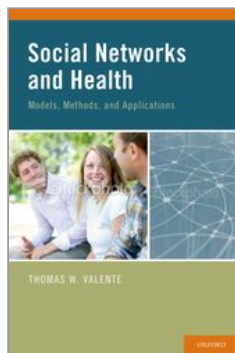


University Press Scholarship Online

Oxford Scholarship Online



Social Networks and Health: Models, Methods, and Applications

Thomas W. Valente

Print publication date: 2010

Print ISBN-13: 9780195301014

Published to Oxford Scholarship Online: May 2010

DOI: 10.1093/acprof:oso/9780195301014.001.0001

Positions

Thomas W. Valente

DOI:10.1093/acprof:oso/9780195301014.003.0007

Abstract and Keywords

This chapter introduces positional analysis which is conducted by defining positions in a network and reducing a network to these positions and mapping the relations between positions. The chapter also covers individual measures of positions in which nodes are defined as occupy the same position based on their connections to others or the similarity in their distances to others in the network. These individual equivalence positions can be used to group people into positions and to model the ways in which network positions influence behavior. Positions are composed of people who seem to be in the same space in the network, regardless of whether they are directly connected to one another, though they might be.

Keywords: network positions, roles, hierarchical, structural equivalence, image matrix

This chapter introduces positional analysis, which is conducted by defining positions in a network and reducing a network to these positions and mapping the relations between positions. The chapter also covers individual measures of positions in which nodes are defined as occupying the same position based on their connections to others or the similarity in their distances to others in the network. These individual equivalence positions can be used to group people into positions and to model the ways in which networks influence behavior.

Groups are aggregates of nodes/people who communicate or are connected to one another at a higher rate than others in the network. A network position, in contrast, is a set of nodes that occupy the same place or have similar relations with others in the network. Positions are composed of people who seem to be in the same space in the network, regardless of whether they are directly connected to one another, though they might be. Generally, a *position* is a set of nodes that has the same links to the same others or the same types of others.

The theoretical basis for defining positions comes from the sociological insight that people who occupy the same roles often act similarly. For example, fathers are alike even though they are fathers to different children. Network-defined positions constitute roles in the network and, consequently, **(p.115)** people in the same position may behave similarly. Positional network analysis consists of (1) using mathematical algorithms to define distinct positions in the network, (2) studying how those positions relate to one another, and (3) determining how membership in a position might influence behavior.

Hummon and Carley (1993) studied publications in the *Social Networks* journal and concluded that the study of positions was the central theme in research on social networks up to the early 1990s. Position analysis has been popular in social network analysis because it uncovers macro-level structure of the network from micro-level analysis of network relations. Structural analysis is, to some extent, identifying positions in the network and then discovering how those positions relate to one another. Finding groups and measuring centrality and other network analysis activities are often considered less-structural analysis because they do not explicitly examine the interaction between micro- and macro-level network properties.

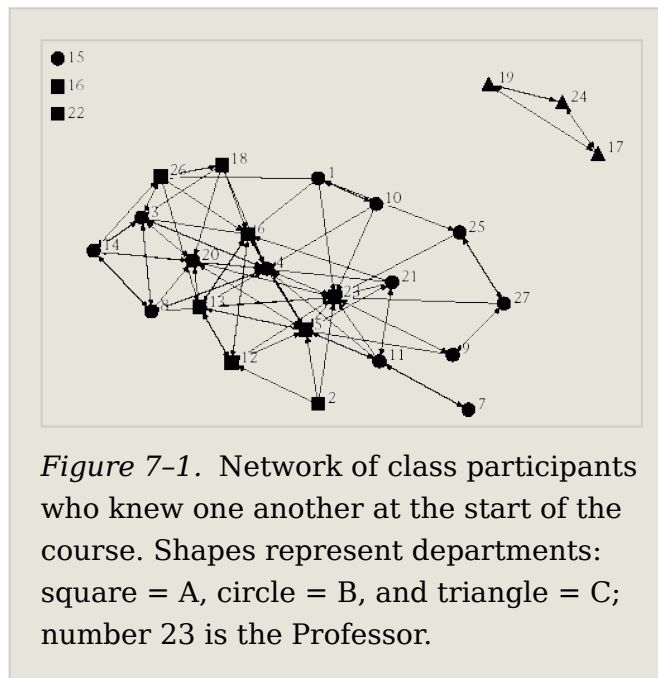
Position analysis is conducted at both the individual and network level. Individual position analysis is conducted by creating a measure of how equivalent two people are and then assigning a score for each pair of individuals based on that equivalence. Network-level position analysis is conducted by using mathematical algorithms to find positions in the network and then studying the relationships between these positions. Network-level position analysis consists of a set of positions and their interrelationships, whereas individual position analysis consists of a matrix of positional equivalence scores.

Network-Level Positions

Lorain and White (1971) wrote an influential paper in which they proposed that the relations between people in a network can be reduced to a set of positions and the relations between these positions also treated as a network (a meta-level network). In some research, the network representing the relations between positions is referred to as the *reduced-form network*. Lorain and White (1971) proposed that the network could be reduced to a set of blocks and the relations between blocks studied. This was called blockmodeling. Network researchers have developed different methods to identify the blocks (positions) and determine relations between positions (Doreian et al., 2005). The matrix or network that reports how the positions interact is referred to as the *image matrix*, the reduced-form network or matrix.

As in group definitions (Chapter 6), a strict criterion can be set for people to be members of a position and then the criterion gradually relaxed. As the definition is relaxed, it is possible to examine the pattern of positions created. The pattern of how members join positions can reveal facets of **(p.116)**

network structure and, of course, many different algorithms can be proposed for defining positions. In short, positional analysis and blockmodeling can get quite complex. To illustrate how positions are defined, we use UCINET 6.0



(Borgatti et al., 2006) and the network shown in Figure 7-1, which shows who knew whom at the start of a course on network analysis. This course had 24 people initially, 3 of whom were faculty members, 1 who was the primary instructor, and 2 who sat in on many of the lectures to learn about social network analysis.

The class members were spread across three departments: A, B, and C. There were three faculty members in the course with one being the Professor, number 23. Casual inspection of the network shows some obvious structural patterns: (1) there were three isolates who were people registered for the wrong class and never attended; (2) there were three students from Department C who enrolled in the course and knew one another but no one in other departments; and (3) there might be some clustering based on other department affiliations but that is not clear from the graph.

A positional algorithm was applied to the data to determine if the different departments or statuses (faculty versus student) constituted distinct positions. The algorithm used was an optimization algorithm, which attempts to find the positions so that the blocks (the set of links within and between positions) have the least number of changes required to make the blocks all zeros or ones. A network divided into positions such that the blocks are all zeros or **(p.117)** all ones is a perfectly partitioned network. This would indicate that all of the ties are within positions and there are no ties between positions. Notice the similarity between the positional approach and the Girvan-Newman method described in Chapter 6. The difference here is that the researcher specifies beforehand (a priori) which nodes belong to which positions.

Figure 7-2 reports the results of a blockmodel analysis of the network in Figure 7-1 using UCINET VI (Borgatti et al., 2006). Six positions were specified to test whether the four departments and two statuses (student and faculty) would yield different positions. Deciding on how many positions expected from a positional analysis should be driven by theory so that positional analysis confirms hypotheses rather than being treated as an exploratory approach in which one specifies various numbers in hopes of finding an interesting partition.

The results were interesting in that they partially confirmed intuition: department membership accounted for much of the variation in the network ties. Number 23 (B_F3), the instructor, was assigned a position by himself, which is not surprising because most of the students knew the instructor before the class. Positions 1 and 5 consisted of all Department B students (except number 2, A_F1), position 4 consisted of Department C students, and position 6 consisted of Department A students, one faculty member, and the “other” faculty member. Position 3 was mixed half Department A and half Department B.

UCINET also reports the number of “errors,” which is the number of links that would need to be added or deleted to have a perfect blockmodeling structure. A perfect blockmodeling structure occurs when each block is composed of all ones (links) or all zeros (no links). For example, position one has 10 links to other people in position one. Since this value is below 50% of the total possible [$28 = (8*7)/2$], it is more efficient to delete links (rather than add 18) to achieve a zero or one block. There are no errors between position 1 and position 2 because everyone in position 1 knew the person who occupies position 2. For links between position 1 and 3 through 6, “errors” consist of deleting links in those blocks as well, for a total of 22 errors. The errors in position 2 consist of one link within position 1, one with position 5, and one with position 6. Positions 2 through 6 had errors of 3, 11, 0, 8, and 9 for a total of 53 errors.

UCINET provides a display of the blockmodel network, the set of ties in the original matrix but with the people (nodes) sorted according to their position. One can easily see if there are links within and between positions; blocks with lots of zeros represent no connection between positions, whereas blocks with lots of ones indicate connections between positions. To determine how much connection is a *lot* of connection, a general rule is to compare densities within and between blocks with the density of the **(p.118)**

(p.119)
whole network.
If block
densities are
larger than the
overall
network
density, there
is a connection
or link
between these
blocks. In this
case, the
overall
network
density is
12.7%. This
value (12.7%)
can be
compared to
the densities of
the
subnetworks
indicated by
links between
positions (the
block of links
indicating
connections
between
positions). If
the density
between two
blocks
(positions)
exceeds 12.7% (the overall average), then the blocks are
connected. When densities are below the average, the blocks/
positions are not connected.

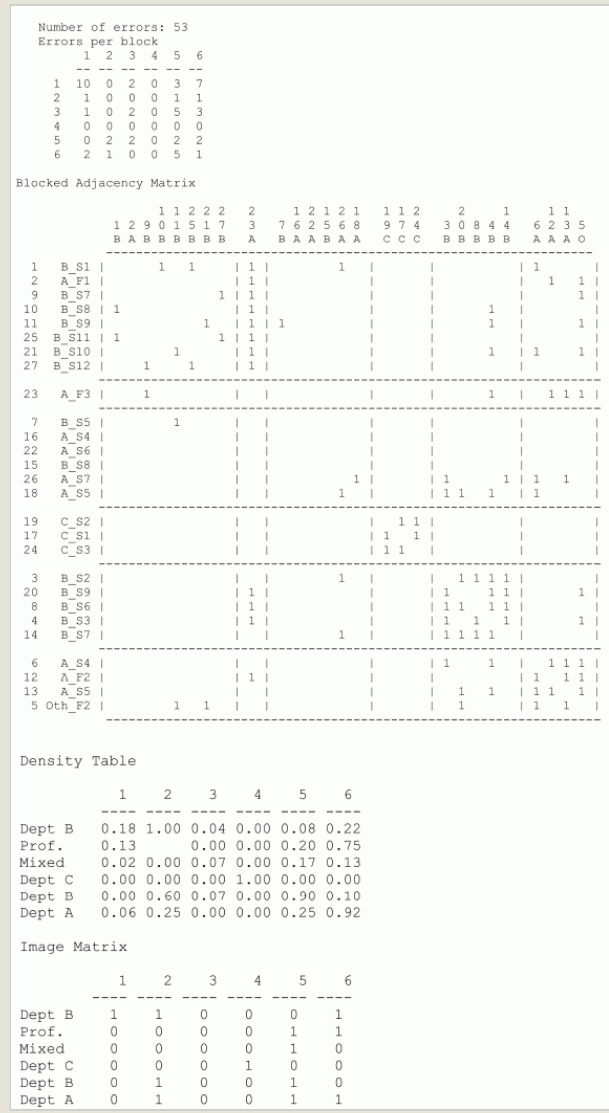


Figure 7-2. Blockmodel results for the network class data.

It is worth reiterating the approach here. The overall network is partitioned into a set of positions. The links in the network are then reconfigured to show the links within a position and between positions. These are the “blocks.” If the proportion of links (the density) connecting any two positions (within a block) is greater than that which occurs on average in the network (the overall density), then a link between positions can be inferred. The network has been reduced to a set of positions, and it has been determined at the macro-level how these positions relate to one another. This network of 27 people has been reduced to a network of six positions. The resulting reduced-form network is referred to as the image matrix. So the image network or matrix is the set of links between positions—in this case, a six-by-six matrix.

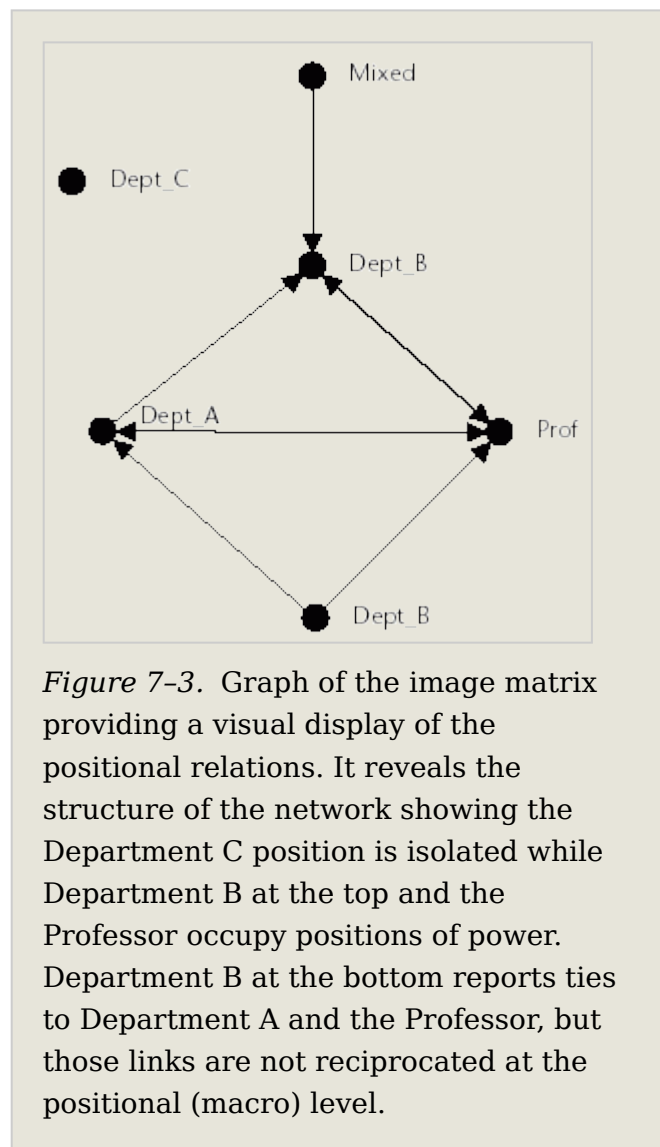
In this class example, position 1 is connected to positions 1, 2, and 6, because the densities (0.18, 1.00, and 0.22) all equal or exceed the overall density of 0.127. Similarly, position 2 is connected to positions 5 and 6, and so on. Note the image matrix does not have to be symmetric: relations from one position can be directed to another without those relations being reciprocated. For example, position 1 is connected to position 2, but 2 is not connected to 1.

A substantive aspect of blockmodeling and positional analysis is to label the positions and interpret the relations between positions. So rather than just say positions 1, 2, 3, and so on, the research can examine the attributes of the people in positions to determine if they can be logically grouped. For example, in an organizational study, the members of positions may conform mostly to organizational roles or task responsibilities (management versus sales, for example). Often, if a position is linked asymmetrically to another, researchers interpret this relationship as one of power or control. In this example, the positions are clearly indicated by department affiliations: Departments A, B, and C and the Professor. The positions conform almost exactly to the departments with the one exception being the mixed department.

Figure 7-3 graphs the image matrix providing a visual display of the positional relations. It reveals the structure of the network showing that Department C is isolated. It also shows the Mixed position reports knowing people in Department B, but those ties are not reciprocated. The Department A position has reciprocated ties with the Professor but not with the two (p.120)

Department B positions. A naïve interpretation would lend one to believe the Professor is in the most comfortable position having two symmetric ties with Departments A and B and one incoming tie from Department B. Department B at the top, however, has a symmetric relation with the Professor and two incoming links from the Mixed and Department A positions. It is important to note that this

type of position analysis can be conducted using attributes as the positional indicators. For example, the researcher could repeat this analysis by assigning four positions based on the department affiliations rather than letting the computer generate the positions.



It is perhaps instructive to compare this positional analysis to a group analysis (Chapter 6) of the same network. We ran the Newman-Girvan algorithm on these data and selected the best-fitting model. Some key differences emerged, including that the Professor was grouped with a new Mixed department group, which included students from Departments A and B and one other faculty member, and the other two faculty were grouped together. Thus, the resulting structure, using the Newman-Girvan group algorithm, was quite different from the positional analysis.

A network with several positions and an image network consisting of directional links pointing to one of those positions is an example of a hierarchical network. In this case, the network analysis has uncovered a situation **(p.121)** in which one position holds a high status position and possibly members of this position enjoy a position of power. Although such a network is also likely to be centralized (see Chapter 5), the positional analysis and resulting image diagram provide a different, and possibly more informative, analysis of the overall network structure.

CONCOR

One of the most common structural equivalence position generators is CONCOR, which is an abbreviation of “converging correlations.” CONCOR uses correlation analysis to find positions in a network and uses a square binary network as input. The first step is to compute correlations between the rows in the matrix. This produces a matrix of correlations with each cell indicating the degree of correlation between cases (the degree of similarity between two nodes). The rows in the correlation matrix are then correlated. This process is repeated successive times. Ultimately, the repeated correlations produce a matrix of all positive and negative ones, indicating cases perfectly similar and dissimilar. The network is then partitioned into two sets based on the positive and negative ones (the perfect correlations). CONCOR has thus produced a partition of the network into two positions. The process is repeated to divide the initial two groups into four groups (two for each initial position) and so on until each position has as few as two members. The researcher decides at which point to stop the CONCOR process and accept the positions identified.

CONCOR analysis is appealing because it provides an unbiased, mathematical partition of the network into positions, requiring little substantive or theoretical insight from the researcher. One limitation of the CONCOR approach is that the analysis is driven entirely by the performance of the mathematical process of correlating rows and columns. A second limitation is that using correlations to measure node equivalence may not necessarily be the best method for doing so (see later). Further, because successive iterations of the correlation matrix eventually produce ones and negative ones, correlations of quite different magnitude are eventually treated the same. That is, a correlation of 0.51 and one of 0.90 may both be eventually converted to 1s during the same partition. A third limitation is that CONCOR forces a bifurcation of the network into two positions, and then two more, and so on, whereas many networks and subnetworks may not conveniently be divided into two distinct positions.

These limitations aside, CONCOR does a good job of identifying positions based on node similarities. The process is automatic and correlations are a good measure of similarity. Further, the researcher can examine the **(p.122)** successive partitions of the network into positions and choose to use a partition that conforms to some substantive knowledge of the data.

Individual Positional Measures

The preceding discussion provided a means of reducing a network to a set of positions and to map the relations between positions. It is often desirable to calculate the degree to which two people are similar, or occupy the same position in a network, without necessarily creating network positions. Individual positional measures provide measures of node similarity based on their connections in the network and in some cases based on all the links in the network. Individual position measures are dyadic measures as they are not measures for each person but, instead, are measures for every pair of people in the network.

The simplest positional measure would indicate the similarity of two people's contacts. For example, two people who are connected and not connected to exactly the same other people in the network are perfectly equivalent. Equivalent nodes occupy the same position. From a network perspective, two equivalent nodes are substitutable as they have the same exact relations with others in the network. Notice that two equivalent nodes do not have to be connected to one another to be equivalent. They can be perfectly equivalent and not connected. This measure of equivalence is referred to as *matches* in UCINET and is a simple comparison of two node's ties. Matches compare the links and nonlinks that each pair of nodes has with everyone else in the network, yet it may be more accurate to compare only the connections and ignore the nonconnections in the calculation. Because some network data are collected by asking for a person to list their five closest friends, the nonconnections that two people share are not as informative as the connections they have in common. Positive matches are the percentage of ties two people have in common and provide a good intuitive measure of structural similarity.

Figure 7-4 displays a structurally equivalent pair based on positive matches. Person A and Person B are connected to the same people and so they are perfectly substitutable and structurally equivalent from a network perspective. If Person A had one other connection, then it would slightly diminish the degree of structural equivalence or similarity between Persons A and B.

One useful measure of structural equivalence was developed by Burt (1987) based on distances. The distance measure of structural equivalence is derived from geographic measures in which two cities are similar if they are the same distances from other cities. For example, Chicago and St. Louis (two U.S. cities) are similarly distant from Los Angeles (LA), Denver, San Francisco, Seattle, New York (NY), Washington, DC, Atlanta, and so on. **(p.123)**

The difference in the distances between Chicago and every large city and between St. Louis and every large city is fairly small because Chicago and St. Louis are similarly distant from all U.S. cities. In contrast, the differences in distances

between NY and LA to all the large US cities is quite large because the cities near NY are far from LA and vice versa.

The calculation of distance differences is referred to as *Euclidean distance*, and this provides a measure of structural equivalence. Two people are structurally equivalent to the extent they have the same distances to everyone else in the network. So a measure of Euclidean distances is given by Burt (1987) as:

7-1

$$D_{ij} = \sum_k (z_{ik} - z_{jk})^2 + \sum_k (z_{ki} - z_{kj})^2$$

where z_{ik} is the distance from i to k measured as the number of steps in the network. This formula takes the difference of distances, squares each one (to get rid of negative values), and sums those differences. It is calculated in both directions (the distance from i to k and the distance from k to i) because networks can be asymmetric. The distance from i to k and the distance from k to i are not necessarily the same in network analysis. Using Euclidean distances to calculate structural equivalence is appealing because the measure considers the overall pattern of ties in the network, not just the direct ties for every pair of nodes.

Individual Measures as Positions

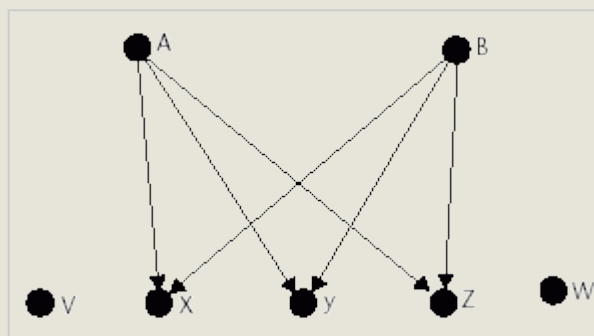


Figure 7-4. People are structurally equivalent when they are linked to the same other people. A and B have ties and non-ties to the same others in the network and so are structurally equivalent.

The individual measures of structural equivalence (SE), such as matches and Euclidean distance, can be used to find network-level positions. Once the matrix of SE scores are calculated, people can be grouped together based on **(p.124)** their similarity of SE scores. For example, everyone with similarity scores of 1.0 can be considered a position; then that criterion can be relaxed to 0.95, 0.90, and so on, down to zero. The result is a series of position grouping based on a relaxation of the criterion of what it means to be in a position.

The pattern of position identifications can be graphically displayed so that the researcher can see which nodes belong to which positions at various cut-off levels. One can then examine the positional definitions and choose a cut-off value to use to define the positions based solely on the empirical pattern or based on some substantive criterion. Researchers can compare different individual (read dyadic) measures of equivalence to determine whether there are differences in the identified positions.

Positions and Behavior

Clearly, finding positions in a network and assessing the degree of equivalence between actors is interesting from a research standpoint and can have substantive meaning when trying to understand networks. Does positional equivalence influence behavior? From a disease perspective, it is not immediately obvious how position equivalence affects one's likelihood of contracting a disease because communicable diseases are spread by person-to-person contact. On the other hand, people in the same positions may be exposed to the same kinds of pathogens and so positional equivalence could be important to know.

For example, a positional analysis of sexual relations among a large network might reveal a set of positions such that one position is obviously a core group of sexually active individuals. Other positions that have contact with the core group may be at increased risk for sexually transmitted infections due to their contact with that group. All members of these positions may be at increased risk for disease. Although the risk factor is sexual contact, the positional analysis indicates why some people are at increased risk compared with others, even though their rate of sexual activity may be the same.

For behavioral influence, positional and structural equivalence measures may matter very much. In any given community, two people who occupy the same position in the network are likely to be very similar and they are likely to monitor each other's behavior. This is particularly true in business; firms that occupy the same position in a network of suppliers, clients, and regulatory agencies are likely to be quite aware of what their structurally equivalent firms are doing because these are cues to appropriate strategic action. In short, as Burt (1987) argued, social influence may in some cases flow via structurally equivalence rather than direct ties.

There are several reasons why influence may pass between SE people. First, competition often motivates action and SE people or organizations may **(p.125)** feel in competition with one another and so monitor each other's behavior. Once one or a few people SE people engage in a behavior, others are likely to do so to remain competitive. Second, SE people, being connected to the same others, may be influenced by those to whom they are connected. For example, two SE professors are connected to the same students and so may get information and influence from those students. Finally, persons in the same position may have other attributes that put them in that position and these characteristics are associated with the adoption behavior. For example, managers in the same position may also be recipients of bonuses for adopting specific practices.

The influence of direct ties was labeled cohesive influence, and throughout the 1990s there were numerous comparisons between cohesion influence and that of structural equivalence. The central research question was: What influenced a person more in their adoption decisions—the behavior of people they are directly connected to, say as someone they turn for advice (cohesion), or the behavior of people who occupy the same position in the network structure (structural equivalence)?

In earlier analyses, Valente (1995) conducted position analysis (CONCOR) on 40 networks from three studies. Once the positions were established in each network, analysis (using ANOVA [analysis of variance]) was conducted to determine if behavioral adoption times were significantly different between positions. It was expected that people in the same position would have similar adoption times and people in different positions would have different adoption times. In 35 of 40 of the networks, the adoption times did not differ significantly between positions; thus, there was no support for the hypothesis that network position was associated with adoption.

The analysis was not necessarily comprehensive because there are multiple ways to determine positions in a network. Further, Valente (1995) did not examine each network to see if the CONCOR analysis produced similar positional structures. For example, did the networks have similar image matrix patterns or were they different between networks? Perhaps diffusion was a function of how the positions related to one another. It was also not clear how the network level positional analysis would compare to influences based on direct ties. One possible means to determine if network position affects behavioral adoption is to use individual positional analysis.

Network Weights

Valente (1995, 2005) compared network exposure based on cohesion with that based on structural equivalence in three diffusion network datasets. Exposure via direct contacts did not influence adoption, whereas in one **(p.126)** dataset, Brazilian farmers' adoption of hybrid seed corn, exposure via structural equivalence did influence adoption. These models also included terms for infection (the degree one's number of nominations received affected adoption) and susceptibility (the degree one's number of nominations sent affected adoption). It was surprising that the network exposure terms were not more significantly associated with adoption, and in Chapter 10 more extensive treatment of these findings and analysis are provided.

It is important to note here, though, how we can use the construction of structural equivalence weights to create general social influence models based on many theoretical ideas of how social influence occurs. The cohesion social influence model states that a person is influenced by the behavior of those they are connected to. The exposure model is:

7-2

$$E_{Nt} = W_{ij} B_{it} A_{ij} +$$

where E is the exposure matrix, W is the network weight matrix, B is the behavior matrix, A is the original adjacency matrix, N is network size, and t is time. The behavior matrix indicates who has adopted the behavior at each time period. Exposure is calculated by multiplying the weight matrix by the adoption matrix and dividing by the number of network choices made. In the basic model, W indicates who is connected to whom (the adjacency matrix). This model can be expanded to include other W s, types of connectivity or similarity.

W , for example, could be structural equivalence scores and exposure calculated on the degree structurally equivalent alters have adopted the behavior. W can be any permutation of A that indicates the degree of similarity (or difference) among the dyads. (Note W can also be constructed from attribute data by calculating the degree dyads are similar on some characteristics such as their education level, ethnicity, or attitude toward a behavior.) Different social influence processes can be modeled by changing W . For example, if the researcher thought that social influence would be stronger between people who shared some characteristics, say gender, then W could be created as a product of network connections and gender such that a tie in W existed only among people who nominated one another and were of the same gender.

Summary

This chapter provided an introduction to defining positions in networks. Positions are often defined as a grouping of nodes in a network based on the pattern of their connections to others in the network. Two people can be **(p.127)** in the same position in the network if they have the same ties to the same others even though they might not be directly connected. There are many different algorithms used to identify structurally equivalent positions, notably the percentage of matches between two nodes or the degree of similarity in their distances to other nodes (structural equivalence).

Once these positions are identified, the researcher may reduce the network to these positions and create what is referred to as an image matrix. In an image matrix, each position is a node in the network and links are created between the nodes to indicate if there is a relationship between the positions. The image matrix provides a macro-structural view of the network. The chapter also discussed CONCOR, a popular algorithm used to find positions in a network.

The chapter closed with a discussion of positions and behavior and how diffusion and positional analysis may be used to understand how behaviors flow through networks. We then discussed how matrices of SE relations may be used to calculate SE-weighted network exposure. Persuasive influence may be a function of people directly connected (cohesion exposure) or may be driven by monitoring of structurally equivalent others, people who occupy the same position in the network.



Access brought to you by: