

Structural Measures for Complete Networks

In: Social Network Analysis and Education: Theory, Methods & Applications

By: Brian V. Carolan

Pub. Date: 2013

Access Date: July 6, 2021

Publishing Company: SAGE Publications, Inc.

City: Thousand Oaks

Print ISBN: 9781412999472

Online ISBN: 9781452270104

DOI: <https://dx.doi.org/10.4135/9781452270104>

Print pages: 97-110

© 2014 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE Research Methods. Please note that the pagination of the online version will vary from the pagination of the print book.

Structural Measures for Complete Networks

Objectives

The primary objective of this chapter is to introduce important concepts and measures that you use to describe static properties of social networks. More specifically, these concepts and measures apply to the analysis of complete networks, those networks for which you have complete relational information on each pair of actors, i and j , that constitute any given network. These measures are derived from relational data that are organized in a matrix whose rows and columns represent actors and whose cell values indicate the presence/strength of a tie. The concepts and measures that are discussed in this chapter reflect a variety of structural properties that have been of interest to network researchers for some time. This chapter allows you to ask a question that has been at the core of most network analyses: What does this network look like?

Understanding these properties will eventually allow you to assess a network's dynamics (Hanneman & Riddle, 2011b). So, if a network of 25 teachers in the same school has few connections (low density), does this structure look the same a year after the introduction of teacher professional communities? To address questions that probe a network's dynamics (i.e., its change over time), it is necessary to first figure out what the network looks like at one point in time, what is commonly referred to as the network's topography. This chapter will provide the conceptual tools and precise measures that are needed to perform this first set of analytical steps.

Example Data

To demonstrate these concepts and measures, this chapter will rely on a fairly simple network data set: Newcomb's Fraternity Data (referred to as Fraternity Data). The original sociometric data collected by Newcomb required each of the 17 actors (all members of the same fraternity) to rank all the others in terms of friendship preferences, ranging from 1 to 16, with 1 indicating first preference. These rankings were done across the entire semester, resulting in 15 separate 17×17 single-mode, directed, and valued matrices. However, to better illustrate these concepts and measures, these data have been transformed to keep things a little simpler. In addition, the focus will be on one of these networks at a single point in time (week 0, the beginning of the study). These recoded data have been dichotomized, with friendship rankings ranging from 1 to 3, now coded as 1, 0 otherwise. Therefore, using the terminology introduced in Chapter 4, the recoded data set is now directed (asymmetrical) and binary (nonvalued). This was done simply for purposes

of presentation; any manipulation of network data should have some theoretical or empirical basis. Table 5.1 shows the recoded data file in the node-list format, with each row starting with the ID number of the responding student followed by three other ID numbers, the alters

Table 5.1 Transformed Fraternity Data in Node-List Format. These binary and directed data consist of 17 students, numbered 1 through 17 (column 1). The next three columns are the ID numbers of the alters who have “received” a tie. For example, Student 2 has sent a tie to Students 4, 7, and 16.

1	11	13	17
2	4	7	16
3	11	12	17
4	2	7	17
5	11	12	17
6	4	8	13
7	4	12	17
8	6	10	11
9	11	12	17
10	1	17	15
11	9	12	17
12	3	11	17
13	1	6	15
14	7	9	10
15	5	10	11
16	4	9	11
17	4	9	12

that are “receiving” friendship nominations. For example, based on the recoding scheme described above, Student 1 sent a friendship nomination Students 11, 13, and 17.

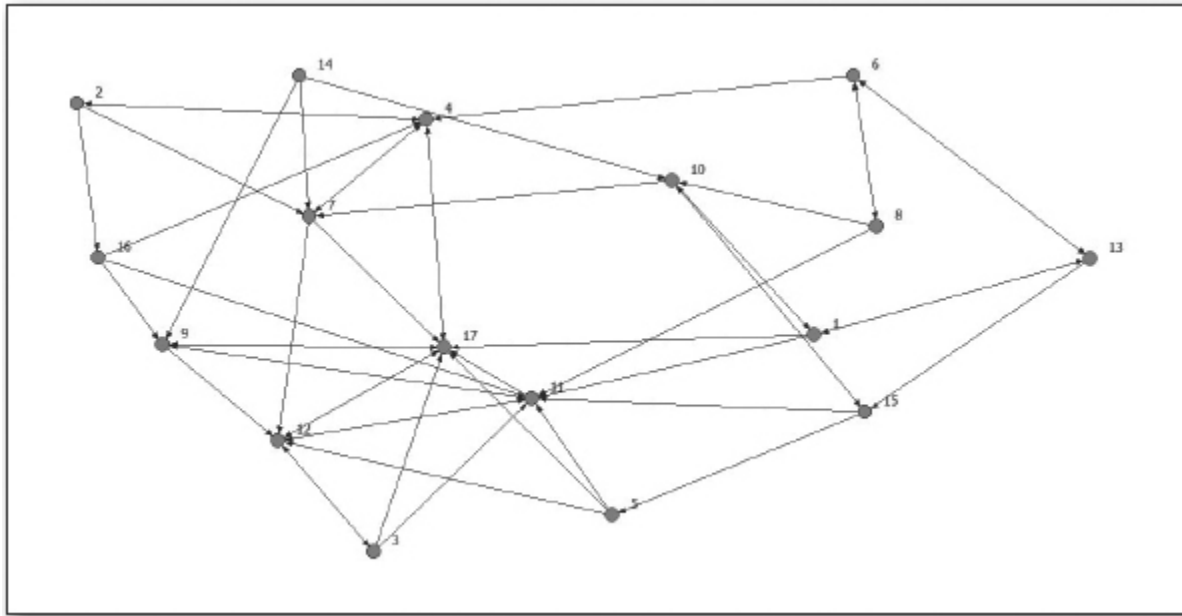
Table 5.2 Transformed Fraternity Data in Square Adjacency Matrix. These are the same data from Table 5.1, but in the form of an adjacency matrix. For example, Student 3 nominated Students 11, 12, and 17 as friends but received only one nomination in return from Student 12.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1
2	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0
3	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1
4	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
5	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1
6	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0
7	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1
8	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1
10	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0
11	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1
12	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1
13	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0
14	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0
15	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0
16	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0
17	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0

These data can also be represented as a sociomatrix, the format in which the data are then processed for analysis. There are several excellent software packages that facilitate these analyses. The measures demonstrated in this chapter have been calculated with UCINET (Borgatti, Everett, & Freeman, 2006), one the most popular and intuitive applications for social network analysis (reviewed along with other software applications in Chapter 12). Table 5.2 shows the recoded Fraternity Data in this matrix format. This matrix shows that, for example, Student 2 sent a friendship nomination to Students 4, 7, and 16 but only received a nomination from Student 4. One can determine this by simply reading the matrix's rows (ties sent) and columns (ties received). These are the very same relational data as seen in the node-list format, Table 5.1.

Before preceding any further, it is helpful to take a look at a directed graph generated from the sociomatrix in Table 5.2. Recall from Chapter 3 that each point on the graph represents a node (student) and each line an arc, that is, a directed tie sent from one node to another. Several striking features of the graph are evident.

Figure 5.1 Graph of Transformed Fraternity Data Generated from the Adjacency Matrix in Table 5.2. From this graph, we can tell that the network is fairly small (size) but pretty well connected (one component), with no isolates (the reason for the latter is obvious, as each student had to rank all the others). Few of the friendships, however, are reciprocal (reciprocity), and some parts of the network are “thicker” than others (density).



When looking from this vantage point at Figure 5.1, it is apparent that the network is fairly small (size) but pretty well connected (one component), with no isolates (the reason for the latter is obvious, as each student had to rank all the others). Few of the friendships, however, are reciprocal (reciprocity), and some parts of the network are “thicker” than others (density). That indicates that there are subgroups within the graph—certain clusters of students who are friends with each other (cliques and clans). There are also a couple of students, Students 11 and 17, for example, who occupy equivalent positions and, consequently, perform the same role in the network (positions and roles).

This chapter introduces the formal algorithms that precisely describe these key features of social networks. Social network and graph theorists have developed an array of indices to characterize these features. These indices rely on some basic—and not-so-basic—mathematical computation. There is no way to avoid this fact without distorting and oversimplifying the advances made by mathematics, specifically graph theory and algebra, to the study of social networks (a point established in Chapter 2). However, an effort is made not to let this fact get in the way of your understanding and application of these concepts and measures. The rest of this chapter introduces the most common ways in which features of complete networks are described, first starting with a set of basic network measures, then, in Chapter 6, moving on to the ways in which subgroups are derived from relational data and, finally, the related ideas of positions and roles.

Network-Level Structural Measures

Network-level structural measures are those that are calculated from the entire network. Therefore, they provide an excellent snapshot of the network's structure—the pattern of relations among the network's actors. This section introduces seven indicators of network structure. In addition to providing a more global view, these indicators can be used to address questions such as whether the friendship, social support, and advice-seeking networks within the same fraternity, for example, are comparable.

Size

The simplest structural property of a social network is its size, an often ignored but important feature. Size plays an important role in determining what happens in the network—what resources are exchanged among actors, for example. In a network of 17 students, it is not hard to imagine that the pattern of who studies with whom, for example, would look much different than if the network consisted of 200 students. Size affects other network measures, but on a conceptual level, it influences the structure of relations, as actors only have so many resources and capacities for creating and maintaining ties with others. Also, in smaller networks, it is easier for actors to know each other. In formal terms, size is simply a measure of the number of nodes in the network. In the case of the Fraternity Data, it is not a terribly interesting property, as it simply reflects the network's boundary.

Density

Density is intricately linked to network size. Density refers to the number of ties in the network reported as a fraction of the total possible number of ties. Equation 5.1 below shows how a network's density is calculated:

$$D = \frac{L}{N(N-1)} \quad (\text{Equation 5.1})$$

where L is the number of lines (directed ties) in the network and N is network size. So, in the Fraternity Data with 17 actors and 51 directed ties, density equals 0.19. The closer this number is to 1.0, the denser the network. A density score of 1.0 indicates that all possible relations are present.

It is also possible to calculate the density of the network's subgroups. For example, you might hypothesize that network density differs by some attribute, GPA, for example. That is, students with denser friendship networks might have access to resources—class notes, information about upcoming assignments, support, and so forth—that may contribute to their academic performance. Density could be calculated on ties among those with GPAs above 3.0 and then compared to those below that threshold.

However, the density of 0.19 reported above is somewhat deceptive, at least in terms of what it means for the Fraternity Data set. This formula needs to be adjusted to account for the limited number of friendship nominations (three) that was imposed when the data were recoded. In a nomination study in which you ask

the respondent to list a certain number of alters (or you impose this limit after the data have been collected), it is more sensible to report what is referred to as effective density, which is the number of lines (ties) multiplied by the number of possible alters:

$$D_E = \frac{L}{N(\lambda)} \quad (\text{Equation 5.2})$$

where L again is the number of lines (directed ties) in the network, N is network size, and λ is the maximum number of alters requested or permitted. Using this formula, the density of the Fraternity Data is 1.0: All possible relations are present, which is unsurprising given that the original ranked data were recoded (1–3 = 1, all others = 0) and each respondent had the maximum number of three friendship nominations.

Reciprocity

A third network measure that indicates an aspect of the networks social structure is reciprocity, defined as the degree to which actors in a directed network select one another. Stated another way, reciprocity indicates the mutuality of the network's ties (this is why reciprocity is also referred to as mutuality). This property is important because it reveals the direction through which resources such as help, advice, and support flow. It also indicates the network's stability, as reciprocated ties tend to be more stable over time. Finally, networks with high reciprocity may be more “equal,” while those with lower reciprocity may be more hierarchical.

While some relations are by definition reciprocal, such as “studied with,” others such as friendship are potentially asymmetrical and indicate imbalances in influence, status, or authority.

Borgatti, Everett, and Freeman (2006) provide the following measure of reciprocity:

$$R = \frac{(A_{ij} = 1) \text{ and } (A_{ji} = 1)}{(A_{ij} = 1) \text{ or } (A_{ji} = 1)} \quad (\text{Equation 5.3})$$

where A_{ij} indicates a tie from actor i to j . A high degree of reciprocity means that a network's actors choose one another. It could also mean that while some actors choose one another, they are not choosing others, which results in a high degree of clustering within the network. At the complete network level, reciprocity is reported as the proportion of reciprocated ties in the network. Therefore, values closer to 1.0 indicate higher reciprocity.

Applying Equation 5.3 to the Fraternity Data, reciprocity for the entire network is 0.31, indicating that slightly less than one-third of all ties are mutual. This low value should be considered unsurprising, as this particular network snapshot was taken at the beginning of the study, well before the participants got to know each other. One could reasonably expect that the network's reciprocity, as well as its density, will increase over time, which would be associated with a greater likelihood of engaging in behavior together and sharing similar attitudes. For example, Moolenaar and Sleegers (2010) hypothesize that teacher networks with high reciprocity would be positively associated with teachers' perceptions of their schools' innovative climate and trust. Their analyses, however, do not support this hypothesis. According to their models, dense networks

(on two different network relations: “discuss work” and “regard as friend”) are significantly related to both a teacher's perceptions of trust and the school's innovative climate, while reciprocity is not. However, despite this null finding, others have shown that networks with high reciprocity are associated problem solving and the exchange of complex knowledge (Uzzi & Spiro, 2005) but are also associated with less desirable outcomes such as risky drug and sexual behaviors (Valente & Vlahov, 2001).

Transitivity

Transitivity is another feature of complete networks that reflects the social structure's tendency toward stability and consistency. Whereas reciprocity focuses on the ties between two actors (dyads), transitivity is based on the triad, any “triple” of actors. The addition of this third “other” reflects a more authentic character of social life and has been the basis for much theorizing about social networks (Simmel, 1908/1950). Much of this theorizing revolves around a triad's tendency toward transitivity: if Teacher A collaborates with Teacher B and Teacher B collaborates with Teacher C, then Teacher A collaborates with Teacher C. In this instance, the triad is considered transitive because Teachers A and B have the same relationship with C.

Valente (2010) notes that early theorizing about this tendency focused on the concept of balance, first introduced by Heider (1958) to explain a person's observed preference for equilibrium with the people around them. The theory of cognitive dissonance grew out of this work, attempting to explain how people felt when their immediate environment was unbalanced (Festinger, 1954). The argument then extended to postulate that increasing the number of transitive triads of which one is a part could reduce cognitive dissonance. Even Granovetter's strength of weak ties theory (1973) emerged from this work, showing that intransitive triads are relatively rare, resulting in networks with fewer weak ties. The paradox, of course, is that the relative paucity of weak ties is what makes them “strong,” as they provide early access to diverse information.

Measuring triadic relations within a network of undirected (symmetric) relations is straightforward (Hanneman & Riddle, 2011b). There are only four possible types of triadic relations: no ties, one tie, two ties, or all three ties. Performing what is called a “triad census” on these data can reveal the extent to which the network can be characterized as consisting of isolates (actors not connected to any other actors), dyads (actors who are only connected to one other actor), structural holes (one actor connected to two others who, in turn, are not connected to each other), and clusters (groups of three actors that are all connected to each other). With undirected data, one can simply count how often each of these four types occurs across all possible triples.

However, things get more interesting when performing a triad census on directed data, such as the transformed Fraternity Data. There are 16 possible types of relations among any three of the network's actors (Holland & Linehardt, 1979). Different forms of relationships can be observed among these groups, including hierarchy, equality, and exclusivity. This is one of the reasons why researchers with an interest in social group relations suggest that the most fundamental forms of social relationships can be observed at the triadic level. A triad census for a network with directed data counts the proportion of triads that fall within each of these 16 categories. These categories are typically called MAN categories, which reflects the number of mutual (M),

asymmetric (A), and null (N) ties in each triad. Therefore, a triad with a code of 111 would have one mutual tie, one asymmetric, and one null tie, indicating that one actor serves as a go-between for the triad's other two actors. Of these 16 possible types of directed triads, only four can be classified as transitive (Wasserman & Faust, 1994). In the Fraternity Data there are, for example, five 030 triads (zero mutual, three asymmetric, and zero null ties); these triads are considered transitive because Students 1 and 2 have the same relationship to 3.

Going one step further, a transitivity analysis divides a network's number of transitive triads by the number of triads of all kinds. More formally, the density of transitive triples is the number of triples, which are transitive, divided by the number of paths of length two (i.e., the number of triples that have the potential to be transitive). In the Fraternity Data, this value is 10.70%. This lack of transitivity suggests that there are numerous null ties, indicating a relatively sparse network in which resources will have difficulty flowing from one part of the network to another. This is not too surprising given that these relations were measured at the start of the study (week 1) and, as a result of the data transformation, each actor has three out-degrees (that is, they sent three friendship nominations). Here, too, this can be attributed to a measurement issue. Faust (2008) has shown that fixed-choice survey instruments (recall that these manipulated data “fixed” the number of friendship nominations to three), artificially limit the triad census and the subsequent analysis of transitivity. Regardless, establishing a network's transitivity is important, as it is theoretically connected to actors' tendencies to divide into exclusive subgroups over time, especially relational data that capture positive affect, such as friendship.

Diameter and Distance

Similarly, the next two network-level properties, diameter and distance, also indicate how well resources can move from one part of the network to another. Whereas transitivity focuses on the importance of certain configurations of triads, these two related concepts are more straightforward. A network's diameter refers to the longest path between any two actors. This property is important, as networks that have the same size (equal numbers of actors) and even the same density (equal percentages of ties present) can have different diameters. Consider the network's diameter in Figure 5.1. The length of the longest path between two actors is five. To “get from” Student 15 to 16 in this directed network requires five steps: 15 → 5 → 17 → 4 → 2 → 16. This is the only five-step path in the network and is the maximum distance between any two actors.

Related to this is another property that captures how “far” actors are from one another. The average path length measures the mean distance between all pairs of actors in the network. When this value is small, this indicates that there is a cohesive network with minimal clustering. Conversely, when this value is high, the network likely has little cohesion, thereby making it difficult for resources—whether these resources be advice, trust, information, or the like—to move from one part of the network to another. The average path length in the Fraternity Data is 2.18. The interpretation of this is that, on average, all students in the network are slightly more than two steps from everyone else in the network, assuming they could reach one another (an assumption that is incorrect given that there are no paths between some pairs of students). To calculate

this, you first have to identify the distances between each pair of actors in the network and then calculate the average of all these pairs.

Both diameter and average path length are important network-level structural properties. Valente (2010) goes as far as to note that you need not know anything else about a network's topography to draw several important conclusions. For example, a network with a large diameter and small average path length suggests a structure in which there are parts of the network that some network actors may be unable to access. This is evident for Student 15 who, in a network with a total of 17 actors, needs five steps to reach another member (Student 16).

Clustering

Another structural network-level property that is of interest to network researchers is a network's clustering. Clustering is a measure of a network's actors' tendency to "group together" into pockets of dense connectivity (Valente, 2010). High clustering indicates that there are numerous pockets in which some actors are connected to each other but not to others. Low clustering, on the other hand, suggests that relations are more evenly distributed across the network with very few pockets of dense connectivity among subsets of actors. The tendency, of course, is for people to interact with a fairly small set of actors who share similar attributes—the "birds of a feather flock together" (homophily) phenomenon. Knowing the degree to which actors cluster says a great deal about the structure of these everyday relations, a structural pattern that may even seem somewhat paradoxical (Watts, 1999). This is related, in part, to what is referred to as the small-world phenomenon.

The small-world phenomenon combines two ideas (Valente, 2010). The first is a network with a short average path length. That is, relatively few steps connect each pair of actors, on average. The second part of this is that "small-world" networks also have a high degree of clustering—groups of actors who interact almost exclusively in their own immediate "neighborhoods." This second part is the network's tendency to form dense local pockets of connectivity: clusters. "Small worlds" are those that paradoxically have a low average path length but high clustering.

Measuring the extent to which a network displays clustering involves two steps. First, one calculates the density of each actor's local neighborhood (the density of each ego's network, which equals the total number of ties present divided by the total number of possible ties). Then, after doing this for all actors in the network, the degree of clustering can be characterized by the average of all the ego neighborhoods in the complete network. For example, the Fraternity Data network has a network-level clustering coefficient of 0.24, which reflects the average neighborhood densities of all 17 actors. This seems pretty low; students are surrounded by local neighborhoods that are sparse. However, this should be considered in light of the network's overall density, which was 0.19 (Equation 5.1). So the average density of these local neighborhoods (clustering) is not much different than the density of the entire network. There is also a weighted version of the measure that accounts for the size of each ego's neighborhood. That is, actors with more alters in their immediate

neighborhoods get more “weight” in computing the average density. When computing this weighted version of the Fraternity Data, there is little difference with the nonweighted version.

Centralization

A final structural property of complete networks is centralization. A network that is highly centralized is one in which relations are focused on one or a small set of actors. This is, however, different than a structural property such as density, which measures the presence of relations, not whether they are focused on a small set of actors. In other words, a network can be dense but have low centralization (many relations that are spread evenly across the network's actors) and vice versa (few relations that are concentrated on a small set of actors). Networks that are centralized, regardless of their density, are ones in which only a small and exclusive set of actors hold positions of power and control. Decentralized networks, conversely, are those in which power and control are diffuse and spread over a number of actors.

Centralization can be based on a number of metrics (these are discussed in Chapter 7 when individual ego centrality measures are introduced), with degree being one of the most common (Freeman, 1979). *Degrees* refers to the number of ties an actor either sends (out-degree) or receives (in-degree). Valente (2010) summarizes the calculation of network centralization. Network centralization is calculated by determining the highest individual centrality score (either in-degree or out-degree) and then subtracting it from all the other individual scores in the network. Next, these differences are added and that total is divided by the maximum sum of differences theoretically possible in a network of that size. Freeman's (1979) formula for centralization degree (CD) is:

$$CD = \frac{\sum (Max(C_{Di}) - C_{Di})}{n^2 - 3n + 2} \quad (\text{Equation 5.4})$$

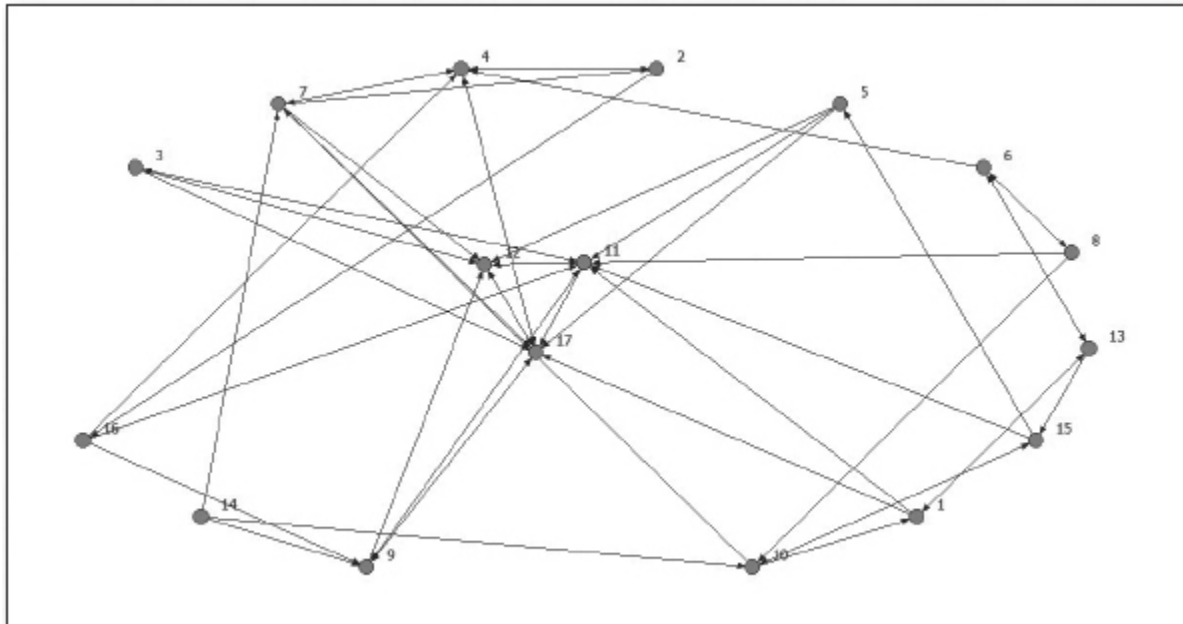
where $Max(C_{Di})$ is the network's maximum centrality score, C_{Di} indicates individual actors' centrality scores, and n is the network's size. This formula can also be applied to the other types of centrality, including betweenness and closeness, to be discussed in Chapter 7.

Applying Equation 5.4 to the Fraternity Data, the network's (in-degree) centrality score is 33.20%. Figure 5.2 shows this network with a different layout. This circle layout includes three students (numbers 12, 11, and 17), who receive the most friendship nominations (six, eight, and eight in-degrees, respectively), who are at the center of the circle. This layout demonstrates just how “central” these actors are to the network. However, this overall centralization score is not too high, indicating that friendship nominations are spread more or less evenly across the network's actors. Why is this property important?

A network's centralization affects the process through which resources traverse the network. Imagine if Students 12, 11, and 17 have information about what was to be included on an upcoming exam. Their location can either accelerate or prevent the spread of that information to other students. These central actors likely wield a disproportionate amount of influence on the network. Therefore, high centralization provides fewer

actors with more power and control.

Figure 5.2 Circle Layout of Fraternity Members Network With Three Central Students. This circle layout includes three students (numbers 12, 11, and 17) who receive the most friendship nominations (six, eight, and eight in-degrees, respectively), who are at the center of the circle. This layout demonstrates just how “central” these actors are to the network.



Summary

This chapter introduced structural measures that you can be calculated on complete network data. The introduction of these concepts, definitions, and measures parallels the analytic process that is often employed in network studies. First, you identify the main structural properties of a network, including those related to its size, density, and connectivity. After performing the analyses covered in this chapter, you will typically move on to a group or positional analysis, which is addressed in Chapter 6. Although these structural measures for complete networks are considered a baseline for most analyses, their use should be guided by questions of theoretical interest. Each measure will give you a slightly different perspective on the network's structure, so it is best to use measures that reflect your theoretical interests.

Chapter Follow-Up

Assume you have complete network-level data on school leaders in a large urban district that is transitioning to a new teacher-evaluation system. You have relational data on the frequency with which ego discusses this new system with each alter (0 = never, 1 = sometimes, 2 = regularly, and 3 = frequently) and whether ego turns to alter for advice regarding general professional matters

(1 = yes, 0 = no). Which structural properties of the complete network might be of interest to you? Please explain why these properties might be of interest.

Given the same network described above, what would high centralization scores on both relations indicate about this network's ability to successfully transition to a new evaluation system?

How might your response to #2 differ if you knew that the networks also had high density scores? Given this new information, what would you predict about the transition to a new evaluation system?

Essential Reading

Freeman, L. C. Centrality in social networks. *Social Networks*, (1979).1,215–239.

Watts, D. Networks, dynamics, and the small-world phenomenon. *American Journal of Sociology*, (1999).105(2),493–527.

Wellman, B. (1988). Structural analysis: From method and metaphor to theory and substance. In Edited by: **B. Wellman & S. D. Berkowitz** (Eds.), *Social structures: A network approach* (pp. 19–61). New York: Cambridge University Press.

<http://dx.doi.org/10.4135/9781452270104.n5>