

Center and Periphery

6.1 Introduction

In this chapter, we present the concepts of **centrality** and **centralization**, which are two of the oldest concepts in network analysis. Most social networks contain people or organizations that are central. Because of their position, they have better access to information and better opportunities to spread information. This is known as the *ego-centered approach* to centrality. Viewed from a *sociocentered perspective*, the network as a whole is more or less centralized. Note that we use **centrality** to refer to positions of individual vertices within the network, whereas we use **centralization** to characterize an entire network. A network is highly centralized if there is a clear boundary between the center and the periphery. In a highly centralized network, information spreads easily but the center is indispensable for the transmission of information.

In this chapter, we discuss several ways of measuring the centrality of vertices and the centralization of networks. We confine our discussion of centrality to undirected networks because we assume that information may be exchanged both ways between people or organizations that are linked by a tie. Concepts related to importance in directed networks, notably prestige, are discussed in Part IV of this book.



6.2 Example

Studies of organizations often focus on informal communication: who discusses work matters with whom and to whom do people turn for advice? Informal communication is important to the operation of the organization and it does not always coincide with the organization's formal structure. For the diffusion and retrieval of information, it is crucial to know the people who occupy central positions in the communication network.

Our example is a communication network within a small enterprise: a sawmill. All employees were asked to indicate the frequency with which they discussed work matters with each of their colleagues on a 5-point

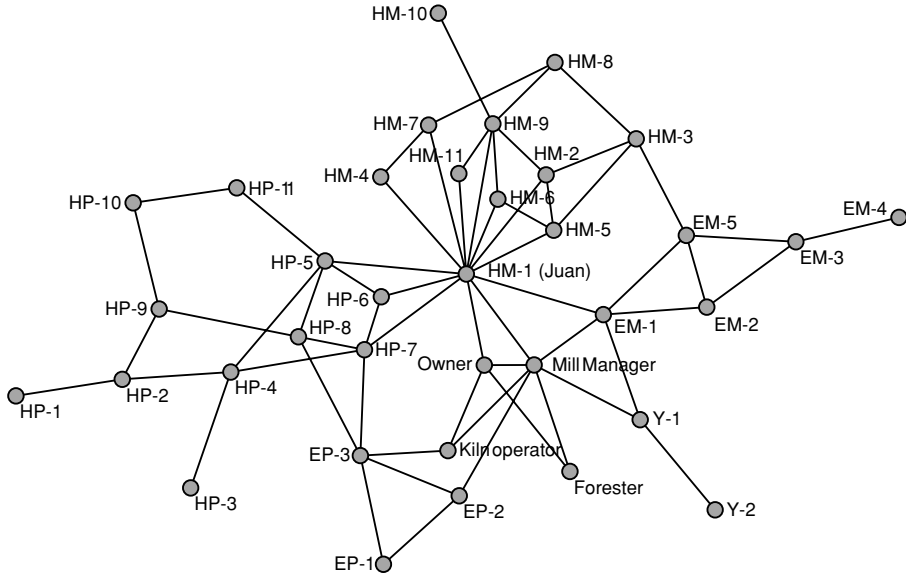


Figure 58. Communication ties within a sawmill.

scale ranging from less than once a week to several times a day. Two employees were linked in the communication network if they rated their contact as three or more. We do not know whether both employees had to rate their tie in this way or that at least one employee had to indicate a strength of three or more. The network is stored in the file *Sawmill.net*.

In the sawmill, the employees are Spanish speaking (H) or English speaking (E), which, of course, is relevant to their communication. The sawmill contains two main sections: the mill (M), where tree trunks are sawn into logs, and the planer section (P), where logs are planed. Then there is a yard (Y) where two employees are working and some managers and additional officials.

Figure 58 shows the communication network in the sawmill. Note that vertex labels indicate the ethnicity and the type of work of each employee, for example, HP-10 is an Hispanic (H) working in the planer section (P). In this figure, vertex labels instead of vertex colors identify the attributes of employees. It is quite easy to see that work-related communication is structured along work section (planers at the left, sawyers at the right) and ethnicity: Hispanics at the top and English-speakers at the bottom – assuming that management, forester, kiln operator, and employees in the yard are English-speakers.

Intuitively, HM-1 (Juan) is a central, perhaps the most central, person in this network. He communicates with many colleagues directly and through his direct contacts it is easy for him to reach most of the people working in the sawmill. Juan seems to occupy a pivotal position in the flow of information between the planers, the mill section, and management. This chapter presents formal measures of centrality and centralization, which capture these intuitions.

6.3 Distance

One approach to centrality and centralization is based on the simple idea that information may easily reach people who are central in a communication network. Or, to reverse the argument, people are central if information may easily reach them.

The larger the number of sources accessible to a person, the easier it is to obtain information; for instance, an elderly person will acquire information about where to look for help more easily if his or her social support network is larger. In this sense, social ties constitute a *social capital* that may be used to mobilize social resources. Hence, the simplest indicator of centrality is the number of its *neighbors*, which is his or her degree in a simple undirected network (see Chapter 3). The higher the degree of a vertex, the more sources of information it has at its disposal, the quicker information will reach the vertex, so the more central it is. In the sawmill network, Juan communicates with no fewer than thirteen colleagues, whereas the manager of the mill has only seven communication ties (Figure 58). In this respect, Juan is more central than the manager and information from the shop floor will reach him more easily than it will reach the manager. If degree is the simplest measure of the centrality of a vertex, what is the associated measure of centralization for the entire network, which expresses the extent to which a network has a center? Let us first answer another, related question: Given a fixed number of lines, what is the most efficient structure to exchange information? We should note that this network must be connected, otherwise information cannot reach all vertices. In this case, the *star-network* is known to be the most efficient structure given a fixed number of lines. A star is a network in which one vertex is connected to all other vertices but these vertices are not connected among themselves (e.g., network A in Figure 59).

Compare the star-network in Figure 59 to the line-network, containing the same number of vertices and lines (network B). It is much easier to identify the central vertex in the star-network than in the line-network because the difference between the central vertex (v5) and the peripheral vertices (v1, v2, v3, and v4) is much more apparent than in the line-network. This leads to an idea that may be counterintuitive, namely that a network is more centralized if the vertices vary more with respect to



Figure 59. Star- and line-networks.

their centrality. More variation in the centrality scores of vertices yields a more centralized the network.

Now we can define degree centralization as the variation in the degree of vertices divided by the maximum variation in degree which is possible given the number of vertices in the network. In a simple network of a particular size, the star-network has maximum degree variation. The division by maximum degree variation ensures that degree centralization ranges from zero (no variation) to 1 (maximum variation) in the case of a star-network.

The **degree centrality** of a vertex is its degree.

Degree centralization of a network is the variation in the degrees of vertices divided by the maximum degree variation which is possible in a network of the same size.

Variation is the summed (absolute) differences between the centrality scores of the vertices and the maximum centrality score among them. In network A (Figure 59), for instance, one vertex (v5) has degree 4, which is the maximum degree in a simple undirected network of this size because this vertex is connected to all other vertices. The other four vertices have minimum degree, which is 1 in a connected undirected network. Hence, the degree variation amounts to 12: (vertices v1 to v4 contribute) $4 \times (4 - 1)$ and (vertex v5 contributes) $1 \times (4 - 4)$. In a simple undirected network, the degree of vertices cannot vary more than this, so 12 is the maximum variation and dividing 12 by itself, of course, yields a degree centralization of 1.00.

In network B, two vertices have a degree of 1 (v1 and v2) and the other vertices have a degree of 2. Because 2 is the maximum degree in this network, the degree variation equals $2 \times 2 - 1$ (for vertices v1, v2) and $3 \times 2 - 2$ (for vertices v3 to v5), which is 2. To obtain the degree centralization of network B, we divide 2 by 12, which is the maximum variation in a simple undirected network, and we obtain 0.17. If we add a line between v1 and v2, degree centralization becomes minimal (0.00) because all vertices have equal degree, so variation in degree is zero and degree centralization is zero.

We should issue a warning here. In a network with multiple lines or loops, the degree of a vertex is not equal to the number of its neighbors. Therefore, the star-network does not necessarily have maximum variation and we may obtain centralization scores over 1.00 if we compare the variation in a network with multiple lines or loops to the variation in a simple star-network of the same size. In this case, we advise not using degree centralization.

In a simple undirected network, degree centrality is just the number of neighbors of a vertex. In some cases, this is all we know about the network position of people, for instance, when data are collected by means of a survey in which people are asked to indicate the size of their personal network. If we want to analyze the communication structure of the network, however, we need to know who is connected to whom in the

entire network and we must pay attention to indirect ties because information can flow from one person to the next and on to other people. In a communication network, information will reach a person more easily if it does not have to “travel a long way.” This brings us to the concept of *distance* in networks, namely the number of steps or intermediaries needed for someone to reach another person in the network. The shorter the distance between vertices, the easier it is to exchange information.

In Chapter 3, we defined paths as a sequence of lines in which no vertex in between the first and last vertices occurs more than once. Via a path, we can reach another person in the network: we can inform our neighbor, who passes the information on to his neighbor, who in turn passes it on, until the information finally reaches its destination. We say that a person is *reachable* from another person if there is a path from the latter to the former. Note that two persons are mutually reachable if they are connected by a path in an undirected network, but that two paths (one in each direction) are needed in a directed network.

In an undirected network, the distance between two vertices is simply the number of lines or steps in the shortest path that connects the vertices. A shortest path is also called a *geodesic*. In a directed network, the geodesic from one person to another is different from the geodesic in the reverse direction, so the distances may be different. This sounds strange if you are used to geographic distances but think of a directed network as a system of one-way streets: it is easy to imagine that the route from A to B differs from the journey back. In this chapter, however, we use only undirected networks, so you do not have to worry about this now.

A *geodesic* is the shortest path between two vertices.

The *distance* from vertex u to vertex v is the length of the geodesic from u to v .

With the concept of distance, we can define another index of centrality, which is called closeness centrality. The closeness centrality of a vertex is based on the total distance between one vertex and all other vertices, where larger distances yield lower closeness centrality scores. The closer a vertex is to all other vertices, the easier information may reach it, the higher its centrality.

Just like degree centralization, we can conceptualize closeness centralization as the amount of variation in the closeness centrality scores of the vertices. Again, we compare the variation in centrality scores to the maximum variation possible, that is, the variation in closeness centrality in a star-network of the same size.

The *closeness centrality* of a vertex is the number of other vertices divided by the sum of all distances between the vertex and all others.

Closeness centralization is the variation in the closeness centrality of vertices divided by the maximum variation in closeness centrality scores possible in a network of the same size.

In star-network A (Figure 59), vertex v5 has maximum closeness centrality because it is directly linked to all other vertices. The sum of distances to the other vertices is minimal, namely four geodesics of length 1 combine into a summed distance of 4. Because there are four vertices other than v5, the closeness centrality of vertex v5 is maximal: $4/4 = 1.00$. The other vertices of network A have a closeness centrality score that is considerably lower (0.57) because three vertices are two steps away from them.

In network B, v5 also has the highest closeness centrality because it is in the middle, but now its closeness centrality is not maximal (0.67) and it differs less from the other vertices, which have closeness centrality 0.57 (vertices v3 and v4) and 0.40 (v1 and v2). Because the variation of closeness centrality scores in network B is less than in network A, network B is less centralized. Its closeness centralization is 0.42, whereas the maximum centralization of network A is 1.00.

Note that complications arise if the network is not (strongly) connected. If an undirected network is not connected or a directed network is not strongly connected, there are no paths between all vertices, so it is impossible to compute the distances between some vertices. The solution to this problem is to take into account only the vertices that are reachable to or from the vertex for which we want to calculate closeness centrality and weight the summed distance by the percentage of vertices that are reachable. This solution works fine for the closeness centrality of vertices. However, it does not allow us to compute the closeness centralization of the entire network because the star-network does not necessarily have the highest variation in closeness centrality scores if the network is not (strongly) connected. Therefore, we do not use closeness centralization in the case of a network that is not (strongly) connected.

Application

*Net>Partitions>
Degree*

*Net>Transform>
Edges→Arcs*

*Net>Transform>
Remove>
multiple lines*

*Net>Transform>
Remove>loops*

In Chapter 3, we explained how to compute the degree of vertices in Pajek. Note that the *Net>Partitions>Degree>All* command counts edges only once, which is fine if the network is undirected. In a network containing edges and arcs, however, you may want to count the edges as incoming and outgoing arcs. If so, replace the edges by bidirectional arcs (*Net>Transform>Edges→Arcs*) before you calculate degree with the *All* command. In addition, we advise removing multiple lines (*Net>Transform>Remove>multiple lines*) and loops (*Net>Transform>Remove>loops*) from the network before you compute degree centrality and degree centralization.

The degree partition contains the degree centrality scores of the vertices and the normalized degree vector contains the degree centrality of the vertices expressed as a proportion of the maximum degree, which is the number of other vertices in the network. If you inspect these scores, you will see that Juan has thirteen neighbors, which we already noted in Figure 58, which is more than one-third (0.37) of all thirty-five people with whom he can be connected. In addition to the degree partition, Pajek automatically writes the degree centralization of the entire network to the Report screen provided that the network contains no multiple lines and

loops. Otherwise Pajek reports a message explaining that degree centralization is not valid for this network. Here, the degree centralization of the network is 0.289, which is meaningful only in comparison to other networks.

If you want to know the distance between one vertex (e.g., Juan) and all other vertices in the network, you can use the commands in the *Net>k-Neighbours* submenu, which create a partition of classes containing the distances between one vertex and all other vertices. The *Input* command calculates the distances *from* the selected vertex, whereas the *Output* command computes distances *to* the vertex. The *All* command disregards the direction of the lines. In an undirected network, you may choose the *Input*, *Output*, or *All* commands: they yield the same results. The *From Cluster* command is useful if you want to compute distances for a subset of vertices in a large network.

Net>
k-Neighbours

When you execute a *k-Neighbours* command, you must first specify the vertex number or the label of the vertex from which distances will be computed. In the case of Juan, enter 12 (his vertex number) or HM-1 (the start of his vertex label). Next, you can set a limit to the maximum distance that will be computed. In very large networks, setting a limit may speed up computation considerably. In this dialog box, 0 means that you want all distances, which is usually the right choice in the case of a small network. The distances are stored in a partition and unreachable vertices or vertices further away than maximum distance are placed in class number 9999998, which indicates that their distance is not known.

In Figure 60, vertex colors and class numbers indicate the distances between Juan and other employees. Most employees are directly connected to Juan (black or yellow) or indirectly connected with one intermediary (light gray or green, distance 2). Two employees are four steps away from Juan, namely HP-1 and EM-4.

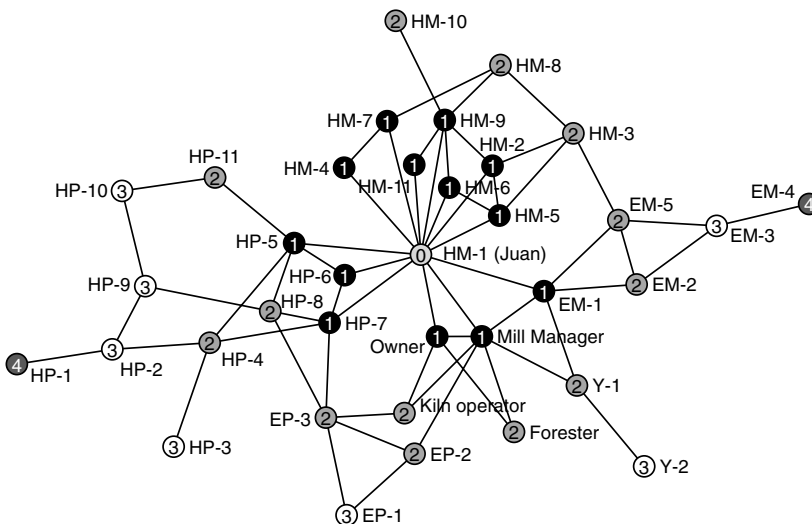


Figure 60. Distances to or from Juan (vertex colors: Default Grey-Scale 1).

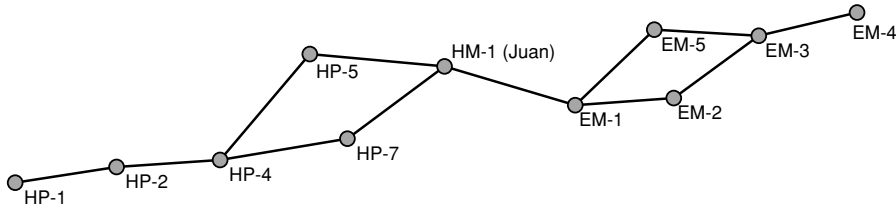


Figure 61. Geodesics between HP-1 and EM-4.

*Net > Paths
between 2
vertices > All
Shortest*

Employees HP-1 and EM-4 seem to be furthest apart in the communication network because their distance to Juan is 4. However, their geodesic does not necessarily include Juan, so they may be connected in fewer than eight steps. In Pajek, the geodesics between two vertices can be found with the command *Net > Paths between 2 Vertices > All Shortest*. In the first dialog box, enter the vertex number or label of HP-1. In the second dialog box, enter EM-4 and subsequently answer *Yes* to the question “Forget values of lines?” because you do not want to weight the lines by their values. This is the right thing to do unless line values indicate distances, for instance, geographical distances. Finally, a dialog box asks whether the paths must be identified in the source network. If you answer *Yes* to this question, Pajek produces a partition for the original network that assigns vertices on the geodesics to class 1 and other vertices to class zero. Regardless of your choice in this dialog box, Pajek creates a new network with the vertices and lines that constitute the geodesics (Figure 61). In addition, it prints the distance in the Report screen. In our example, all geodesics between HP-1 and EM-4 include Juan (see Figure 61), so the distance between HP-1 and EM-4 cannot be less than 8, which is the sum of their distances to Juan.

*Net > Vector >
Centrality >
Closeness*

In Pajek, the computation of closeness centrality is straightforward. Because closeness centrality scores are continuous rather than discrete, the centrality commands are located in the *Net > Vector* submenu. The *Net > Vector > Centrality* submenu has commands to compute closeness centrality for all vertices in the network. For undirected networks, you may choose the commands *Input*, *Output*, or *All*, which yield the same results. If the network is not (strongly) connected, Pajek creates a vector with closeness centrality scores but it does not compute closeness centralization, which is undefined in such a network. Closeness centrality of vertices that are not reachable to or from all other vertices is set to zero. For medium-sized and large networks, closeness centrality demands a lot of computing time so it should be applied with care.

Pajek creates a vector with the closeness centrality scores of the vertices. You may inspect this vector or use it for computations in the ways explained in previous chapters. In our example, closeness centrality scores range from 0.20 to 0.51 and Juan (0.51) turns out to be more central than the manager (0.42). In addition, Pajek computes the closeness centralization of the network, which is printed in the Report screen. The sawmill communication network has a closeness centralization score of 0.38, which, again, must be interpreted in comparison to other networks.

Exercise 1

What will happen to the network if Juan (HM-1) disappears? Remove the vertex, compute closeness centrality and centralization, and interpret the results.

6.4 Betweenness

Degree and closeness centrality are based on the reachability of a person within a network: How easily can information reach a person? A second approach to centrality and centralization rests on the idea that a person is more central if he or she is more important as an intermediary in the communication network. How crucial is a person to the transmission of information through a network? How many flows of information are disrupted or must make longer detours if a person stops passing on information or disappears from the network? To what extent may a person control the flow of information due to his or her position in the communication network?

This approach is based on the concept of betweenness. The centrality of a person depends on the extent to which he or she is needed as a link in the chains of contacts that facilitate the spread of information within the network. The more a person is a go-between, the more central his or her position in the network. If we consider the geodesics to be the most likely channels for transporting information between actors, an actor who is situated on the geodesics between many pairs of vertices is very important to the flow of information within the network. This actor is more central.

Juan, for instance, is important to the communication between HP-1 and EM-4 in the sawmill because all (four) geodesics include Juan (Figure 61). In contrast, HP-5 and HP-7 or EM-2 and EM-5 are less important because if one fails to pass on information, the other may fulfill this role and the communication chain between HP-1 and EM-4 is still intact.

Each pair of vertices may contribute to the betweenness centrality of a vertex. HP-5 and EM-1, for example, contribute to the betweenness centrality of Juan, because their geodesic includes Juan. In contrast, the pair HP-4 and HP-5 does not contribute to Juan's betweenness centrality, because he is not included in their geodesic. In general, we may say that the betweenness centrality of a vertex is the proportion of all geodesics between other vertices in the network that include this vertex. Betweenness centralization is the ratio of the variation in betweenness centrality scores to the maximum variation.

The *betweenness centrality* of a vertex is the proportion of all geodesics between pairs of other vertices that include this vertex.

Betweenness centralization is the variation in the betweenness centrality of vertices divided by the maximum variation in betweenness centrality scores possible in a network of the same size.

It is easy to see that the center of a star network (vertex v5 in Figure 59) has maximum betweenness centrality: all geodesics between pairs of other vertices include this vertex. In contrast, all other vertices have minimum betweenness centrality (0) because they are not located between other vertices. The centrality scores of vertices in a star have maximum variation, so the betweenness centralization of the star is maximal: remove its central vertex and all communication ties are destroyed. In the line-network (B in Figure 59), removal of a vertex may also break the flow of information, but parts of the chain remain intact. Therefore, centrality indices vary less than in the star-network and betweenness centralization is lower.

Net>Vector>
Centrality>
Betweenness

Application

The *Betweenness* command in the *Net>Vector>Centrality* submenu creates a vector of betweenness centrality scores for the vertices in the network. In addition, the betweenness centralization of the network is printed in the Report screen. In directed networks, the procedure automatically searches for directed paths, so there are no separate commands for *Input*, *Output*, and *All*. Even in unconnected networks, betweenness centrality can be computed.

The sawmill communication network has a betweenness centralization of 0.55. Betweenness centrality scores of employees range from 0.00 to 0.59. In Figure 62, vertex size indicates betweenness centrality. Several vertices are invisible because they have zero betweenness centrality: they do not mediate between other vertices. In this example, the betweenness centrality of vertices varies more than their closeness centrality because vertices at the outer margin of the network have zero betweenness, whereas they are still close to part of the network. As a consequence, betweenness centralization is higher than closeness centralization.

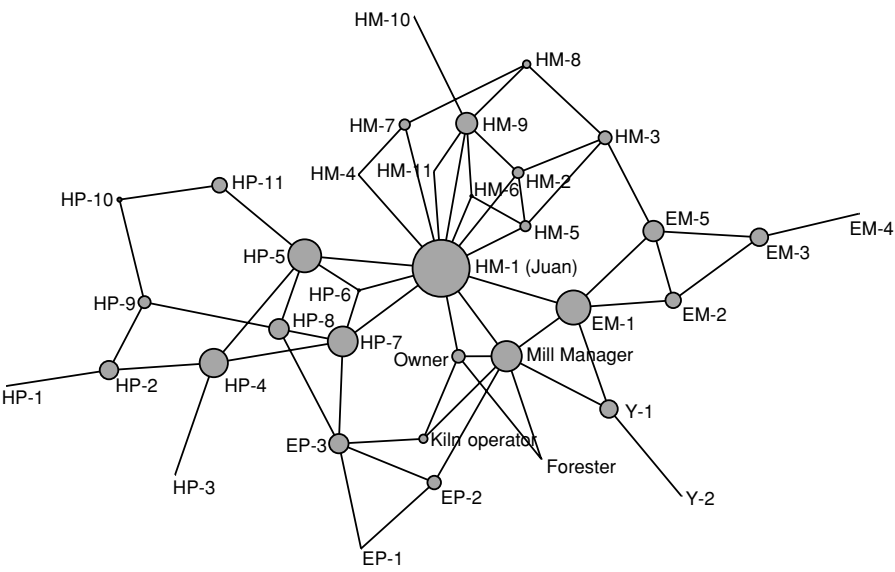


Figure 62. Betweenness centrality in the sawmill.

It is interesting to note that Juan (0.59), EM-1 (0.21), and HP-5 (0.20) are more central than the manager of the mill (0.17). Each ethnic group within the mill's departments – with the exception of the English-speaking planers – seems to have an informal spokesman who is taking care of the communication with other departments or ethnic groups. Juan, who is the spokesman of the Hispanic employees at the mill, is clearly most central.

Exercise II

Compute betweenness centrality on the sawmill network and draw it with vertex sizes corresponding to the betweenness centrality like Figure 62.

6.5 Summary

The concepts of vertex centrality and network centralization are best understood by considering undirected communication networks. If social relations are channels that transmit information between people, central people are those who either have quick access to information circulating in the network or who may control the circulation of information.

The accessibility of information is linked to the concept of distance: if you are closer to the other people in the network, the paths that information has to follow to reach you are shorter, so it is easier for you to acquire information. If we take into account direct neighbors only, the number of neighbors (the degree of a vertex in a simple undirected network) is a simple measure of centrality. If we also want to consider indirect contacts, we use closeness centrality, which measures our distance to all other vertices in the network. The closeness centrality of a vertex is higher if the total distance to all other vertices is shorter.

The importance of a vertex to the circulation of information is captured by the concept of betweenness centrality. In this perspective, a person is more central if he or she is a link in more information chains between other people in the network. High betweenness centrality indicates that a person is an important intermediary in the communication network. Information chains are represented by geodesics and the betweenness centrality of a vertex is simply the proportion of geodesics between pairs of other vertices that include the vertex.

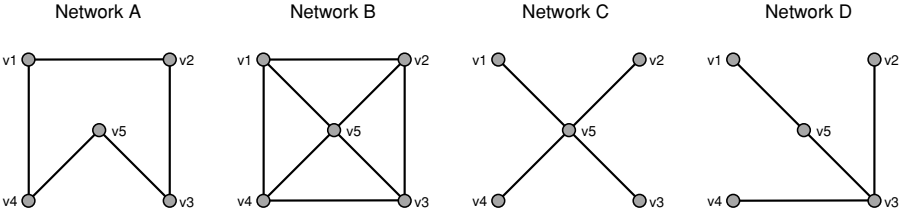
The centralization of a network is higher if it contains very central vertices as well as very peripheral vertices. Network centralization can be computed from the centrality scores of the vertices within the network: more variation in centrality scores means a more centralized network. There is an index of network centralization for each measure of centrality but some centralization measures need special networks: degree centralization is applicable only to networks without multiple lines and loops, and closeness centralization requires a (strongly) connected network.

In this book, we apply centrality and centralization only to undirected networks. It is easy to devise centrality measures for directed networks. We could base degree centrality on the outdegree of vertices, compute closeness centrality from the distances from a vertex to all other vertices

(and not in the reverse direction), and consider only shortest directed paths in the case of betweenness centrality. In fact, other books on social network analysis advocate such an approach. We think, however, that it is conceptually more clear to restrict centrality and centralization to undirected networks and to apply other concepts (e.g., prestige) to directed networks.

6.6 Questions

1. Which of the following statements is correct?
 - a. The centrality of a network equals the degree of its vertices.
 - b. Centralization depends on the variation of centrality scores.
 - c. A single vertex is always the center of a network.
 - d. The center of a network is always a cohesive subgroup.
2. Put the four networks (below) in order of ascending centralization.



- a. B, A, C, D
 - b. A, D, B, C
 - c. D, A, B, C
 - d. A, B, D, C
3. Manually compute the closeness centrality of vertices v1 and v3 in network D of Question 2.
 4. The betweenness centrality of vertex v3 in network D of Question 2 is 0.83. List the geodesics that include v3 and the geodesics that do not.
 5. The file `question5.net` contains the nominations made by thirty-two employees of an organization who were asked to name the colleagues with whom they discuss work matters. Note that not all nominations are reciprocated. Omit the unilateral nominations, which are less reliable, from the network and find out who is most central in this communication network.

6.7 Assignment

In the 1970s, Rogers and Kincaid studied the diffusion of family planning methods in twenty-four villages in the Republic of Korea. The files `Korea1.net` and `Korea2.net` contain the communication networks among women in two villages: a village with a successful family planning program (`Korea1.net`) and a village in which family planning was not adopted widely (`Korea2.net`). In both networks, a line indicates that two women discussed family planning. In addition,

we know which women adopted family-planning methods at least temporarily (class 1 in the partitions `Korea1-adopters.clu` and `Korea2-adopters.clu`) and which women were members of the local Mothers' Club, which played an important role in the diffusion of family-planning methods (class 1 in the partitions `Korea1-members.clu` and `Korea2-members.clu`) in both networks. The project file `Korea.paj` contains all files.

Analyze the networks and find whether centrality and centralization are associated with the success of the family-planning program in one village and its relative failure in the other village. Explain the effects of centrality and centralization by discussing the role of communication in the adoption of family planning methods.

6.8 Further Reading

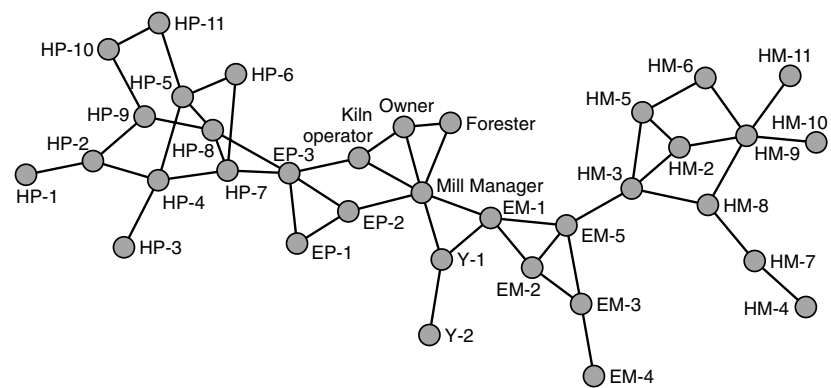
- The sawmill example is taken from J. H. Michael and J. G. Massey, "Modeling the communication network in a sawmill." In: *Forest Products Journal* 47 (1997), 25–30.
- The data on the Korean villages stem from E. M. Rogers and D. L. Kincaid, *Communication Networks: Toward a New Paradigm for Research* (New York: The Free Press, 1981), which offers an overview over network analysis from the perspective of communication studies. Note that some of the methods and software packages discussed in the book are obsolete.
- Many more measures of centrality have been proposed, notably information centrality, which considers all paths and not just geodesics in computing a betweenness score, and eigenvector centrality or power, which takes into account how central those are to whom an ego is connected. Read more about these and other measures of centrality in A. Degenne and M. Forsé, *Introducing Social Networks* (London: Sage, 1999, Chapter 6), J. Scott, *Social Network Analysis: A Handbook* (London: Sage (2nd ed. 2000), 1991, Chapter 5), and S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications* (Cambridge: Cambridge University Press, 1994, Chapter 5).
- The method for calculating closeness centrality on networks that are not (strongly) connected was proposed by G. Sabidussi ["The centrality index of a graph." In *Psychometrika* 31 (1966), 581–603].

6.9 Answers

Answers to the Exercises

- Remove Juan by creating a partition identifying him in one class and then extracting the other vertices from the network. When Juan's vertex is deleted, the network changes considerably (see figure below). Now, the mill's manager is most central (his or her closeness centrality

score is 0.29). The network's closeness centralization decreases to 0.16 and the planers (left) are clearly separated from the mill (right). It is interesting to see that the English-speaking employees are more central, because they are closer than the Spanish-speaking employees to the manager. Juan seemed to function as the informal Hispanic manager.



II. Section 6.4 explains how to compute betweenness centrality. To obtain a sociogram such as that in Figure 62, draw it with the command *Draw>Draw-Vector* (or *Ctrl-u*) from the Main screen.

Answers to the Questions in Section 6.6

1. Answer b is correct: network centralization measures the variation of the centrality of the vertices within the network. The more variation, the easier it is to distinguish between the center and the periphery, the more centralized the network. Answer a is incorrect because centrality is a property of a vertex, not of a network. Answers c and d are incorrect, because the center of a network, if it has one, may either be a single vertex or a cohesive subgroup (e.g., a clique) consisting of a number of vertices that are equally central.
2. Answer d is correct. The star-network is most central, so answer a is not correct. In the circle network (network A), all vertices have degree 2, each vertex is equally distant from all other vertices, and each vertex is situated on one geodesic between pairs of other vertices, so there is no variation in centrality, hence minimum centralization. Network A is least centralized, so answer c is not correct. Network D is more centralized than network B, so answer d is correct.
3. The distances between vertex v1 and vertices v2, v3, v4, and v5 are 3, 2, 3, and 1 respectively. The sum distance is 9, so the closeness centrality of vertex v1 is 4 (the number of other vertices) divided by 9, which is 0.44.
The sum distance of vertex v3 to v1, v2, v4, and v5 is 2 + 1 + 1 + 1 = 5, so its closeness centrality is 4/5 = 0.8.
4. The geodesics between vertex v1 and v2 (v1-v5-v3-v2), v1 and v4 (v1-v5-v3-v4), v5 and v2 (v5-v3-v2), v5 and v4 (v5-v3-v4), and v2 and v4 (v2-v3-v4) include vertex v3, whereas the geodesic between v1 and v5

(v1-v5) does not. Five of six geodesics include v3, so its betweenness centrality is 0.83.

5. The easiest way to omit all unilateral nominations is to change bi-directional arcs into edges (*Net>Transform>Arcs->Edges>Bidirected Only*) and remove the remaining arcs subsequently (*Net>Transform>Remove>all arcs*). Now you have an undirected network and it is easy to compute the three kinds of centrality (degree, closeness, and betweenness) and display the vertices with highest scores with the *Info>Partition* and *Info>Vector* commands. Vertex v9 has highest degree (17), highest closeness centrality (0.67), and highest betweenness centrality (0.20). Person v9 consistently ranks highest on the three centrality indices.