischer 1982; Giordano et al. 1998), international networks (Rossem 1996; Smith and White 1992), kinship networks (Hage and Harary 1983; Schweizer and White 1998; Houseman and White 1998), community and regional integration (White et al. in press), social class integration (Brudner and White 1997), the state (White et al., 1999), disease and information diffusion (Anderson and May 1991; Klovdahl 1985; Morris 1993), or social influence (Friedkin 1998), just to name a few. Our hope is that by providing a clear and concise definition and operationalization of structural cohesion, researchers in all such fields will better be able to conduct their work.

Bibliography

Alba, R. D. 1973. "A Graph-Theoretic Definition of a Sociometric Clique." *Journal of Mathematical Sociology* 3:113-26.

Albert, Reka, Hawoon. Jeong, and Albert-László Barabási. 1999. Diameter of the World Wide Web.*Nature* 401: 130-131.http://www.nd.edu/~networks/Papers/401130A0.pdf

Anderson, R. M. and R. M. May. 1991. *Infectious Diseases of Humans*. Oxford University Press.

Astone, N., C. Nathanson, R. Schoen, and Y. Kim. 1999. "Family Demography, Social Theory, and Investment in Social Capital." *Population and Development Review* 25:1-31.

Auletta, V., Ye. Dinitz, Z. Nutov, and D. Parente. 1999. "A 2-Approximation Algorithm for Finding an Optimum 3-Vertex Connected Spanning Subgraph." *Journal of Algorithms* 32:21-30.

Baker, W. E. and R. R. Faulkner. 1993. "The Social Organization of Conspiracy: Illegal Networks in the Heavy Electrical Equipment Industry." *American Sociological Review* 58:837-60.

Barabási, Albert-László and Réka Albert. 1999. The emergence of scaling in random networks. *Science* 286:509-512 . http://www.nd.edu/~networks/papers/

Ball, M. O. and J. S. Provan. 1983. "Calculating Bounds on Reachability and Connectedness in Stochastic Networks." *Networks* 13:253-78.

Baum, J. A. and Christine Oliver. 1992. "Institutional Embeddedness and the Dynamics of Organizational Populations." *American Sociological Review* 57:540-559.

Bearman, P.S., J. Jones, and J.R. Udry. 1996. “Connections Count: The Add Health Design” http://www.cpc.unc.edu/projects/addhealth/design.html

Blau, P. M. 1964. *Exchange and Power in Social Life*. New York: Wiley.

Bollen, K. and R. H. Hoyle. 1990. "Perceived Cohesion: A Conceptual and Empirical Examination." *Social Forces* 69:479-504.

Brinton, M. C. 1988. "The Social-Institutional Bases of Gender Stratification: Japan As an Illustrative Case." *American Journal of Sociology* 94:515-35.

Brudner, L. A. and D. R. White. 1997. "Class, Poverty, and Structural Endogamy: Visualizing Networked Histories." *Theory and Society* 26:161-208.

Burt, R. S. 1978. "Cohesion Versus Structural Equivalence As a Basis for Network Sub-Groups." *Sociological Methods and Research* 7:189-212.

———. 1990. "Detecting Role Equivalence." *Social Networks* 12:83-97.

———. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge, Massachusetts: Harvard University Press.

Carron, A. V. 1982. "Cohesiveness in Sport Groups: Implications and Considerations." *Journal of Sport Psychology* 4:123-38.

Cohen, J. 1979-1980. "Socio-Economic Status and High-School Friendship Choice: Elmtown's Youth Revisited." *Social Networks* 2:65-74.

Cook, K. S., R. M. Emerson, M. R. Gillmore, and T. Yamagishi. 1983. "The Distribution of Power in Exchange Networks: Theory and Experimental Evidence." *American Journal of Sociology*:275-305.

Cooley, C. H. 1912. *Social Organization: A study of the Larger Mind.* New York: Charles Scribner's Sons

Doreian, P. and T. Fararo. 1998. *The Problem of Solidarity: Theories and Models*. Canada: Gordon and Breach Publishers.

Durkheim, E. 1984. *The Division of Labor in Society*. Translator W. D. Halls. New York: The Free Press.

Emirbayer, M. 1997. "Manifesto for Relational Sociology." *American Journal of Sociology* 103:281-317.

Even, S. and Endre Tarjan. 1975. "Network Flow and Testing Graph Connectivity." *SIAM Journal of Computing* 4:507-18.

Fershtman, M. 1997. "Cohesive Group Detection in a Social Network by the Segregation Matrix Index." *Social Networks* 19:193-207.

Festinger, L., S. Schachter, and K. Back. 1950. *Social Pressures in Informal Groups: A Study of Human Factors in Housing*. Stanford, CA: Stanford University Press.

Fischer, C. S. 1982. *To Dwell Among Friends: Personal Networks in Town and City*. Chicago: University of Chicago Press.

Frank, K. A. 1995. "Identifying Cohesive Subgroups." *Social Networks* 17:27-56.

Freeman, L. C. 1977. "A Set of Measures of Centrality Based on betweenness." *Sociometry* 40:35-41.

———. 1992. "The Sociological Concept of "Group": An Empirical Test of Two Models." *American Journal of Sociology* 98:152-66.

———. 1996. "Cliques, Galois Lattices, and the Structure of Human Social Groups." *Social Networks* 18:173-187.

Friedkin, N. E. 1984. "Structural Cohesion and Equivalence Explanations of Social Homogeneity." *Sociological Methods and Research* 12:235-61.

———. 1998. *A Structural Theory of Social Influence*. Cambridge: Cambridge.

Fussell, D. S., V. Ramachandran, and R. Thurimella. 1993. "Finding Triconnected Components by Local Replacement." *SIAM Journal of Computing* 22:587-616.

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

Giordano, P. C., S. A. Cernkovich, H. T. Groat, M. D. Pugh, and S. Swinford. 1998. "The Quality of Adolescent Friendships: Long Term Effects?" *Journal of Health and Social Behavior* 39:55-71.

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

———. 1985. "Economic Action and Social Structure: The Problem of Embeddedness." *American Journal of Sociology* 91:481-510.

Gross, N. and W. E. Martin. 1952. "On Group Cohesiveness." *American Journal of Sociology* 52:546-54.

Hage, P. and F. Harary. 1983. *Structural Models in Anthropology*. Cambridge: Cambridge University Press.

———. 1996. *Island Networks: Communication, Kinship, and Classification Structure in Oceania*. New York: Cambridge University Press.

Harary, F. 1969. *Graph Theory*. Reading, Massachusetts: Addison-Wesley.

Harary, F., R. Z. Norman, and D. Cartwright. 1965. *Structural Models: An Introduction to the Theory of Directed Graphs*. New York: John Wiley & Sons, Inc.

Healy, K. 1999. "The Emergence of HIV in the US Blood Supply: Organizations, Obligation, and the Management of Uncertainty." *Theory and Society* 28:529-58.

Houseman, M. and White, D.R. 1998 "Taking Sides: Marriage Networks and Dravidian Kinship in Lowland South America" in *Transformations of Kinship*, pp. 214-243 eds. Maurice Godelier, Thomas Trautmann and F.Tjon Sie Fat, Washington D.C., Smithsonian Institution Press

Hopcroft, J. E. and R. E. Tarjan. 1973. "Dividing a Graph into Triconnected Components." *SIAM Journal of Computing* 2:135-58.

Kanevsky, A. 1990. "On the Number of Minimum Size Separating Vertex Sets in a Graph and How to Find All of Them." *Proceedings of the 1st ACM-SIAM Symposium on Discrete Algorithms* :411-21.

———. 1993. "Finding All Minimum-Size Separating Vertex Sets in a Graph." *Networks* 23:533-41.

Khuller, S. and B. Raghavachari. 1995. "Improved Approximation Algorithms for Uniform Connectivity Problems." *Proceedings of the 27th Annual ACM Symposium on the Theory of Computing* :1-10.

Klovdahl, A. S. 1985. "Social Networks and the Spread of Infectious Diseases: The AIDS Example." *Social Science Medicine* 21:1203-16.

Lawler, E. J. and J. Yoon. 1993. "Power and the Emergence of Commitment Behavior in Negotiated Exchange." *American Sociological Review* 58:465-81.

Libo, L. M. 1953. *Measuring Group Cohesiveness*. Ann Arbor: University of Michigan Institue for Social Research.

Lofland, J. and R. Stark. 1965. "Becoming a World Saver: A Theory of Conversion to a Deviant Perspective." *American Sociological Review* 30:862-75.

Lorrain, F. and H. C. White. 1971. "Structural Equivalence of Individuals in Social Networks." *Journal of Mathematical Sociology* 1:49-80.

Luce, D. 1950. "Connectivity and Generalized Cliques in Sociometric Group Structure." *Psychometrika* 15:169-90.

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

Markovsky, B. and E. J. Lawler. 1994. "A New Theory of Social Solidarity." Pp. 113-37 in *Advances in Group Processes*, vol. 11, Eds Barry Markovsky, Jodi O'Brien, and Karen Heimer. Greenwich, Conn: JAI Press.

McCarthy, B., J. Hagan, and L. E. Cohen. 1998. "Uncertainty, Cooperation and Crime: Understanding the Decision to Co-Offend." *Social Forces* 77:155-84.

McPherson, J. M. and L. Smith-Lovin. 1986. "Sex Segregation in Voluntary Associations." *American Sociological Review* 51:61-79.

Milgram, S. 1969. "The Small World Problem." *Psychology Today* 22:61-67.

Mintz, B. and M. Schwartz. 1981. "Interlocking Directorates and Interest Group Formation." *American Sociological Review* 46:781-96.

Mizruchi, M. S. 1982. *The American Corporate Network, 1904-1974*. Beverly Hills: Sage.

———. 1992. *The Structure of Corporate Political Action*. Cambridge, MA and London, England: Harvard University Press.

Moody, James. 1999. *SPAN: SAS Programs for Analyzing Networks*. Vers. .30. The Ohio State University.

Morris, M. 1993. "Epidemiology and Social Networks: Modeling Structured Diffusion." *Sociological Methods and Research* 22:99-126.

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

Newman, Mark E. J. 2001. The Structure of Scientific Collaboration Networks. *Proceedings of the National Academy of Science* 98: 404-409. http://xxx.lanl.gov/abs/cond-mat/0007214

Newman, Mark E. J., Steven H. Strogatz and Duncan J. Watts, 2000. Random graphs with arbitrary degree distributions and their applications. Submitted to *Physical Review E* 63. http://xxx.lanl.gov/abs/cond-mat/0007235

O'Reilly, C. A. I. and K. H. Roberts. 1977. "Task Group Structure, Communication and Effectiveness in Three Organizations." *Journal of Applied Psychology* 62:674-81.

Owen, W. F. 1985. "Metaphor Analysis of Cohesiveness in Small Discussion Groups." *Small Group Behavior* 16:415-24.

Palmer, D., R. Friedland, and J. V. Singh. 1986. "The Ties That Bind: Organizational and Class Bases of Stability in a Corporate Interlock Network." *American Sociological Review* 51:781-96.

Pescosolido, B. A. 1992. "Beyond Rational Choice: The Social Dynamics of How People Seek Help." *American Journal of Sociology* 97:1096-138.

Pool, I. d. S. and M. Kochen. 1978. "Contacts and Influence." *Social Networks* 1:5-51.

Portes, A. and J. Sensenbrenner. 1993. "Embeddedness and Immigration: Notes on the Social Determinants of Economic Action." *American Journal of Sociology* 98:1320-1350.

Powell, Walter W. 1990. Neither Market Nor Hierarchy: Network Forms of Organization. *Research in Organizational Behavior* 12:295-336.

Powell, Walter W. 1996. Inter-Organizational Collaboration in the Biotechnology Industry. *Journal of Institutional and Theoretical Economics* 120(1): 197-215.

Powell Walter W., Kenneth Koput, and Laurel Smith-Doerr. 1996. Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology. *Administrative Science Quarterly* 41(1): 116-45.

Powell, Walter W., Douglas R. White and Kenneth W. Koput. n.d. The Evolution of a Knowledge Industry: Network Movies and Dynamic Analyses of Biotechnology. http://eclectic.ss.uci.edu/~drwhite/Movie/

Rapoport, A. and W. J. Horvath. 1961. "A Study of a Large Sociogram." *Behavioral Science* 6:279-91.

Richards, William D. 1995. *NEGOPY*. Vers. 4.30. Brunaby, B.C. Canada: Simon Fraser University.

Rossem, R. V. 1996. "The World System Paradigm As General Theory of Development: A Cross-National Test." *American Sociological Review* 61:508-27.

Roy, W. G. 1983. "The Unfolding of the Interlocking Directorate Structure of the United States." *American Sociological Review* (248-257).

Roy, W. G. and P. Bonacich. 1988. "Interlocking Directorates and Communities of Interest Among American Railroad Companies, 1905." *American Sociological Review* 53:368-79.

Ruef, M. 1999. "Social Ontology and the Dynamics of Organizational Forms: Creating Market Actors in the Healthcare Field, 1966-1994." *Social Forces* 77:1403-32.

Schweizer, T. and D. R. White. 1998. *Kinship, Networks and Exchange*. Cambridge: Cambridge University Press.

Schwartz, Michael. 2001. *The Rise and Fall of Detroit*. Beverly Hills, Sage Publication

Simmel, G. 1950. *The Sociology of Georg Simmel*. Editor K. H. Wolf. New York: The Free Press.

Smith, D. A. and D. R. White. 1992. "Structure and Dynamics of the Global Economy: Network Analysis of International Trade 1965-1980." *Social Forces* 70:857-93.

Stockard, J. and M. Mayberry. 1992. *Effective Educational Environments*. Newbury Park, CA: Corwin Press.

Useem, M. 1984. *The Inner Circle*. New York: Oxford University Press.

Uzzi, B. 1996. "The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect." *American Sociological Review* 61:674-98.

———. 1999. "Embeddedness in the Making of Financial Capital: How Social Relations and Networks Benefit Firms Seeking Financing." *American Sociological Review* 64:481-505.

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

Watts, D. J. 1999. *Small Worlds: The Dynamics of Networks Between Order and Randomness*. Princeton: Princeton University Press.

Watts, D. J. and S. H. Strogatz. 1998. "Collective Dynamics of 'Small-World' Networks." *Nature* 393:440-442.

Weber, M. 1978. *Economy and Society*. eds. Guenther Roth and Claus Wittich. Berkeley: University of California Press.

White, D. R. 1998. "Concepts of Cohesion, Old and New: Which Are Valid Which Are Not?" University of California - Irvine. Manuscript.

White, D. R. and F. Harary. 2001 "The Cohesiveness of Blocks in Social Networks: Connectivity and Conditional Density." Manuscript. Submitted to *Sociological Methodology 2001*.

White, D. R., M. Schnegg, L. A. Brudner. 1999. "Decentralized Systems and The Invisible State: Low Density Multiconnected Cohesion in Large-Scale Social Networks in Tlaxcala, Mexico." Web publication *http://eclectic.ss.uci.edu/~drwhite/cases/decentralized.html*

White, D. R., M. Schnegg, L. A. Brudner, and H. Nutini. In Press. "Conectividad Múltiple y sus Fronteras de Integración: Compadrazgo y Parentesco en Tlaxcala" In, *Redes Sociales: Teoría y Aplicaciones*, eds. Jorge Gil and Samuel Schmidt. Mexico, D.F.: UNAM Press. Translated from "Multiple Connectivity and Its Boundaries of Reticulate Integration: A Community Study." University of California Irvine. manuscript.

White, H. C., S. A. Boorman, and R. L. Breiger. 1976. "Social Structure From Multiple Networks I." *American Journal of Sociology* 81:730-780.

Willer, D. 1999. *Network Exchange Theory*. Westport, Connecticut: Praeger.

Appendix: Cohesive Blocking Procedure for Identifying Connectivity Sets

1. Identify the connectivity, *k*, of the input graph.
2. Identify all *k*-cut-sets at the current level of connectivity.
3. Generate new graph components based on the removal of these cut-sets (nodes in the cut-set belong to both sides of the induced cut.)
4. If the graph is neither complete nor trivial, return to (1), else end.

This procedure is repeated until all nested connectivity sets have been enumerated.[39] Walking through the example in figure 2, we would first identify the bicomponents (**step 1**), and identify the cut-node {7} (**step-2**).[40] Separating the two subgraphs at node 7 (**Step 3**) induces two new components: {7-16} and {1-7,17-23} which are neither complete nor trivial. Within each induced sub-graph we repeat the process, starting by identifying the sub-graph connectivity. Within the {7-16} bicomponent, we identify {8,10}, {10,16}, and {14,16} as the 2-cuts for this sub-graph, each of which leads to a single minimum degree cut (we call these types of cuts *singleton cuts*, e.g., of 9, 13, or 15). The graph remaining after the singleton cuts have been removed is {7, 8, 11, 14, 10, 12, and 16}, which is 1- connected, with {7 8 11 and 14 the largest included tricomponent). Because {7,8,11, and 14} form a completely connected clique, we stop here and return to the other graph induced by removing node {7}, ({1-7,17-23}). Again, this graph is a bicomponent. Cut-sets {5,7} and {21,19}, {21,7} and {5,19} induce two graphs of higher cohesion: {1-7} and {17-23} that are of maximal connectivity, as further cuts will induce only singleton partitions.

One can represent the hierarchical nesting of connectivity groups as a directed tree, with the total graph as the root, and each sub-graph that derives from it a new node. The *cohesive blocking* of a network consists of identifying all cohesive substructures within the network and

---

[39] SAS IML programs for identifying the full connectivity sets of a network are available (Moody 1999).

[40] An efficient algorithm for doing so can be found in Gibbons (1985).

relating them to each other in terms of the nested branching of the subgroups. The blocking for the example above is given in figure 3 (with singleton cuts not represented).

Testing for *k*-connectivity (**Step 1**) can be accomplished with a network maximum flow algorithm developed by Even and Tarjan (1975).[41] An algorithm for identifying all *k*-cut-sets of the graph (**Step 2**) has been developed by Ball and Provan (1983), extended to finding all minimum-size separating vertex sets by Kanevsky (1993, 1990).[42] One must apply this pair of procedures for every induced sub-graph, and thus the total running time of the algorithm can be high. Two steps can be taken to reduce the computation time. First, there are linear time algorithms for identifying *k*-connected components for $k \leq 3$, and one can start searching with these algorithms, limiting the number of levels at which one has to run the full connectivity algorithms (Hopcroft and Tarjan 1973, Fussell et al. 1993). Second, in many empirical networks the most common cut-set occurs for singleton cuts. Because the procedure is nested, one can search for nodes with degree less than or equal to the connectivity of the parent graph,[43] remove them from the network, and thus apply the network flow search only after the singleton cuts have been removed.[44]

---

41 This is an extension of Dinic's algorithm and runs in $O(V^{1/2} * E^2)$ time.

42 Which runs in $O(2^k V^3)$ time.

43 The graph from which the current graph was derived.

44 Additionally, there are approximation approaches (Auletta et al. 1999; Khuller and Raghavachari 1995) that could be used to identify graph connectivity within a certain amount of error, which would be faster.