The Development of Social Network Analysis—with an Emphasis on Recent Events

Linton C. Freeman
University of California, Irvine

In a recent book I reviewed the development of social network analysis from its earliest beginnings until the late 1990s (Freeman, 2004). There I characterized social network analysis as an approach that involves four defining properties: (1) It involves the intuition that links among social actors are important. (2) It is based on the collection and analysis of data that record social relations that link actors. (3) It draws heavily on graphic imagery to reveal and display the patterning of those links. And (4) it develops mathematical and computational models to describe and explain those patterns.

In that book I reviewed both the history and the prehistory of social network analysis. I showed that as early as the thirteenth century, and probably even earlier, people began to produce work that drew on one or more of the four properties listed above. Until the 1930s, however, no one had used all four properties at the same time. The modern field of social network analysis, then, emerged in the 1930s.

In its first incarnation, modern social network analysis was introduced by a psychiatrist, Jacob L. Moreno, and a psychologist, Helen Jennings (Freeman, 2004, Chapter 3). They conducted elaborate research, first among the inmates of a prison (Moreno,1932) and later among the residents in a reform school for girls (Moreno, 1934).

Moreno and Jennings named their approach *sociometry*. At first, sociometry generated a great deal of interest, particularly among American psychologists and sociologists. But that interest turned out to be short lived; by the 1940s most American social scientists had returned to their traditional focus on the characteristics of individuals.

During the same period another group, led by an anthropologist, W. Lloyd Warner, also adopted the social networks approach (Freeman, 2004, Chapter 4). Their efforts were centered in the Anthropology Department and the Business School at Harvard, and their approach was pretty clearly independent of Moreno and Jennings work. Warner designed the "bank wiring room" study, a social network component of the famous Western Electric research on industrial productivity (Roethlisberger and Dixon, 1939). And he involved business school colleagues and anthropology students in his community research. They conducted social network research in two communities, Yankee City (Warner and Lunt, 1941) and Deep South (Davis, Gardner and Gardner, 1941).

The Warner people never stirred up as much interest as did Moreno and Jennings. And when Warner moved to the University of Chicago in 1935 and turned to other kinds of research the whole Harvard movement fell apart.

The third version of social network analysis emerged when a German psychologist, Kurt Lewin, took a job at the University of Iowa in 1936 (Freeman, 2004, pp. 66-75). There, Lewin worked with a large number of graduate students and post-docs. Together, they developed a structural perspective and conducted social network research in the field of social psychology (e. g. Lewin and Lippit, 1938).

The Lewin group moved to the Massachusetts Institute of Technology in 1945, but after Lewin's sudden death in 1947, most of the group moved again, this time to the University of Michigan. This Michigan group made important contributions to social network research for more than twenty years (e. g. Festinger and Schachter, 1950; Cartwright and Harary, 1956; Newcomb, 1961).

One of Lewin's students, Alex Bavelas, remained at MIT where he spearheaded a famous study of the impact of group structure on productivity and morale (Leavitt, 1951). This work was influential in the field of organizational behavior, but most of its influence was limited to that field.

All three of these teams began work in the 1930s. None of them, however, produced an approach that was accepted across all the social sciences in all countries; none provided a standard for structural research.

Instead, after the 1930s and until the 1970s, numerous centers of social network research appeared. Each involved a different form and a different application of the social network approach. Moreover, they worked in different social science fields and in different countries. Table 1 lists thirteen centers that emerged during those thirty years.[1]

Place	**Field**	**Team Leaders**	**Country**
Michigan State	Rural sociology	Charles P. Loomis Leo Katz	USA
Sorbonne	Linguistics	Claude Lévi-Strauss André Weil	France
Lund	Geography	Thorsten Hägerstrand	Sweden
Chicago	Mathematical Biology	Nicolas Rashevsky	USA
Columbia	Sociology	Paul Lazersfeld Robert Merton	USA
Iowa State	Communication	Everett Rogers	USA
Manchester	Sociology	Max Gluckman	Great Britain
MIT	Political Science	Ithiel de Sola Pool Manfred Kochen	USA
Syracuse	Community Power	Linton C. Freeman Morris H. Sunshine	USA
Sorbonne	Psychology	Claude Flament	France
Michigan	Sociology	Edward Laumann	USA
Chicago	Sociology	Peter Blau James A. Davis	USA
Amsterdam	Sociology	Robert Mokken	Netherlands

Table 1. Centers of Social Network Research from 1940 to 1969

By 1970, then, sixteen centers of social network research had appeared. With the development of each, knowledge and acceptance of the structural approach grew. Still, however, none of these centers succeeded in providing a generally recognized paradigm for the social network approach to social science research.

[1] Important publications from each of these centers are listed in Freeman (2004).

That all changed in the early 1970s when Harrison C. White, together with his students at Harvard, built a seventeenth center of social network research. In my book I described the impact of this group (Freeman, 2004, p. 127):

> From the beginning, White saw the broad generality of the structural paradigm, and he managed to communicate both that insight and his own enthusiasm to a whole generation of outstanding students. Certainly the majority of the published work in the field has been produced by White and his former students Once this generation started to produce, they published so much important theory and research focused on social networks that social scientists everywhere, regardless of their field, could no longer ignore the idea. By the end of the 1970s, then, social network analysis came to be universally recognized among social scientists.

Following the contributions of White and his students, social network analysis settled down, embraced a standard paradigm and became widely recognized as a field of research.

In the late 1990s, however, there was a revolutionary change in the field. It was then that physicists began publishing on social networks.[2] First, Duncan Watts and Stevan H. Strogatz (1998) wrote about small worlds. And a year later Albert-Lásló Barabási and Réka Albert (1999) examined the distribution of degree centralities. I ended the earlier account in my book by commenting on the entry of Watts, Strogatz, Barabási and Albert into social network research. I expressed the pious hope that, like all the earlier potential claimants to the field, our colleagues from physics would simply join in the collective enterprise.

That hope, however, was not immediately realized. These physicists, new to social network analysis, did not read our literature; they acted as if our sixty years of effort amounted to nothing. In a recent article, I contrasted the approach of these new physicists with that of earlier physicists who had been involved in social network research (Freeman, 2008):

> Other physicists had already been involved in social network analysis. Notable among these were Derek de Solla Price, Harrison White and Peter Killworth (*e. g.* Price, 1965, 1976; White,

[2] Scott, in the current volume, also describes the entry of physicists into social network analysis. His description centers on their theoretical perspective.

> 1970; White, Boorman and Breiger, 1976; Killworth, McCarty, Bernard, Johnsen, Domini and Shelley, 2003; Killworth, McCarty, Bernard and House, 2006). These physicists read the social network literature, joined the collective effort and contributed to an ongoing research process.

But neither Watts and Strogatz nor Barabási and Albert did any of these things. They simply took research topics that had always been part of social network analysis and claimed them as topics in physics.

The result was a good deal of irritation (and perhaps a certain amount of jealousy) on the part of many members of the social network research community. Bonacich (2004) put it this way:

> Duncan Watts and Albert-Lásló Barabási are both physicists who have recently crashed the world of social networks, arousing some resentment in the process. Both have made a splash in the wider scientific community, as attested by their publications in high status science journals (*Science*, *Nature*). . . . Both have recently written scientific best-sellers: *Six Degrees* ranks 2547 on the Amazon list, while *Linked* ranks 4003.

Watts, Strogatz, Barabási and Albert opened the door. They managed to get a huge number of their physics colleagues involved—enough to completely overwhelm the traditional social network analysts. Their impact, then, was to produce a revolution in social network research. In the present essay I will focus on that revolution and its aftermath. Here I will review the developments that have occurred since those two articles were published.

The Origins of the Revolution

The article by Watts and Strogatz (1998), addressed a standard topic in social network analysis, the "small world." Concern with that issue stemmed from one of the classic social network papers, "Contacts and influence," written by Ithiel de Sola Pool and Manfred Kochen in the mid-1950s. It circulated in typescript until 1978 when it was finally published as the lead article in Volume 1, Number 1 of the new journal, *Social Networks*.

The questions raised by Pool and Kochen concerned patterns of acquaintanceship linking pairs of persons. They speculated that any two people in the United States are linked by a chain of acquaintanceships involving no more than seven intermediaries.

Various students picked up on Pool and Kochen's ideas, including Stanley Milgram who used them as the basis for his doctoral dissertation on the "small world." Milgram published several papers on the subject, one of which one was a popularization that appeared in *Psychology Today* (1967).

Watts and Strogatz cited the *Psychology Today* article as well as a later book edited by Kochen (1989) on the small world idea. But they apparently did not discover any of the other literature on the subject. In any case, they introduced an entirely new model that was designed to account for both the clustering found in human interaction and the short paths linking pairs of individuals.

The Watts and Strogatz model begins with an attempt to capture clustering—the universal tendency of friends of friends to be friends. They represent links among individuals as a circular lattice like the one shown in Figure 1, where each node is an individual and each edge is a social link connecting two individuals. They go on to define an *average clustering coefficient* ***C(p)*** that measures the degree to which each node and its immediate neighbors are all directly linked to one another. The structure in Figure 1 embodies a good deal of clustering—neighbors of neighbors are, for the most part, neighbors—thus the clustering coefficient ***C(p)*** is high. But, at the same time, ***L(p)***, the average length of the path linking any two individuals in the whole lattice, is relatively large.

Place Figure 1 about here

Since ***L(p)*** is large, the world represented by this circular lattice is certainly not small. But Watts and Strogatz showed that they could produce a small world effect—where no individual is very far from any other individual—simply by removing just a few of the links between close neighbors and substituting links to randomly selected others. As Figure 2 shows, under those conditions some links span clear across the lattice. The result is, that as random links are substituted for links to close neighbors, path length ***L(p)*** drops abruptly, but the clustering coefficient ***C(p)*** is hardly diminished at all. Thus, for the most part, friends of friends are still friends, but the total world has become dramatically smaller.

Place Figure 2 about here

The article by Barabási and Albert (1999) also took up a standard network analytic topic, degree distribution. The degree of a node is simply the number of other nodes to which it is directly connected by edges. Much of the earliest research on social networks was focused on the distributions of degrees. Research in sociometry often involved asking people whom they would choose, say, to invite to a party or to work with on a project (Moreno, 1934). As soon as the responses to such questions began to be tallied, it became apparent that the distribution of being chosen was dramatically skewed. A few individuals were chosen extremely often while a large number were chosen rarely, if at all.

Moreno and Jennings (1938) reported two empirical results: (1) such skewed distributions were universally observed, and (2) they departed from expectations based on random choices. As they described it, "A distortion of choice distribution in favor of the more chosen as against the less chosen is characteristic of all groupings which have been sociometrically tested."

Barabási and Albert (1999) studied the distribution of connections in networks that grew as a consequence of adding new nodes. Their examples included links between sites in the World Wide Web, links between screen actors who worked together on films and links between generators, transformers and substations in the U. S. electrical power grid. Although Barabási and Albert were apparently unaware of the earlier findings of Moreno and Jennings, they discovered that the connections in the networks they examined were not random. Instead, the links were skewed; just as Moreno and Jennings had reported, Barabási and Albert found a few nodes that displayed too many connections and a great many nodes that displayed too few.

Barabási and Albert went on to propose a simple model designed to account for the pattern of skewness they had observed. Consider a collection of existing nodes. Let $\boldsymbol{k_i}$ be the number of links already established to node $\boldsymbol{i}$. Then let the probability that a new node is going to link to any node $\boldsymbol{i}$, depend on $\boldsymbol{k_i}$. The model specifies the probability of that link connecting to node $\boldsymbol{i}$ as

$\boldsymbol{P(k_i) \approx k_i^{\gamma}}$ where $2 \leq \gamma \leq 3$.[3] The distribution of connections, then, follows a power law, or as Barabási and Albert characterize it, it is "scale free."

The Growth of the Revolution

As a consequence of the interest generated by Watts and Strogatz and by Barabási and Albert, the revolution began in earnest. As Figure 3 shows, physicists followed up on the Watts and Strogatz small world paper. Within five years, the physics community had produced more small world papers than the social network community had turned out in forty-five years (Freeman, 2004, pp. 164-166).

Moreover, Figure 3 also shows that, at that point, 98% of the citations were made within either the physics community or the social network community. For the most part, physicists ignored the earlier work by social network analysts. And social network analysts responded in kind.

Place Figure 3 about here

Physicists were also quick to follow up on Barabási and Albert's work on degree distributions. According to Google Scholar their first paper had received over 4000 citations as of mid-November 2008. But practically none of those citations was produced by a social network analyst.

It soon became evident that the physicists' interest in social networks was not going to be confined to small world phenomena and degree distributions. Members of the physics community quickly began to explore other problems that had traditionally belonged to social network analysts. Nor was that interest restricted to physicists. At the same time, physicists succeeded in getting biologists and computer scientists involved their efforts. Two main foci of this new thrust involved the study of cohesive groups or what physicists call *communities* and the study of the positions that nodes occupy in a network—particularly their centrality. I will review these foci in the next two sections.

---

[3] The Barabási and Albert model, however, turns out to be essentially the same as that proposed by a social network analyst, Derek de Solla Price, in 1976.

Cohesive Groups or Communities

The notion of cohesive group is foundational in sociology. Early sociologists (Tönnies, 1855/1936; Maine, 1861/1931; Durkheim, 1893/1964; Spencer, 1897; Cooley, 1909/1962) talked about little else. Their work provided an intuitive "feel" for groups, but it did not define groups in any systematic way.

When the social network perspective emerged, however, network analysts set out to specify groups in structural terms. Freeman and Webster (1994) described the observation behind this structural perspective on groups:

> . . . whenever human association is examined, we see what can be described as thick spots—relatively unchanging clusters or collections of individuals who are linked by frequent interaction and often by sentimental ties. These are surrounded by thin areas-where interaction does occur, but tends to be less frequent and to involve very little if any sentiment.

Thus, the social ties within a cohesive group will tend to be dense; most individuals in the group will be linked to a great many other group members. Moreover, those in-group ties will tend to display clustering—where, as described above, friends of friends are friends. In contrast, relatively few social ties will link members of different groups, and clustering will be relatively rare.

An early social network analyst, George Homans (1950, p. 84) spelled out the intuitive basis for the social network conception of cohesive groups:

> . . . a group is defined by the interactions of its members. If we say that individuals A, B, C, D, E . . . form a group, this will mean that at least the following circumstances hold. Within a given period of time, A interacts more often with B, C, D, E, . . . than he does with M, N, L, O, P, . . . whom we choose to consider outsiders or members of other groups. B also interacts more often with A, C, D, E, . . . than he does with outsiders, and so on for the other members of the group. It is possible just by counting interactions to map out a group quantitatively distinct from others.

Over the years, network analysts have proposed dozens of models of cohesive groups. These models serve to define groups in structural terms and provide procedures to find groups in network data. They all try to capture something close to Homans' intuition in one way or another. Some of them represent groups in terms of on/off or binary links among actors (e. g. Luce and Perry, 1949; Mokken, 1979). Others represent them in terms of quantitative links that index the strength of ties inking pairs of actors (e. g. Sailer and Gaulin, 1984; Freeman, 1992).

Currently, then, we have a huge number of models of cohesive groups. Most of them were reviewed by Wasserman and Faust (1994). Some were algebraic (e. g. Breiger, 1974; Freeman and White, 1993), some were graph theoretic (e. g. Alba, 1973; Moody and White, 2003), some were built on probability theory (e. g. Frank, 1995; Skvoritz and Faust, 1999) and some were based on matrix permutation (Beum and Brundage, 1950; Seary and Richards, 2003). All, however, were designed to specify the properties of groups in exact terms, to uncover group structure in network data, or both.

Over the years social network analysts have also drawn on various computational algorithms in an attempt to uncover groups. These include multidimensional scaling (Freeman, Romney and Freeman, 1987; Arabie and Carroll, 1989), various versions of singular value decomposition, including principal components analysis and correspondence analysis (Levine, 1972; Roberts, 2000), hierarchical clustering (Breiger, Boorman and Arabie, 1975; Wasserman and Faust, 1994, pp. 382-383), the max-cut min-flow algorithm (Zachary, 1977, Blythe, 2006), simulated annealing (Boyd, 1991, p.223; Dekker, 2001) and the genetic algorithm (Freeman, 1993; Borgatti and Everett, 1997).

In social network research, the general tendency over the years has been to move from binary representations to representations in which the links between nodes take numeric values that represent the strengths of connections. At the same time social network analysts have gradually shifted from building algebraic and graph theoretic models to developing models grounded in probability theory. And, as time has passed, they have relied more often on the use of computational procedures to uncover groups.

A notable exception to this trend can be found in the recent article by

Moody and White (2003). There, they used graph theory to define *structural cohesion*. They defined structural cohesion "... as the minimum number of actors who, if removed from a group, would disconnect the group." Then they went on to define *embeddedness* in terms of a hierarchical nesting of cohesive structures. This approach represents a new and sophisticated version of the traditional social network model building.

Since the early 1970s, mathematicians and computer scientists had also been interested in groups or communities. They defined that interest in terms of *graph partitioning* (Fiedler, 1973, 1975; Parlett, 1980; Fiduccia and Mattheyses, 1982, Glover, 1989, 1990; Pothen, Simon and Liou, 1990). Social network analysts recognized this tradition when the work by Glover was cited and integrated into the program UCINET (Borgatti, Everett and Freeman, 1992). And in 1993 the link in the other direction was made when a team composed of an electrical engineer and a computer engineer, Wu and Leahy, cited the work of the statistician-social network analyst, Hubert (1974). And in 2000 three computer scientists, Flake, Lawrence and Giles cited the social network text by Scott (1992).

Until quite recently, however, these efforts did not stir up much interest in the physics community. Instead, the physicists turned to the procedures developed in social network analysis. Michelle Girvan and Mark Newman (2002), adapted the social network model of betweenness centrality (Freeman, 1977) to the task of uncovering groups. Their adaptation was based on the betweenness of graph edges, rather than nodes, and the result was a new algorithm for partitioning graphs.

Edge betweenness refers to the degree to which an edge in the graph falls along a shortest path linking every pair of nodes. A path in a graph is a sequence of nodes and edges beginning and ending with nodes. Girvan and Newman reasoned that since there should be relatively few edges linking individuals in different groups, those linking edges should display a high degree of betweenness. So they began by removing the edge with the highest betweenness, and continued that process until the graph was partitioned.

Two years later Newman and Girvan (2004) published a follow-up article. Their second paper again focused on edge removal, but this time they introduced an alternative model that had two intuitive foundations. In one, they showed that random walks between all pairs of nodes would determine the betweenness of edges—not just along shortest paths—but along all the paths linking pairs of nodes. The other intuition was motivated by a physical model where edges were defined as resistors that impeded the flow of current between nodes. The edge with the lowest current flow was removed. If that did not yield a partition the process was continued until partitioning did take place. These two models produced the same partitions.

Newman and Girvan went on to show that all of their algorithms always partitioned the data even though some of the partitionings might not reflect the presence of actual communities. So they introduced a measure called *modularity*. Modularity is based on the ratio of within partition ties to those that cross partition boundaries and compares that ratio to its expected value when ties are produced at random. Thus, it provides an index of the degree to which each partition embodies a group- or community-like form.

The result of the two papers by Girvan and Newman was dramatic. Both physicists and computer scientists quickly developed an interest in groups or communities. Radicchi, Castellano, Cecconi, Loreto and Parisi (2004) specified two kinds of communities. One was characterized as "strong"; it defined a partition as a community if it met the condition that every node had more within-group ties than cross-cutting ones.[4] The other they characterized as "weak". It proposed that a partition was a community if the total number of ties within each partition was greater than the total number of ties linking nodes in the partition to nodes outside the partition.

Radicci et al. also pointed out that the Girvan and Newman betweenness-based algorithm was computationally slow. So they introduced a new, more efficient, measure. They reasoned that edges that bridge between communities are likely to be involved in very few 3-cycles (where friends of friends are friends). So they based their measure on the number of 3-cycles in which each edge is involved, and they showed that their measure had moderate negative correlation with the Girvan-Newman measure. The number of 3-cycles in which an edge is involved, then, turns out to be inversely related to the betweenness of that edge.

[4] They did not cite the similar social network models introduced by Sailer and Gaulin (1984).

Newman (2004) quickly jumped back in. He, too, was troubled by the slowness of the Girvan-Newman algorithm for finding communities. So he proposed a fast "greedy" algorithm. A greedy algorithm makes the optimal choice at each step in a process, without regard to the long-term consequences of that choice.[5] In this case, Newman proposed starting a process by having each cluster contain a single node. Then, at each stage in the process, the pair of clusters that yields the highest modularity is merged.

The concern with computing speed seems to have started a race to see who could develop the fastest algorithm to cluster nodes in terms of their modularity. A computer scientist, Clauset, working with two physicists, Newman and Moor (2004) were able to speed up Newman's "greedy" algorithm. Two more computer scientists, Duch and Arenas (2005), devised an algorithm to speed it up even more. And in 2006 Newman showed how to gain still more speed by applying singular value decomposition to the modularity matrix. Then, in 2007, a computer scientist, Djidjev, developed a still faster algorithm for constructing partitions based on modularities.

Continuing the search for speed, two other computer scientists, Pons and Latapy (2006) took an entirely different approach. They reasoned that since communities are clusters of densely linked nodes that are only sparsely linked together, a short (2 or 3 step) random walk should typically stay within the community in which it is started. They proposed an algorithm that began with a series of randomly selected starter nodes. Then each starter is used to generate a random walk. Then the starter, along with the nodes that are reached, are tallied as linked. The likelihood is that once these results are cumulated, they will display the clustered communities. And finally, two industrial engineers and a physicist, Raghavan, Albert and Kumara (2007) produced a very fast algorithm based on graph coloring. Nodes begin with unique colors, then, iteratively, acquire the color of the majority of their immediate neighbors.

Other, quite different, procedures were also introduced. A physicist and a computer scientist, Wu and Huberman (2004), developed a model based on assuming edges are resistors, as was the case in the earlier model introduced by Newman and Girvan. But Wu and Haberman's model turns out to be much

[5] Hierarchical clustering is an example of a greedy algorithm.

more complicated and *ad hoc*. Four physicists, Capocci, Servedio, Caldarelli and Colaiori (2004) suggested using singular value decomposition to uncover communities. And three others, Fortunato, Latora and Marchiori (2004) proposed a variation of edge centrality, called "information centrality." Their centrality is based on the inverse of the shortest path length connecting each pair of nodes. Physicists Palla, Derényi, Farkas and Vicsek (2005) defined communities as cliques and focused on patterning of clique overlap. Reichardt and Bornholdt (2006) used simulated annealing to search for partitions that yield communities that have a large number of ties within groups and a small number of ties that cut across groups.

Some of these ideas, like overlapping cliques and simulated annealing, will be familiar to seasoned social network analysts. Many others, however, are new and several are quite creative. In particular, edge betweenness, modularity, the use of 3-cycles, short random walks and graph coloring appear to have promise.

Almost all of these contributions focused on building new tools to uncover groups or communities. They all reported applications to data, but for the most part, their applications were merely illustrative. The main thrust of this research has been to build better and faster group-finding algorithms. That preoccupation with developing ever faster algorithms may not seem too important to most social network analysts, but many applications—particularly those in biology—involve data sets that involve connections linking hundreds of thousands or millions of nodes. For those applications speed is essential.

Positions

Concern with the positions occupied by individual actors has been the second main theme in social network analysis. Four kinds of positions have been defined. First, positions in groups—*core* and *periphery*—have been specified. Second, a good deal of attention has been focused on *social roles*. Third, some attention has also been devoted to the study of the positions of nodes in *hierarchical structures*. And fourth, social network analysts have been concerned with the structural *centrality* of nodes in networks.

Core and peripheral positions in groups were first defined by early network analysts, Davis, Gardner and Gardner (1941). As they described this idea (p. 150):

> Those individuals who participate together most often and at the most intimate affairs are called *core members*; those who participate with core members upon some occasions but never as a group by themselves alone are called *primary members*; while individuals on the fringes, who participate only infrequently, constitute the *secondary members* of a clique.

Various others followed up on this observation and algorithms for finding core and peripheral positions in groups were proposed by Bonacich (1978), Doreian (1979), Freeman and White (1993) and Skvoretz and Faust (1999). Finally, in a pair of articles, Borgatti and Everett (1999) and Everett and Borgatti (2000) developed a full model of core/periphery structure.

The intuitive idea of social role was introduced by the anthropologist, Ralph Linton (1936). The notion was that two individuals who were, say, both fathers of children, occupied a similar position as a consequence of their being fathers. They could, it was assumed, be expected to display similar behaviors.

This idea was spelled out by Siegfried Nadel (1957) and formalized by Lorrain and White (1971) in their model of *structural equivalence*. In that model, two individuals are structurally equivalent if they have the same relations linking them to the same others.

Other social network analysts concluded that structural equivalence was too restrictive to capture the concept of social role (Sailer, 1978). So they were quick to propose other models that relaxed the restrictions of structural equivalence. These include *regular equivalence*, *isomorphic equivalence*, *automorphic equivalence*, and *local role equivalence*. These ideas are all thoroughly reviewed in Wasserman and Faust (1994).

The third kind of positional model used in social network analysis is focused on hierarchies or dominance orders. The study of dominance began with Pierre Huber's (1802) observations of dominance among bumblebees. Huber was an ethologist, and most of the research and model building about dominance has remained in ethology. But Martin Landau (1951), who was both an ethologist and a social network analyst, created a formal model of hierarchical structure for social network analysts. And another social network analyst, James S. Coleman (1964), proposed an alternative model. More recently, Freeman (1997) adapted an algebraic model from computer science (Gower, 1977) to be used in social network analysis. And Jameson, Appleby and Freeman (1999) took a model from psychology (Batchelder and Simpson (1988) and applied it to the study of social networks.

The fourth and final kind of model of social position is based on the notion of centrality. Alex Bavelas (1948) and Harold Leavitt (1951) originally developed the idea of structural centrality at the Group Networks Laboratory at the Massachusetts Institute of Technology. Their conception of centrality, based on the distance of each node to all the others in the graph, was used to account for differences in performance and morale in an organization.

Very soon a large number of other conceptions of centrality were introduced. Those based on graph theory were reviewed (Freeman, 1979) and reduced to a set of three. They included Sabidussi's (1966) measure based on *closeness*, Nieminen's (1974) measure based on *degree* and Freeman's (1977) measure based on *betweenness*.

In addition to these graph theoretic measures, Bonacich (1972, 1987) introduced an algebraic centrality measure. His measure is based on the concept of *eigenstructure*; it is determined by a combination of the degree of a node, the degrees of its neighbors, the degrees of their neighbors and so on.

The community of physicists has not displayed any major interest in the first three of these kinds of positions developed in social network analysis. Physicist Petter Holme (2005) did write an article about core/periphery structures. And in a review article, Mark Newman (2003) introduced structural equivalence to physicists. Petter Holme and Mikael Huss (2005) reviewed the social network equivalence measures and applied them in the study of protein function in yeast. Finally, Juyong Park and Mark Newman

(2005) introduced a new model of dominance and applied it to ranking American college football teams.

The physicists, however, were quick to adopt the ideas about centrality that had been developed in social network analysis. And they immediately passed them on to biologists. Figure 4 displays the number of articles on centrality published each year by social network analysts and the number published by physicists and biologists. It is clear that once they began publishing in this area, the physicists and biologists quickly overtook the social network analysts.

Figure 4. Articles on Centrality by Date and by Field (From Freeman, 2008)

In working with centrality, though, the physicists took a very different approach than the one they used when they dealt with the group or community concept. As we saw above, most of their contributions to the study of groups involved the development of new models and the introduction of refined procedures for finding groups. But, with centralities, most of the physicists' work has involved applications; they simply found new problems to which standard centrality measures could be fruitfully applied.

Many of the areas in which physicists applied centrality may seem quite surprising. Only a few of their applications fall into what most outsiders would think of as belonging to physics. These include packet switching in the internet, electronic circuitry and the electric power grid (Freeman, 2008).

A great many more of these applications involve areas that traditionally are considered to fall in the domain of social network analysis. These include studies of friendships linking students, contacts among prisoners, email contacts, telephone conversations, scientific collaboration, corporate interlock and links among sites in the World Wide Web (Freeman, 2008).

By far the most common application of centrality has been to problems in biology. This work was started by physicists (Jeong, Mason, Barabási and Oltavi, 2001) who studied interactions among proteins. But, almost immediately, biologists themselves began to use centrality ideas in their research. Two biologists, Wagner and Fell (2001) examined centrality in a study of metabolic networks. And a year later, four molecular biologists, Vendruscolo, Dokholyan, Paci and Karplus (2002) used centrality in a study of protein folding. These three themes, protein-protein interaction, metabolic networks and protein folding have all come to rely heavily on the use of various centrality models and have produced a great deal of research (Freeman, 2008).

Summary and Conclusions

In social network analysis we have a field with a long history. It began in the late 1930s. And it emerged again and again in different social science disciplines and in various countries. But in the 1970s all these separate research efforts came together and merged into a single coherent research effort embodying a structural perspective.

But in the late 1990s a new kind of situation arose. A completely alien field, physics, embraced the same kind of structural perspective that was embodied in social network analysis. Moreover, a good many of these physicists did not limit their research to the physical realm, but studied the patterning of links among social actors. One physicist, T. S. Evans (2004), reported on this trend to his fellow physicists:

> If you are naturally skeptical about trendy new areas of physics and attempts to mix physics with anything and everything, then the citations of papers in journals of sociology . . . and of books on archeology and anthropology . . . may just be the last straw!

Thus, though it may not be mainstream physics, at least some physicists have defined social network analysis as a proper part of their discipline.

To understand how this occurred, we need to look at physics and biology in the late 1990s. Both fields were suddenly faced with mammoth amounts of structural data. In physics, data on the internet became available. These data involve millions of computers, all linked by wires, fiber-optic cables and wireless connections. And in biology data on genetic and metabolic networks was being produced by all the genome research. In both fields investigators were confronted with data on very large networks.

These investigators needed tools—both intellectual and computational—that would help them to grapple with these huge new network data sets. So they turned to a field that had been dealing with network data for sixty years, social network analysis. They drew on ideas from social network analysis and they used analytic tools developed in that field. They refined existing tools and developed new ones. Sometimes they reinvented established tools and sometimes they rediscovered known results, but often they contributed important new ways to think about and analyze network data.

More important, at least some of these physicists have become increasingly involved in social network research. They have developed new tools aimed toward the study of social networks (Watts and Strogatz, 1998). They have reanalyzed standard social network data sets (Girvan and Newman, 2002; Holme, Huss and Jeong, 2003; Kolaczyk, Chua and Barthelemy, 2007; Newman, 2006).

Physicists have increasingly begun to cite social network articles. Girvan and Newman (2002), for example, cited 8 social network articles among their 29 citations. Fortunato, Latora and Marchiori (2004) cited 9 social network articles in 27 citations. And Holme and Huss (2005) cited 5 in 34 citations. On the other hand, most social network analysts have resisted citing physicists. Many, I suspect, still view the physicists as "alien invaders."

Physicists have used computer programs produced by social network analysts in their data analyses, and they have produced new programs that include some of the models developed in social network analysis (Freeman, 2008). In addition, a few physicists have attended the annual Sun Belt social network meetings.[6] And a few social network analysts have been invited to the meetings of the physicists.[7] Representatives of each discipline are beginning to publish in journals usually associated with the other.[8] There are even some joint publications (e. g. Reichardt and White, 2007; Salganik, Dodds, Sheridan and Watts, 2006).

My earlier hope for rapprochement between physics and social network analysis, it seems, is beginning to take place. All that is required now is that the social network analysts relax their claim of ownership of the field. The

---

6 Freeman (2004, p. 166) mentions the attendance of physicists Watts, Newman and Hoser at the social network meetings.

7 Social network analysts, Vladimir Batagelj and Linton Freeman were invited to the Summer Workshop in Complex Systems and Networks put on by physicists in Transylvania in 2007.

8 See, for example, physicists Watts (1999), Holme, Edling, Liljeros (2004) and Newman (2005) publishing in *Social Networks* or network analysts, Borgatti, Mehra, Brass and Labianca (2009) appearing in *Science*.

physicists are making important contributions to what could easily end up as a collective effort.[9]

References

Alba, R. D. A graph theoretic definition of a sociometric clique. Journal of Mathematical Sociology, 3:113-126, 1973.

Arabie, P. and J. D. Carroll. Conceptions of overlap in social structure. In L. C. Freeman, D. R. White and A. K. Romney, editors, Research methods in social network analysis. George Mason University Press, Fairfax, VA, 1989.

Barabasi, A-L. and R. Albert. Emergence of scaling in random networks. Science, 286:509-512, 1999.

Batchelder, W. H. and R. S. Simpson. Rating system for human abilities: the case of rating chess skill. , UMAP Modules in Undergraduate Mathematics and its Applications. Consortium for Mathematics and its Applications, Arlington, MA, 1989.

Barabasi, A-L. Linked: The New Science of Networks, Cambridge, MA, Perseus, Cambridge, MA, 2002.

Bavelas, A. A mathematical model for small group structures. Human Organization, 7:16-30, 1948.

Beum, C. O. and E. G. Brundage. A method for analyzing the sociomatrix. Sociometry, 13:141-145, 1950.

Bonacich, P. Factoring and weighting approaches to status scores and clique identification. Journal of Mathematical Sociology, 2:113-120, 1972.

[9] A hopeful sign is that Jeroen Bruggeman (2008) cites 77 reports by physicists in his new book on social network analysis.

Bonacich, P. Using boolean algebra to analyze the overlapping memberships. Sociological Methodology:101-115, 1978.

Bonacich, P. Power and centrality - a family of measures. American Journal of Sociology, 92(5):1170-1182, 1987.

Bonacich, P. The invasion of the physicists. Social Networks, 26(3):285-288, 2004.

Boyd, J. P. Social semigroups, Fairfax, VA, 1991. George Mason University Press, Fairfax, VA, 1991.

Borgatti, S. P. and M. G. Everett. Network analysis of 2-mode data. Social Networks, 19(3):243-269, 1997.

Borgatti, S. P. and M. G. Everett. Models of core/periphery structures. Social Networks, 21(4):375-395, 1999.

Borgatti, S. P., M. G. Everett and L. C. Freeman. Ucinet, version iv, Columbia, SC, 1992. Analytic Technology, Columbia, SC, 1992.

Breiger, R. L. The duality of persons and groups. Social Forces, 53:181-190, 1974.

Breiger, R. L., S. A. Boorman and P. Arabie. An algorithm for clustering relational data, with applications to social network analysis and comparison to multidimensional scaling. Journal of Mathematical Psychology, 12:328-383, 1975.

Cartwright, D. and F. Harary. Structural balance: a generalization of heider's theory. Psychological Review, 63:277-292, 1956.

Capocci, A., V. D. P. Servedio, G. Caldarelli and F. Colaiori. Detecting communities in large networks. Physica A: Statistical Mechanics and its Applications, 352:669-676, 2005.

Clauset, A., M. E. J. Newman and C. Moore. Finding communities in very large networks. Physical Review E, 70(6):066111, 2004.

Coleman, J. S. Introduction to Mathematical Sociology, New York, 1964. Free Press, New York, 1964.

Cooley, C. H. Social organization, New York. Shocken Books, New York.

Davis, A., B. B. Gardner and M. R. Gardner. Deep south, Chicago, 1941. The University of Chicago Press, Chicago, 1941.

Dekker, A. Visualization of social networks using CAVALIER. Proceedings of the 2001 Asia-Pacific Symposium on Information Visualization, pages 49-55. 2001.

Djidjev, H. A fast multilevel algorithm for graph clustering and community detection. arXiv:0707.2387v1 [physics.data-an], 2007.

Doreian, P. On the deliniation of small group structures. In M. Hudson, editor, Classifying social data. Jossey-Bass, San Francisco, 1979.

Duch, J. and A. Arenas. Community detection in complex networks using external optimization. Physical Review E, 72(2027104), 2005.

Durkheim, E. The division of labor in society, New York. The Free Press, New York.

Evans, T. S. Complex networks. Contemporary Physics, 45(6):455-474, 2004.

Everett, M. and S. R. Borgatti. Peripheries of cohesive subsets. Social Networks, 21(4):397-407, 2000.

Festinger, Leon, Stanley Schachter and Kurt W. Back. Social pressures in informal groups, New York, 1950. Harper & Bros., New York, 1950.

Fiduccia, C. M. and R. M. Mattheyses. A linear time heuristic for improving network partitions. Desigb Automation, 1982, 19th Conference on, pages 175-181. 1982.

Fiedler, M. Algebraic connectivity of graphs. Czechoslovak Mathematics Journal, 23:298-305, 1973.

Fiedler, M. A property of eigebvectors of non-negative symmetric matrices and its application to graph theory. Czechoslovak Mathematics Journal, 25:619-633, 1975.

Flake, G. W., S Lawrence and C. L. Giles. Efficient identification of web communities. Proceedings of the Sixth International Conference on Knowledge Discovery and Data Mining, pages 150-160. 2000.

Fortunato, S., V. Latora and M. Marchiori. Method to find community structures based on information centrality. Physical Review E, 70(5):056104, 2004.

Frank, K. A. Identifying cohesive subgroups. Social Networks, 17(1):27-56, 1995.

Freeman, L. C. The development of social network analysis:a study in the sociology of science, Vancouver, B. C., 2004. Empirical Press, Vancouver, B. C., 2004.

Freeman, L. C. Going the wrong way on a one-way street: Centrality in physics and biology. Journal of Social Structure, 9(2), 2008.

Freeman, L. C. A set of measures of centrality based on betweenness. Sociometry, 40:35-41, 1977.

Freeman, L. C. Centrality in social networks: conceptual clarification. Social Networks, 1:215-239, 1979.

Freeman, L. C. On the sociological concept of "group": a empirical test of two models. American Journal of Sociology, 98:152-166, 1992.

Freeman, L. C. Finding groups with a simple genetic algorithm. Journal of Mathematical Sociology, 17:227-241, 1993.

Freeman, L. C. Uncovering organizational hierarchies. Computational and Mathematical Organization Theory, 3:5-18, 1997.

Freeman, L. C., A. K. Romney and S. C. Freeman. Cognitive structure and informant accuracy. American Anthropologist, 89:311-325, 1987.

Freeman, L. C. and D. R. White. Using galois lattices to represent network data. In P. Marsden, editor, Sociological methodology 1993. Blackwell, Cambridge, MA, 1993.

Freeman, L. C. and C. M. Webster. Interpersonal proximity in social and cognitive space. Social Cognition, 12:223-247, 1994.

Girvan, M. and M. E. J. Newman. Community structure in social and biological networks. PNAS, 99(12):7821-7826, 2002.

Glover, F. Tabu search - part I. ORSA Journal on Computing, 1:190-206, 1989.

Glover, F. Tabu search - part II. ORSA Journal on Computing, 2:4-32, 1990.

Gower, J. C. The analysis of asymmetry and orthoganality. In J. Barra, F. Brodeau and G. Romier, editors, Recent developments in statistics. North Holland, Amsterdam, 1977.

Holme, P. Core-periphery organization of complex networks. Physical Review E, 72(4), 2005.

Holme, P. and M. Huss. Role-similarity based functional prediction in networked systems: application to the yeast proteome. Journal of the Royal Society Interface, 2(4):327-333, 2005.

Holme, P., M. Huss and H. W. Jeong. Subnetwork hierarchies of biochemical pathways. Bioinformatics, 19(4):532-538, 2003.

Homans, G. C. The human group, New York, 1950. Harcourt, Brace and Company, New York, 1950.

Huber, P. Observations on several species of the genus apis, known by the name of humble bees, and called bombinatrices by linneaus. Transactions of the Linnean Society of London, 6:214-298, 1802.

Hubert, L. Approximate evaluation techniques for the single-link and complete-link hierarchical clustering procedures. Journal of the American Statistical Association, 69:698-704, 1974.

Jameson, K. A., M. C. Appleby and L. C. Freeman. Finding an appropriate order for a hierarchy based on probabilistic dominance. Animal Behaviour, 57:991-998, 1999.

Jeong, H., S. P. Mason, A. L. Barabasi and Z. N. Oltvai. Lethality and centrality in protein networks. Nature, 411(6833):41-42, 2001.

Killworth, P. D., C. McCarty, H. R. Bernard, E. C. Johnsen, J. Domini and G. A. Shelley. Two interpretrations of reports of knowledge of subpopulation sizes. Social Networks, 25:141-169, 2003.

Killworth, P. D., C. McCarty, H. R. Bernard and M. House. The accuracy of small world chains in social networks. Social Networks, 28:85-96, 2006.

Kochen, M. The small world, Northwood, NJ, 1989. Ablex Publishing Company, Northwood, NJ, 1989.

Kolaczyk, E. D., D. B. Chua and M. Barthelemy. Co-betweenness: A pairwise notion of centrality. arXiv.org>physics, 2007.

Landau, H. G. On dominance relations and the structure of animal societies: ii. some effects of possible social factors. Bulletin of Mathematical Biophysics, 13:245-262, 1951.

Leavitt, H. J. Some effects of communication patterns on group performance. Journal of Abnormal and Social Psychology, 46:38-50, 1951.

Levine, J. H. The sphere of influence. American Sociological Review, 37:14-27, 1972.

Lewin, Kurt and Ronald Lippitt. An experimental approach to the study of autocracy and democracy: a preliminary note. Sociolmetry, 1(3/4):292-300, 1938.

Linton, R. The study of man, New York, 1936. D. Appleton-Century

Company, New York, 1936.

Lorrain, F. P. and H. C. White. Structural equivalence of individuals in social networks. Journal of Mathematical Sociology, 1:49-80, 1971.

Luce, R. D. and A. Perry. A method of matrix analysis of group structure. Psychometrika, 14:95-116, 1949.
Maine, H. Ancient law, London. Oxford University Press, London.

Milgram, S. The small world problem. Psychology Today, 22:61-67, 1967.

Mokken, R. J. Cliques, clubs and clans. Quantity and Quality, 13:161-173, 1979.

Moody, J. and D. R. White. Structural cohesion and embeddedness: a hierarchical concept of social groups. American Sociological Review, 68(1):103-127, 2003.

Moreno, Jacob L. Application of the group method to classification, New York, 1932. National Committee on Prisons and Prison Labor, New York, 1932.

Moreno, Jacob L. Who shall survive?, Washington, DC, 1934. Nervous and Mental Disease Publishing Company, Washington, DC, 1934.

Moreno, J. L. and H. H. Jennings. Statistics of social configurations. Sociometry, 1:342-374, 1938.

Mullins, N. C. and C. J. Mullins. Theories and theory groups in contemporary American sociology, New York, 1973. Harper & Row, New York, 1973.

Nadel, S. F. The theory of social structure, London, 1957. Cohen and West, London, 1957.

Newman, M. E. J. The structure and function of complex networks. Siam Review, 45(2):167-256, 2003.

Newman, M. E. J. Detecting community structure in networks. European

Physical Journal B, 38(2):321-330, 2004.

Newman, M. E. J. Finding community structure in networks using the eigenvectors of matrices. Physical Review E, 74(3):19, 2006.

Newman, M. E. J. and M. Girvan. Finding and evaluating community structure in networks. Physical Review E, 69, 026113, 2004.

Newcomb, T. M. The acquaintance process, New York, 1961. Holt, Rhinehart, and Winston, New York, 1961.

Nieminen, J. On centrality in a graph. Scandinavian Journal of Psychology, 15(4):332-336, 1974.

Park, J. and M. E. J. Newman. A network-based ranking system for U. S. college football. Journal of Statistical Mechanics:P Theory and Experiment, P10014, 2005.

Parlett, B. N. A new look at the Lanczos method for solving symmetric systems of linear equations. Linear Algebra Appl., 29:323-346, 1980.

Pool, I. D. and M. Kochen. Contacts and influence. Social Networks, 1(1):5-51, 1978.

Pons, P. and M. Latapy. Computing communities in large networks using random walks. Journal of Graph Algorithms and Applications, 10(2):191-218, 2006.

Pothen, A., H. Simon and K-P. Liou. Partitioning sparse matrices witheigenvalues of graphs. SIAM Journal of Matrix Analysis Appl., 11(3):430-452, 1990.

Price, D. J. Networks of scientific papers. Science, 149:510-515, 1965.

Price, D. J. A general theory of bibliometric and other cumulative advantage processes. Journal of the American Society for Information Science, 27:292-306, 1976.

Radicchi, F., C. Castellano, F. Cecconi, V. Loretto and D. Parisi.

Defining and identifying communities in networks. Proceedings of the National Academy of Sciences, 101:2658-2663, 2004.

Ragnaven, U. N., R. Albert and S. Kumara. Near linear time algorithm to detect community structures in large-scale networks. Physical Review E, 76:036106, 2007.

Reichardt, J. and S. Bornholdt. Statistical mechanics of community detection. Physical Review E, 74:016110, 2006.

Reichardt, J. and D. R. White. Role models for complex networks. The European Physical Journal B, 60:217-224, 2007.

Roberts, J. M. Correspondence analysis of two-mode network data. Social Networks, 22:65-72, 2000.

Roethlisberger, Fritz J. and W. J. Dickson. Management and the worker, Cambridge, MA, 1939. Harvard University Press, Cambridge, MA, 1939.

Sabidussi, G. The centrality index of a graph. Psychometrika, 31(4):581-603, 1966.

Sailer, L. D. and S. J. C. Gaulin. Proximity, sociality and observation: the definition of social groups. American Anthropologist, 86:91-98, 1984.

Salganik, M. J., P. S. Dodds and D. J. Watts. Experimental study of inequality and unpredictability in an artificial cultural market. Science, 311:854-856, 2006.

Scott, J. Social network analysis, Newbury Park, CA, 1992. Sage, Newbury Park, CA, 1992.

Seary, A. J. and W. D. Richards. Dynamic Social Network Modeling and Analysis, Washington, 2003, type Spectral methods for analyzing and visualizing networks: An introduction, pages 209-228. The National Academies Press, Washington, 2003.

Skvoretz, J. and K. Faust. Logit models for affiliation networks. Sociological Methodology, 29:253-280, 1999.

Spencer, H. The principles of sociology, New York, 1897. Appleton-Century-Crofts, New York, 1897.

Tönnies, F. Fundamental concepts of sociology, New York. American Book Company, New York.

Vendruscolo, M., N. V. Dokholyan, E. Paci and M. Karplus. Small-world view of the amino acids that play a key role in protein folding. Physical Review E, 65(6):061910-061913, 2002.

Wagner, A. and D. A. Fell. The small world inside large metabolic networks. Proceedings of the Royal Society of London Series B-Biological Sciences, 268(1478):1803-1810, 2001.

Warner, W. Lloyd and Paul S. Lunt. The social life of a modern community, New Haven, CT, 1941. Yale University Press, New Haven, CT, 1941.

Wasserman, S. and K. Faust. Social network analysis: methods and applications, Cambridge, 1994. Cambridge University Press, Cambridge, 1994.

Watts, D. J. Six degrees: the science of a connected age, New York, 2003. W. W. Norton and Company, New York, 2003.

Watts, D. J. and S. H. Strogatz. Collective dynamics of 'small-world' networks. Nature, 393(6684):440-442, 1998.

White, H. C., S. A. Boorman and R. L. Breiger. Social structure from multiple networks I: blockmodels of roles and positions. American Journal of Sociology, 81:730-781, 1976.

Wu, F. and B. A. Huberman. Finding communities in linear time: a physics approach. European Physical Journal B, 38(2):331-338, 2004.

Wu, Z. and R. Leahy. An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15:1101-1113, 1993.

Zachary, W. An information flow model for conflict and fission in small groups. Journal of Anthropological Research, 33:452-473, 1977.