CHAPTER 4

SOCIAL LINKS

At the same time, as social beings, we are moved in the relations with our fellow beings by such feelings as sympathy, pride, hate, need for power, pity, and so on.
—*Albert Einstein*

Learning Objectives

1. Understand and describe the concept of social circles and how they are formed.
2. Understand and describe the concept of link formation and the influence of Homophily, Reciprocity, Proximity, Prestige, Transitivity, and Balance in link formation.
3. Understand the following concepts and describe their applicability in social networks:
   (a) Social Conformity.
   (b) Structural Holes.
   (c) Social Capital.
   (d) Link Optimization.
   (e) Clusterability.

Why Do We Connect?

There exist a certain set of social dynamics that dictate how people form relationships. Social network analysis provides a context in which to understand how these forces work upon individuals. The framework of a hierarchy of social link motivation presented here is straightforward, designed by the authors, and serves to organize the vast research on social theory that drives link formation in networks.

Social Network Analysis: with Applications, First Edition.
Ian A. McCulloh, Helen L. Armstrong, and Anthony N. Johnson.

4.1 INDIVIDUAL ACTORS

Fundamentally, people need to feel valuable and this feeling is only validated in the presence of others. Cooley (1902) refers to this concept as the "looking glass self." There are three components to this theory: (i) people consider how others perceive them, (ii) they imagine how others judge that appearance, and (iii) they develop self-identity through the perceived judgment of others (Yeung and Martin, 2003). The very activities that people pursue, the accomplishments that they take pride in are those that other people recognize as valuable. Psychologist Erik Erikson's theories on personality and identity have self-worth at the core of all stages of human development (Erikson, 1956; Marcia, 1966). Carl Rogers (1961) theories on personality and the self-concept are examples of social group validation of individual identity and self-worth. In essence people need to belong to a social group.

People are constrained, however, in the number of relationships they can maintain. British anthropologist Robin Dunbar (2010) suggested that there existed a cognitive limitation on the number of stable relationships a person can maintain. This did not include people who the individual knew in the past but had ceased a regular relationship with, nor did it include latent friends, where there was no regular contact. She used a model connecting the neocortical ratio to primary group size to predict that humans should have about 150 individuals in their primary social group. The neocortical ratio is the volume of the neocortex divided by volume of the remaining brain. Dunbar did not provide a precise definition of relationship. Killworth and Bernard (1979) conducted several empirical studies, which have shown this limit to be as high as 290 connections with a median of 231 links, which is based on empirical findings of observed human relations. Mayhew and Levinger (1976a, 1976b) show findings that suggest that the density of connections between actors is a function of the context of the relationship and the size of the population. Intuitively, we can also recognize that an individual has a finite amount of time in the day to invest in relationships. Thus, the number of links an actor maintains is a function of three main variables: the nature of the relationship (online, acquaintance, face-to-face friendship); the required amount of time invested to define that a relationship exists; and the cognitive limitations of the actor.

The need for self-worth to be validated by others and the actor's constraints in terms of time and cognitive limitations drive individuals to maximize the utility of their relationships (Lospinoso et al., 2009). They may choose to establish many low value ties, few high value ties, diverse ties across different social groups, among other strategies. This has several important implications for link formation in networks. First, individuals may derive different utility from different networks. Some networks may provide access to knowledge or resources, which make them valuable in another setting. Other networks may provide social entertainment. Different networks may provide love and security. Yet other networks may provide communities of interest for different hobbies, activities, and employment. Charles Kadushin (2012) identified these different networks as *social circles*.

Most individuals have membership in multiple social circles, and it is rare that two individuals will have all of their social circles in common. Feld (1981) explains this as Focus Theory. He describes social context as a set of foci and individuals. Individuals are related to the extent that they share a focus. Feld considers a focus to be a "social object," which could be group membership, a work place, ethnic identity, attitude, and so on. He proposes that individuals' shared relations to foci create social circles and increase their likelihood for interaction. This is an explanation for overlapping patterns of network clusters. Maintaining membership in multiple social circles is also efficient from an optimization perspective. By maintaining relationships with a few individuals from disparate groups, an individual can have access to a wider variety of resources without the cost of establishing relationships with all members of the group.

4.2 SOCIAL EXCHANGE THEORY

Another key dynamic for network link formation is social exchange theory. People in positions of advantage within the network often exploit their position for advantage. Exchange theory was first presented by Homans (1958) and it suggests that all human relationships or social interactions are exchanges based upon an analysis of costs and benefits as well as analysis of alternatives. Social exchange theory evaluates the social associations analyzed in terms of the value of giving and receiving, and when the costs outweigh the benefits then the relationship breaks down.

The exchange process is complex as relationships can be interdependent and are context driven. Zafirovski (2005) observes that reciprocity compels exchange in social interaction, and relations do not continue where reciprocity is breached, as relations continue based upon the expectation that the interaction will be mutually beneficial. Blau (1994) believes the initial drive for social interaction is initiated by the perceived exchange of benefits, both intrinsic and extrinsic, regardless of cultural norms. Molm (1990) discovered that theories that neglect the effect of actors within given structural conditions are incomplete, and how much power an actor has and how it is used to respond to another's behavior are distinct determinants of exchange outcomes. Later research (Molm et al., 2000) on negotiated and reciprocal direct exchange showed that reciprocal exchanges produced stronger trust and commitment than negotiated exchange, and that behaviors indicating or confirming the partner's trustworthiness had a greater influence upon trust in the reciprocal exchange.

Emerson (1969) posits that reciprocity in social exchange links closely with the concepts of power, dependence, and cohesion. Power is derived from a position to control resources and hence wealth and is the property of a relation rather than an actor, because it is inextricably linked with the other's dependency (Emerson, 1962). It is not difficult to see how power can influence the outcomes of exchanges and stability of the associated networks. As networks can be combinations of complex relationships much research has been conducted at basic levels of exchange on very small networks in an effort to understand what happens at a more cellular

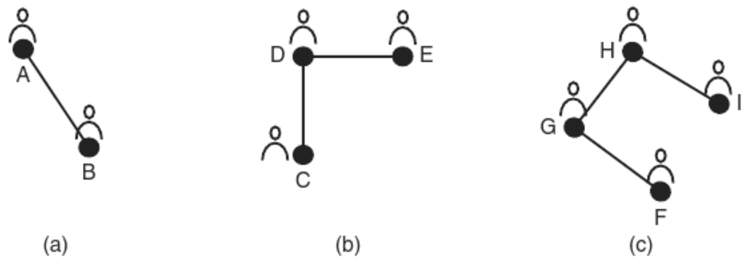

FIGURE 4.1 Power in basic network exchanges: (a) two-node network, (b) three-node network, and (c) four-node network.

level. A number of nodes are required to negotiate the split of given resources, commonly money. Figure 4.1 illustrates simple networks of two to four nodes in an exchange activity.

The results are reasonably simple in a network of two nodes, say nodes A and B. The nodes negotiate and divide the resource, and the result is most commonly an equal split. In a three-node network consisting of nodes C, D, and E in Figure 4.1, node D has greater power as it is positioned between C and E and is able to negotiate with both these nodes. However, nodes C and E are only able to negotiate with node D. In the four-node network consisting of nodes F, G, H, and I, nodes G and H hold the greatest power as G can negotiate with F and H, and node H can negotiate with G and I. If any of the four nodes have a resource to share then nodes G and H are in the most favorable positions and can exert the most power on individual exchanges. They also stand to obtain the best outcomes. When you add other influences to relationships between nodes in a network, such as an organizational reporting responsibility, a homophilous belief or a reciprocal expectation, it becomes more difficult to isolate and quantify the effect of each influencing factor.

Check on Learning

1. Briefly explain the concept of “social circles.”
2. Briefly explain the concept of “social exchange theory.”

Answer

1. Individuals seek confirmation of self-worth and different social networks provide the needed validation. An individual’s needs will provide the impetus for the nature and strength of ties and their diversity of associations.
2. Social exchange theory evaluates links between people based upon the value of giving and receiving, and the exchange will continue as long as the benefits are greater than the costs.

4.3 SOCIAL FORCES

There are six key social forces that drive the formation of social links: homophily, reciprocity, proximity, prestige, transitivity, and balance. All of these drivers of link formation represent costs and benefits associated with establishing and maintaining links in a network.

4.3.1 Homophily

Homophily deals with the tendency of individuals to form relations with those like themselves (Blau, 1977; McPherson et al., 2001). If two actors share common interests, beliefs, goals, race, gender, and/or culture, they are more likely to form connections than if they shared no common interests or features. This is often commonly termed *birds of a feather flock together*. Homophily increases the likelihood that alters will have values that reinforce the individual's self-worth. As people have multiple attribute characteristics, homophily must be evaluated across multiple dimensions, known as *Blau-space* (McPherson, 1983), named after the scientist who proposed the concept in 1977. Homophily is one of the most robust findings in social science.

Interesting studies have explored whether perceived homophily exists in networks, that is, peoples' perceptions of others' interests and beliefs drive their behavior. With perceived homophily, an individual's perceptions are updated through interaction, which may lead to dynamic changes in network behavior over time. Interaction can also create opportunities to modify ideology and value systems, which can contribute to social conformity and group think as actors become more similar in an unconscious effort to seek greater validation of self-worth.

4.3.2 Reciprocity

Reciprocity is whether agents tend to form directed relationships with alters who initiate relationships with them. In other words, if agent A sends an email to agent B, does agent B tend to send emails back to agent A? From a utility perspective, it is easier to maintain a social relationship when the other actors take an active role in reaching out and reciprocating the relationship. Reciprocity may partially be an emergent property of homophily if two similar agents develop a reciprocal relationship due to their shared attributes.

Reciprocity has become a key aspect of social networking using social media and communications such as Facebook and Twitter blogs. Status is seen to be gained from the number of reciprocal contacts and conversations carried out and the number and prestige of connections are important for the perceived prestige of the individual. Communication between two peers would be expected to follow the following interaction format: user A sends user B a message and receives a reply, and the continuing exchange will be bidirectional and symmetric. However, if user A were to send multiple messages to user B who is a celebrity, it is highly unlikely that B will reply and the interaction is unreciprocated. Cheng and colleagues (2011) formulated an approach to predicting the symmetry of the interaction using node attributes that approximate the relative status of the two communication nodes. In closely knit groups based upon homophily the lack of reciprocity can weaken the ties between group members, making the group more fragile as members do not abide by the expected norms of the group.

4.3.3 Proximity

Proximity is the organizational or physical distance between nodes. This can be organizational or physical. Organizational proximity refers to a semantic distance between groups in an organization. People that work together in shipping and receiving are closer in organizational proximity than they are to a person in manufacturing. They may have an even greater distance to a person in accounting, even if that section is physically closer to the shipping and receiving office. Within an organization, there are often common meetings, goals, and joint tasks, which provide opportunities for interaction. An individual in a different organization will often have conflicting demands. This phenomenon makes it easier for members of the same organization to interact than for members of different organizations. Thus, organizational distance creates a cost to link maintenance.

Physical proximity is the actual physical distance between individuals. Sailer and McCulloh (2011) showed that this was also based on perceptions. Individuals in an open office plan perceive distance based on how visible others are within the office, whereas with closed individual offices a metric distance equivalent to the number of paces an individual walks to interact with another is more appropriate. In either case, the closer actors are to one another, the more likely they are to interact and form relationships. Greater physical distance incurs a cost of link maintenance in terms of time to reach another node or in terms of the technological effort, such as a phone or email.

As groups form, social norms are established. A norm is a value of appropriate or inappropriate placed upon feeling, thought, or behavior. Friedkin (2001) suggests that shared belief in these values with influential others reinforces and validates the belief, thus creating the group norm. He further suggests that these norms provide an informal set of expectations, which provide opportunities for individuals to validate their sense of self-worth within the context of the group or social circle. An individual's choice to conform with the group is driven by their need for social acceptance and validation and a perception of how well the group's norms align with their positive attributes. Group membership can form an integral part of a person's individual identity.

4.3.4 Prestige

The establishment of social norms provides a mechanism for two important concepts in social psychology, *prestige* and *social conformity*. Individuals who epitomize the social norms and values of the group not only derive a personal sense of self-worth, but also are perceived by others to be valuable. These individuals have high prestige and hold greater influence over the attitudes and ideology of group members. They are also able to influence group norms that may emerge in the future. Other members of the group will often choose to link to high prestige nodes for greater validation of their own self-worth.

Perceived motivation can influence individual status within the group as well (Ridgeway, 1978; 1982). Low status individuals who are perceived to be group-oriented combined with reasonable task proficiency tend to attain higher status levels over time. By contrast, individuals who are low status and are perceived to be self-oriented tend to remain low status. High status individuals tend to remain high status regardless of their motivation as long as their task proficiency remains high.

Prestige can also be achieved in terms of an individual's access to resources, knowledge, and other social circles. Within this context prestige can be deliberately improved. A node can strategically develop ties between social circles that will benefit one another, thereby increasing their value. They can position themselves within close proximity to others in the network, increasing their opportunities for social interaction. Individuals can demonstrate an adoption of group values making themselves more homophilous to others in the group. Finally, simply being outgoing may leverage the concept of reciprocity by reducing the link maintenance cost of others.

The concept of prestige enables high prestige nodes to maintain more links with less cost. As group members have a desire to connect to a high prestige node, they are willing to overcome other social costs such as a lack of proximity, reciprocity, or homophily, although nodes that exhibit low social cost are even more attractive. This does not mean that a prestigious node is unconstrained, however. They still must make some effort to maintain the relationship from their end. These nodes also must maintain social norms in order to retain their prestige within the group. This can sometimes be challenging if prestigious individuals choose to affiliate with social circles that may have competing social norms.

4.3.5 Social Conformity

Individuals who do not adopt group norms will have reduced opportunity to validate their sense of self-worth. Conforming to the group norms provides them an opportunity to validate their sense of self-worth. This mechanism drives *social conformity*.

There is a social network group effect that contributes to conformity. That is, research shows repeatedly that groups can exert implicit pressure to influence opinions leading to higher conformity, more extreme views, and a group think mentality where contradictory evidence is ignored. For example, Solomon Asch (1955, 1956) conducted a study, where subjects were placed in a group, where the other group members were given information that was different from reality. For example, eight individuals (the subject and seven others) were shown three lines of various lengths and asked to verbally identify which line matched a given fourth line. The subject was positioned to answer after the seven other individuals. Most of the time, the seven other individuals would give a correct response, but in certain cases, the other subjects would agree on a wrong answer. Asch then measured how the subject would modify their answers based on the majority opinion. The objective of the experiment was to understand factors that would enable people to resist group/peer pressure.

The study found that 37% of the subjects conformed to the wrong responses given by the majority. It was not clear if the subjects actually believed their incorrect response (*informational conformity*), or if they were simply conforming to the group so they were not singled out (*normative conformity*). A variation of this experiment had one of the seven other individuals give correct responses, and the conformity dropped from 37% to 5.5%. These findings suggest that an individual's perception of reality, to include their attitudes, may be significantly influenced by other people to whom they are socially connected.

McCulloh (2013) conducted a social network variation of Asch's experiment using a platoon in the US Army. Social network data was collected on the platoon, recording directional ties of self-reported friendship and respect, creating two sets of network data on the same subjects. Four nodes that were highly central in both networks and four nodes that were peripheral in both networks were selected to be the subject of an Asch type experiment. Eight other nodes in the network were selected to serve as confederates of the experiment. Instead of asking participants to offer opinions on line length, the individuals were told that they were preparing for promotion boards and were being tested on military knowledge. Some of the questions were modified to offer an obviously wrong answer. The most ridiculous question was, "when do you apply a tourniquet to a neck wound? (i) when it is spurting blood; (ii) when it is coming from an artery; or (iii) never."

This study demonstrated a profound network effect on conformity. Central nodes conformed to the group and provided wrong answers 7% of the time on average. Peripheral nodes, however, conformed 62% of the time. The strongest correlation between centrality and the number of wrong answers occurred with betweenness centrality in the friendship network with a correlation of -0.84.

This study provides evidence of the need for social acceptance. Individuals who are peripheral and have need of social acceptance may conform as a way to gain group acceptance. Central actors who already have a sense of acceptance are free to identify the "wrong" answers without fear of group sanction.

Mcculloh et al. (2010) studied the social network effects on post-traumatic stress disorder (PTSD). They recorded data on the mental health of 1000 soldiers in a US Army brigade at three time periods; 2 months prior to deployment, 3 months into their deployment, and 2 months following their 1 year deployment to Afghanistan. They found that peripheral individuals in the friendship network were more likely to exhibit symptoms of PTSD and depression. In a similar series of studies on isolated groups at the South Pole, Jeff Johnson et al. (2003) discovered that peripheral individuals would often develop thyroid problems, which is related to depression. Thus, individuals may have a physiological need for social acceptance and conformity is a defense mechanism to achieve that acceptance.

In the famous Milgram Experiment (Milgram 1963, 1974) conducted at Yale University, Milgram studied the willingness of people to obey perceived authority figures, when their instructions conflicted with moral conscience. He found that 26 of the 40 subjects (65%) followed the immoral instructions. This suggests that people can be influenced to a greater degree, by a minority of people, if the minority is seen as authority figures. These authority figures derive their authority informally from epitomizing group norms, as well as formally. Thus nodes with high centrality may serve as an opinion leader for the group.

At the point where people have joined various social circles, derived some benefit from their associations, and potentially conformed to group norms, they are in a position to optimize the utility they derive from their social circles. There are several mechanisms for this behavior: transitivity, social capital, link optimization, and balance. These may not be conscious mechanisms, but they are powerful social forces that describe many dynamics of informal networks.

4.3.6 Transitivity

Transitivity of relations means that if there is a link from actor A to actor B, and a link from actor A to actor C then there is a tendency for actor B and actor C to form a link with each other. In other words, connections are formed through a common connection to a third party. The underlying social mechanism for transitivity may overlap competing social forces. If homophily has driven the links between actors A and B and between actors A and C, then there would be similar attributes that would drive the connection between actors B and C in a homophilous relation. On the other hand, if the connection involves time spent in social interaction, actors B and C would have greater opportunities to meet through their shared proximity to actor A. Therefore there are three potential confounded explanations behind transitivity. Transitivity may be an emergent network property based on (i) homophily, (ii) proximity, or (iii) brokered social relations from a common friend or acquaintance.

Modeling transitivity in statistical models and simulations can present major problems, which argue in favor of the emergent property perspective. If actors attempt to establish transitive links in a network, then the tendency would be for networks to have one giant fully connected component and many isolate nodes with no connection to the network. This usually does not represent observed network behavior. These models are degenerate in that the empirical data is extremely improbable under the models estimated from the same data.

An improvement to the concept of transitivity is geometrically weighted edgewise shared partners (GWESP) (Hunter and Handcock, 2006). This approach models a diminishing marginal return in the utility for additional shared transitive links. In other words, if two actors share a connection to a common alter, they are more likely to establish a link. If they have three alters in common they are even more likely. Four alters are more likely still. However, if they have 20–30 alters in common and have not established a link, there is probably a good reason for it. Thus, each additional shared alter is less important in terms of utility. The rate at which this decays is a geometric parameter in the GWESP model. This approach has much better modeling properties. It is rarely degenerate and thus supports the theory of utility as a mechanism of establishing social links.

Sometimes transitivity can be deliberately employed by an actor to increase their utility within the context of one or more social circles. One of the benefits of social relations is the access to knowledge and resources. Nodes that occupy a network position in between a key resource and others in the network derive informal power as discussed in Chapter 2. However, they also incur a social burden and cost to maintain relationships with the two groups (those with the resource and those requiring the resource). Brokering an introduction can provide an actor the opportunity to derive a sense of self-worth and value by facilitating a group's access to the resource, without the requirement to maintain the links, especially if the volume of resource flow is large (Feld, 1981).

4.3.7 Balance

Links do not always represent positive affinity and in some cases negative affinity may affect positive social relations. *Balance* (Heider, 1946; Cartwright and Harary, 1956) is an aggregate level measure that tests whether "the enemy of my enemy is my friend." If two actors that are "friends" have the same affinity to another there is balance. If the two friends have differing affinity to another there is a cognitive dissonance. *Cognitive dissonance* (Festinger, 1957) is a discomfort resulting from holding conflicting views at the same time. For example, a person who smokes must reconcile their desire for a long healthy life with the fact that smoking is likely to cause lung cancer and other health issues. People have a strong desire to reduce dissonance by altering their views and beliefs or changing their behavior. The smoker might therefore deny the evidence of health concerns associated with smoking or they may attempt to quit smoking.

In a network application, when an actor has positive affinity for two alters who dislike each other, there is dissonance. He must reconcile his positive affinity with his "friends" to their negative affinity toward the third alter. This may cause the actor to change his views toward one of the other alters to reduce the dissonance. Consider the eight possible triads in Figure 4.2 given only positive and negative relationships. You can think of these as "love-hate" triangles. The top four triangles with an even number of negative ties are balanced. The bottom four triangles with an odd number of negative ties have cognitive dissonance. In these four cases there will exist a social force placing pressure on one of the unbalanced triads to move toward balance.

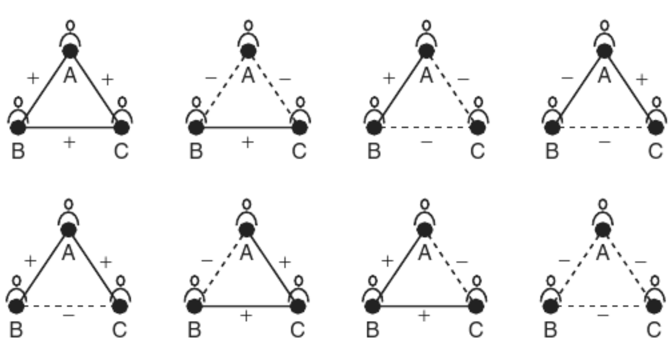

FIGURE 4.2 Eight possible signed triads.

Check on Learning

Note the appropriate social force next to the explanation for the following statements (choose from social forces: Homophily, Reciprocity, Proximity, Prestige, Social Conformity, Transitivity, Balance).

1. Geoff and Ted are friends, but Ted's wife does not like Geoff, so Ted no longer sees Geoff socially.
2. Two individuals belong to the same tennis club.
3. A group of five people who work in the same physical office area trust each other.
4. Joe purchased an iPhone so Bill also purchased an iPhone.
5. Three agents of the same religion spend private time on family outings together.

6. Jake took up golf because the rest of his peers play golf.
7. Mary was invited to dinner by Casey when she took up a new position in the department, Mary returned the invitation and now they share recipes.
8. Lance and Fred are both engineers but are situated at different locations. Despite their separate physical locations they frequently use Skype and emails to share knowledge and ask advice.
9. Charles is respected because he has a degree in law and is very knowledgeable about legal matters.
10. Gillian met Susan at John's party, then again at a local fund raising event managed by John. They got talking and realized they had common interests.

Answers

1. Balance
2. Homophily
3. Proximity
4. Social conformity, Prestige
5. Homophily
6. Social conformity
7. Reciprocity
8. Proximity, Homophily
9. Prestige
10. Transitivity

4.4 GRAPH STRUCTURE

Several theories have natural extensions to graph level measures and properties. Balance, in particular, lays the foundation for Structural Balance and Clusterability in networks that are briefly presented here.

4.4.1 Structural Balance

Structural Balance is a graph level measure that will lie between 0 and 1, just like centralization. A structural balance of 1 means a graph is perfectly balanced, and a 0 means none of the triads are balanced. Figure 4.3a and b is an example of unbalanced and balanced graphs. Each link in the two graphs is labeled as either positive or negative. Balanced triads will reflect one of the four structures in the top row of the signed triads in Figure 4.2 and unbalanced triads will demonstrate a structure in the lower triad row.

4.4.2 Clusterability

Clusterability (Davis, 1967) is an aggregate level measure that describes how well the network can be partitioned into teams. Imagine a parliament of legislative body with two political parties. Suppose that one party's members all like each other, but dislike all members of the opposing party. The other party is similar with all its members liking each other and disliking all members of the opposing party. In this case, we would see a signed adjacency matrix that looks like this:

$$\mathbf{A} = \begin{pmatrix} 0 & + & + & 0 & - & - \\ + & 0 & + & - & 0 & - \\ + & + & 0 & - & - & 0 \\ 0 & - & - & 0 & + & + \\ - & 0 & - & + & 0 & + \\ - & - & 0 & + & + & 0 \end{pmatrix}$$

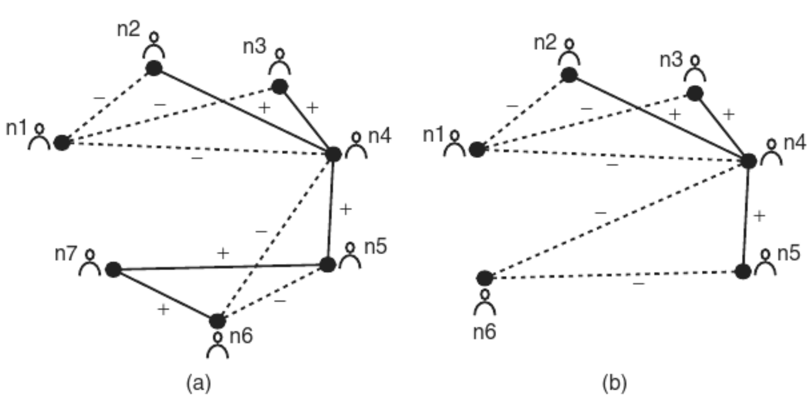

FIGURE 4.3 Two signed graphs: (a) unbalanced network and (b) balanced network.

Check on Learning

Draw the network depicted by the above signed adjacency matrix, A, and determine if any clusters are present.

Answer

Clusterability can be determined by reordering the nodes in the adjacency matrix to place positive and negative relations next to each other as shown in the adjacency matrix above. The first three nodes in the adjacency matrix have positive relations in the top left partition of the adjacency matrix, which shows links to other nodes in this group. There are negative ties to other nodes in the network. The lower right partition also has all positive links. This example network is therefore perfectly clustered with three nodes in each group. If there was any dissent within the parties, the graph is no longer perfectly clusterable. Unlike structural balance, clusterability is NOT a sliding scale measure. A graph is either clusterable or it is not.

4.5 AGENT OPTIMIZATION STRATEGIES IN NETWORKS

Some actors in a network will employ various strategies to optimize the value they derive from a social network. Value can be defined in several ways. It may consist of the access to resources that they gain. It may be the power they derive from brokering resources. It may be the social acceptance value they derive from being central. There are several key strategies, which include structural holes, social capital, and link optimization.

4.5.1 Structural Holes

An actor may attempt to maintain a position as the sole broker between two groups. Ron Burt (1995, 2009) describes the emergent network pattern as *structural holes*. Although the actor's position does provide him an informal power in the network, it can be very difficult to maintain. Most individuals view this type of position negatively and these actors can be sanctioned by the group in other ways. Others in the group will inevitably seek to develop redundant ties separate from the broker to reduce their dependence on him and increase their own utility in the network. Actors attempting to maintain structural holes, therefore, must spend time preventing links from emerging and this behavior reduces collaboration, knowledge, and resource sharing across the organization. Conversely, transitivity increases collaboration, knowledge, and resource sharing and can still provide the actor with informal power derived from his ability to connect to others.

4.5.2 Social Capital

The ability of an actor to connect to others in the network and provide access to knowledge and resources through their social connections is known as *social* *capital* (Bourdieu 1979, 1980, 2008; Coleman, 1988). There is value in collaboration and knowledge. Some managers do not understand this concept and actually unwittingly destroy social capital within their organizations. These managers see their employees chatting in the break-room, going to the designated smoking area to indulge their habit, or otherwise socializing instead of completing some assigned task. The managers view this type of interaction as the employee wasting time or having a poor work ethic. In reality, these social interactions provide unique opportunities for people in disparate parts of an organization to interact and exchange knowledge and information. These activities enable employees to establish new links throughout the organization. These employees now act as high betweenness nodes brokering knowledge and resources. Successful managers will often attempt to create unique opportunities to facilitate social interaction and collaboration and view chatting in the break-room as a valuable investment in organizational capacity and social capital.

Social capital does not come free. There is a cost in terms of time, social expectations, and resources to develop and maintain friendships. Not all actors within an organization will provide the same value, so social capital investments must be made wisely. Link optimization can facilitate the efficient acquisition of social capital.

4.5.3 Link Optimization

Link Optimization is a concept proposed by Cornell sociologist Matthew Brashears (2011). He describes two aspects of social relations: bandwidth and latency, which is referred to in this text as frequency and intimacy, respectively. Frequent positive interactions with an individual will increase the perceived utility of the connection. A friend or coworker that you see several times a week will usually feel like a stronger, more valuable connection than someone you have not seen in years, all else being equal. Intimacy also increases the perceived utility of the connection. As people share secrets, experiences or exchange favors, their level of intimacy increases, but this intimacy does not necessarily entail high frequency of interaction.

Consider two potential actors, a household nanny and your best friend since childhood. The nanny has frequent interaction with you, whereas you might not have seen your childhood friend in years and only speak on the phone a couple times per month. Although the nanny has some level of intimacy in that she works in your home and deals with your children, she has probably not shared the depth of intimate experiences with you as your childhood friend. If you had an important decision to make or needed a favor, who would you feel more comfortable calling upon? On the other hand, if you wanted to know about what happened on last night's episode of your favorite television program, who would you be more likely to ask?

The two aspects of social relations can be plotted as shown in Figure 4.4. The x-axis shows the frequency of interaction and the y-axis shows the intimacy. Different categories of social connections can be expressed in different quadrants

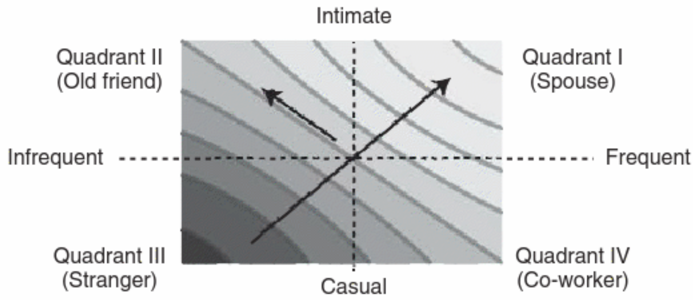

FIGURE 4.4 Pareto optimality space.

of the chart. Quadrant I may include close family members or a spouse. Quadrant II would consist of old, yet close friends or family. Quadrant III consists of strangers. Quadrant IV is made up of coworkers or people with membership in civic organizations. Pareto optimality curves can be expressed over the space represented in Figure 4.4. Relationship strength follows the gradient moving from quadrant III to quadrant I. As the strength of the link increases, the derivative benefit or utility also increases. Unfortunately, the cost associated with maintaining the link may also increase. Connections in quadrant I will have additional expectations associated with maintaining a friendship. In addition to the time cost, they may be called upon for favors and be required to tolerate inconvenience for the sake of maintaining the relationship.

Orthogonal to the strength of relationship is the latency of relationship. Alters classified in quadrant II do not require much investment of time and resources to maintain the association; however, they may still be called upon when needed. This may drive individuals to invest in intimacy to facilitate latent associations as well as develop strong connections. Recently, social media has significantly enabled latency by lowering the cost of maintaining links. This enables people to keep alters in a place where they can be called upon for their utility value when needed with little to no maintenance cost.

In some cases, online media can significantly reduce the costs to maintain a link. Without the cognitive and temporal limitations on the number of links a node can have, some nodes can amass very large numbers of connections. This is common for Web site connections, where once a link is established, there is no cost required to maintain the link. Popular Web sites are able to exploit their prestige and position in the network to acquire more and more links. This enables scale-free networks where there exist hub nodes with many links, although most nodes have relatively few links. The term scale-free is derived from the power-law distribution of node degree, which does not have a scale parameter in the distribution. This type of network structure is not possible in face-to-face networks where cognitive limitations on degree exist. People are not unaware of this phenomenon. A card in the mail is often more meaningful than a "happy birthday" on Facebook.

4.6 HIERARCHY OF SOCIAL LINK MOTIVATION

To present theories of social links in a unified frame, the authors posit a hierarchy of motivations in the formation of social links. The hierarchy comprises four stages: Affiliation, Forming, Conforming, and Optimization (see Fig. 4.5). In stage 1 (Affiliation), nodes attempt to validate a sense of self-worth through social acceptance or access resources by identifying and affiliating with social circles. This stage is characterized with largely weak social ties to loose affiliations, where there exists a perceived benefit. The status of alters may affect interactions when people first meet in this stage. Thus, prestige is an important social force during this stage. In stage 2 (Forming), nodes attempt to establish links with alters that present the greatest probability of success in terms of validation or resources. Social forces such as homophily, reciprocity, and proximity are paramount. As an actor establishes themselves in the social group, they eventually move to stage 3 (Conforming). In this stage, actors exhibit characteristics of social conformity as they attempt to gain greater group acceptance and establish relationships with high prestige alters. Social forces such as prestige become primary. With an adequate level of group acceptance, a node moves to stage 4 (Optimization). In stage 4, nodes attempt to maintain or increase their utility derived from relationships while minimizing their effort and social costs to maintain links that have been established. Concepts such as transitivity, structural holes, social capital, link optimization, balance, and clusterability provide more significant motivators for link formation.

During situations when small groups are forming and all actors enter a social circle at the same time, there are no established opinion leaders, group norms, or high prestige nodes. This affects the dynamics of link formation. In this situation, stage 1 has already occurred by the assignment of individuals to the group. For example, a cohort of graduate students admitted to a masters degree program have all applied for and been accepted into a program of study. The individuals passed through stage 1 Affiliation in their application process. They will have decided that

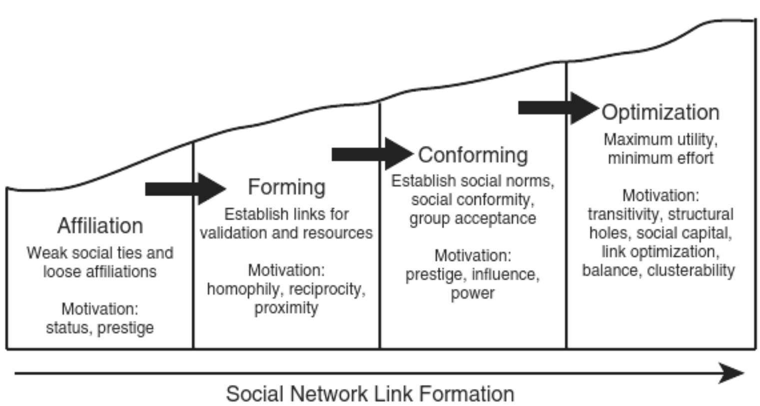

FIGURE 4.5 Four stages of link formation.

TABLE 4.1 Stages of Social Link Motivation and Stages of Small-Group Development

Social link motivation	*Small-group development*
Stage 1—Affiliation	
Stage 2—Forming	Stage 1—Forming: testing and dependence
Stage 3—Conforming	Stage 2—Storming: intra-group conflict
Stage 4—Optimization	Stage 3—Norming: development of group cohesion
	Stage 4—Performing: functional role relatedness

the graduate program is a worthy pursuit that provides them some sense of self-worth or increased status by hopeful attainment of the degree. Thus the small-group formation begins with nodes in stage 2 Forming. Tuckman (1965) and Tuckman and Jensen (1977) describes this as the first stage in small-group formation, which he labels as "testing and dependence." Individuals determine who they will study with, who they will ask for peer assistance on assignments, and who they will seek social relationships with. With no established authority figures, individuals moving into stage 3 conforming have difficulty in determining what social norms to conform to and whom the group sees as prestigious. This manifests itself with Tuckman's "intra-group conflict." Individuals attempt to assert roles and group values that further their own objectives. As individuals progress into stage 4 optimization, they establish their role in the newly formed group as well as the relationships that maximize their perceived value. Through this process, the group will establish social norms and individuals who epitomize those norms will increase in prestige. Tuckman refers to this as "development of group cohesion." This is where the group establishes a group identity and established routines. At this point, Tuckman introduces an additional stage for small-group development, which he labels "functional role relatedness." In this stage, the group reaches its highest level of productivity. Tuckman combines group formation with task activity and summarizes the combined model as "forming, storming, norming, and performing." Table 4.1 shows stages of social link motivation with corresponding stages of small-group development. In this newly formed group, as new individuals join, group norms and functions are already established. New actors will still test relationships and establish dependence through homphily, reciprocity, and proximity. However, as individuals move toward the subsequent stage, they will have greater tendency to conform to established group norms and there is less intra-group conflict. As these actors optimize their relationships, they assume a new role in the group identity and aid the group in fulfilling its functional role.

4.7 SUMMARY

The following are the main points covered in this chapter

- Social forces of homophily, reciprocity, proximity, prestige, transitivity, and balance drive the formation of social links. Homophily posits that people form relations with those who are similar to themselves, that is, similar beliefs or common interests. Reciprocity occurs where links are reciprocated by another. Proximity by physical location or organizational activity will encourage links where the distance between individuals is minimal. Prestige is the value an individual gains by taking on group norms (social conformity) and values resulting in greater influence within the group. Transitivity purports that if there is a link between actors A and B, and a link between actors B and C, then there is a tendency for a link to develop between actors A and C. Balance suggests that the enemy of my enemy is my friend. Dissonance will occur where one individual has positive links with two alters who dislike each other.

- Structural holes develop where an actor positions him or herself between two groups and actively maintains control as sole broker. Members of the group may view this negatively and develop redundant links to reduce their dependence on the broker and increase their value and effectiveness in the network.

- Social capital is the value an actor obtains from ties with others, providing access to knowledge and resources not previously available. Increasing social interactions and collaboration increases social capital.
- Frequency and intimacy of connections will strengthen relationships, thus optimizing links. Social link motivation comprises four stages:
    - Affiliation: where agents strive to make links that confirm their self-worth.
    - Forming: agents establish links that provide success in terms of validation and/or resources.
    - Conforming: agents adopt social conformity for greater group acceptance.
    - Optimization: agents maximize utility gained from links while minimizing the effort and social costs to maintain those links.

EXERCISES

4.1. Use the three concepts of the "looking glass self" to describe your sense of self-identity.

4.2. How many social circles do you affiliate with? List them and provide a short description of the group. What value do you gain from those relationships?

4.3. Construct a Blau-space. Think of two social circles and two attributes of individuals that might be in those social circles. Draw x and y axes corresponding to the two attributes and plot individuals from the groups based on your numerical evaluation of their attributes. Draw circles that encompass each of the two groups. Do they overlap? What conclusions can you draw about relationship formation?

4.4. How does proximity drive the formation of social relations? How can a common employee lounge area in an office increase collaboration? If you were a manager of a knowledge intensive organization, what kind of policies and practices would you implement to increase social interaction?

4.5. What strategies can an individual employ to increase their prestige within a social group?

4.6. Balance theory applied to a two-party political system would suggest that all members of one political party have affinity for each other and animosity for the opposing party. Why is this not observed in real networks? Can you think of any examples of real-world clusterable networks?

4.7. Consider the last time you entered a new social setting. Perhaps it was your arrival at college, starting a new job, or attending church for the first time. What characteristics did you look for in those individuals whom you initiated relationships with? Why did you develop the relationships you did?

REFERENCES

Asch, S. E. (1955). Opinions and social pressure. *Nature*, **176**:1009–1011.

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs: General and Applied*, **70**(9):1–70.

Blau, P. M. (1977). *Inequality and Heterogeneity: A Primitive Theory of Social Structure*. Free Press.

Blau, P. M. (1994). *Structural Contexts of Opportunities Structural Contexts of Opportunities*. University of Chicago Press.

Bourdieu, P. (1979). The three states of cultural capital. *Proceedings of the Social Science Research*, **30**(30):3–6.

Bourdieu, P. (1980). The social capital. *Proceedings of the Social Science Research*, **31**(31):2–3.

Bourdieu, P. (2008). The forms of capital. In Biggart, N.W., editors. *Readings in Economic Sociology*. Blackwell Publishers Ltd.

Brashears, M. (2011). Homophily and link atrophy. In Quantification, Metrics and Modeling, Durham, NC. Duke University Workshop on Weak Social Ties.

Burt, R. (1995). *Structural Holes: The Social Structure of Competition*. Harvard University Press.

Cartwright, D. and Harary, F. (1956). Structural balance: a generalization of heider's theory. *Psychological Review*, **63**(5):277–293.

Cheng, J., Romero, D., Meeder, B., and Kleinberg, J. (2011). Predicting reciprocity in social networks. In Privacy, Security, Risk and Trust (PASSAT), 2011 *IEEE Third International Conference on Social Computing (SocialCom)*, pp. 49–56, October.

Coleman, J. S. (1988). Social capital in the creation of human capital. *The American Journal of Sociology*, **94**:95–120.

Cooley, C. H. (1902). Human nature and the social order. In *Scibner's Conference*, New York, pp. 183–184.

Davis, J. (1967). Clustering and structural balance in graphs. *Human Relations*, **20**(2):181–187.

Dunbar, R. (2010). *How Many Friends Does One Person Need?: Dunbar's Number and Other Evolutionary Quirks*. Faber and Faber.

Emerson, R. (1962). Power-dependence relations. *American Sociological Review*, **27**(1):31–41.

Emerson, R. (1969). Operant psychology and exchange theory. In Robert, B. and Bushell, D., editors, *Bevahiorial Sociology*, pp. 379–408. Columbia University Press, New York.

Erikson, E. (1956). The problem of ego identity. *Journal of the American Psychoanalytic Association*, **4**(1):56–121.

Feld, S. (1981). The focused organization of social ties. *American Journal of Sociology*, **86**(5):1015–1035.

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford University Press.

Friedkin, N. (2001). Norm formation in social influence networks. *Social Networks*, **23**(3):167–189.

Heider, F. (1946). Attitudes and cognitive organization. *The Journal of Psychology: Interdisciplinary and Applied*, **21**(1):107–112.

Homans, G. (1958). Social behavior as exchange. *The American Journal of Sociology*, **63**(6):597–606.

Hunter, D. R. and Handcock, M. S. (2006). Inference in curved exponential family models for networks. *Journal of Computational and Graphical Statistics*, **15**(3).

Johnson, J., Boster, J., and Palinkas, L. (2003). Social roles and the evolution of networks in extreme and isolated environments. *Journal of Mathematical Sociology*, **27**:89–121.

Kadushin, C. (2012). *Understanding Social Networks: Theories, Concepts, and Findings*. Oxford University Press, illustrated edition.

Killworth, P. and Bernard, H. (1979). Informant accuracy in social network data iii. *Social Networks*, **2**:19–46.

Lospinoso, J., McCulloh, I., and Carley, K. (2009). Utility seeking in complex social systems: an applied longitudinal network study on command and control. In *Social Networks and Multi-Agent Systems Symposium (SNAMAS-09)*, Edinburgh, Scotland. Heriot-Watt University, SSAISB.

Marcia, J. E. (1966). Development and validation of ego identity status. *Journal of Personality and Social Psychology*, **3**:551–558.

Mayhew, B. H. and Levinger, R. L. (1976a). On the emergence of oligarchy in human interaction. *American Journal of Sociology*, **81**(5):1017–1049.

Mayhew, B. H. and Levinger, R. L. (1976b). Size and the density of interaction in human aggregates. *American Journal of Sociology*, **82**(1):86–110.

Mcculloh, I., Geraci, J., Turner, J., and Matthews, M. (2010). Network centrality and PTSD. In *Proceedings of SunBelt XXX, Social Networks and Health. International Network for Social Network Analysis*.

McCulloh, I. A. (2013). Social conformity in networks. *Connections*.

McPherson, M. (1983). An ecology of affiliation. *American Sociological Review*, **48**(4):519–532.

McPherson, M., Lovin, L., and Cook, J. (2001). Birds of a feather: homophily in social networks. *Annual Review of Sociology*, **27**(1):415–444.

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal Psychology*, **67**:371–378.

Milgram, S. (1974). *Obedience to Authority: An Experimental View*. Harpercollins.

Molm, L. (1990). Structure, action, and outcomes: the dynamics of power in social exchange. *American Sociological Review*, **55**(3):427–447.

Molm, L., Takahashi, N., and Peterson, G. (2000). Risk and trust in social exchange: an experimental test of a classical proposition. *American Journal of Sociology*, **105**(5):1396–1427.

Ridgeway, C. L. (1978). Conformity, group-oriented motivation, and status attainment in small groups. *Social Psychology*, **41**(3):175–188.

Ridgeway, C. L. (1982). Status in groups: The importance of motivation. *American Sociological Review*, **47**(1):76–88.

Rogers, C. R. (1961). *On Becoming A Person*. Houghton Mifflin.

Sailer, K. and McCulloh, I. (2011). Social networks and spatial configuration–how office layouts drive social interaction. *Social Networks*.

Tuckman, B. W. (1965). Developmental sequence in small groups. *Psychological Bulletin*, **63**(6):384–399.

Tuckman, B. W. and Jensen, M. (1977). Stages of small-group development revisited. *Group & Organization Management*, **2**(4):419–427.

Yeung, K.-T. and Martin, J. L. (2003). The looking glass self: an empirical test and elaboration. *Social Forces*, **81**(3):843–879.

Zafirovski, M. (2005). Social exchange theory under scrutiny: a positive critique of its economic-behaviorist formulations. *Electronic Journal of Sociology*, **2**:1–40.

CHAPTER 5

SUBGROUP ANALYSIS

We must remember that one determined person can make a significant difference, and that a small group of determined people can change the course of history.

—*Sonya Johnson*

Learning Objectives

1. Know the difference between Newman, Consecutive Correlation (CONCOR), and user-defined grouping algorithms.
2. Develop a blockmodel of a social network based on one of the grouping algorithms.
3. Demonstrate the ability to create and interpret a hierarchical clustering diagram, density matrix, reduced network, and attribute analysis table.

5.1 SUBGROUPS

Subgroup analysis is an important approach for understanding the structure of a network as well as assessing organizational risk in networks. Subgroup analysis is carried out to more fully understand the clustering of nodes into subgroups, usually linked by a common factor or positioned close to each other within the network (Wasserman and Faust, 1994; Krackhardt, 1994; Kadushin, 1995). A substantial amount of prior research has been carried out on the groupings of actors who interact frequently and on how individuals are influenced by these interactions. Network groups can have sparse or dense interactions within the group. Early organizational studies reported that large organizations comprise separate subgroups showing dense interactions (Freeman, 1992). Subgroups within organizations were found to link in order for the organization to evolve or for reasons of management efficiency (Simon, 1962; Granovetter, 1973). Organizations of today are focused on teams and group connectivity and can affect the efficiency and effectiveness of work practices, competitive advantage, and ultimately organizational survival. Interactions between groups impact not only strategic alignment and performance

Social Network Analysis: with Applications, First Edition.
Ian A. McCulloh, Helen L. Armstrong, and Anthony N. Johnson.

in the attainment of organizational goals but also the creation and sharing of knowledge that in turn feeds innovation. Analysis of the interactions between group members and external nodes is also of interest when measuring the isolation of the group. In some circumstances isolation may be desirable, but undesirable in others.

In this chapter, we must first paint the context for which subgroup analysis is useful. An overview of organizational theory is presented as a foundation to aid our understanding of groups within organizational contexts. The structural importance of subgroups is then discussed including derivation of the probability of groups forming at random. With this context in place, three popular subgroup heuristics are presented and compared: Newman–Girvan (2004), CONCOR (Breiger et al., 1975), and user-defined grouping approaches. Common methods for analyzing the groups are demonstrated on the Newman–Girvan grouping heuristic.

5.2 ORGANIZATIONAL THEORY

Organizational theory is a collection of sociological and psychological theories of groups for the purpose of improving productivity, maximizing efficiency, and facilitation of group problem solving. There are three common perspectives for understanding organizational theory: classical, neo-classical, and environmental. The *classical perspective* is concerned with increasing efficiency and productivity by designing organizational structures, rules, and expertise. An early pioneer in the study of organizations, Frederick Taylor (1911), introduced *scientific management* for "knowing exactly what you want men to do and then see in that they do it in the best and cheapest way." He identified four principles: (i) scientific measurement of performance, (ii) training of workers, (iii) cooperation between management and workers to ensure productivity, and (iv) division of labor where workers and management specialize in specific tasks and roles.

Another classical theory is *Bureaucracy theory*, introduced by Max Weber (Gerth and Wright Mills, 1948), which focuses on the part of an organization that implements rules, laws, and functions of the group. Individuals assume well-defined roles in this organization with established policies for conduct and performance. Neutral parties, for example, make decisions for career advancement based on individual performance, thus removing personal bias from the process.

There are several important limitations for viewing organizations from the classical perspective. Most important is Taylor's initial assumption . . . "knowing exactly what you want men to do . . . " Scientific management and bureaucracy are better suited to production and manufacturing processes or standard administrative functions, where the task is well defined and repetitive. For knowledge-intensive organizations focused on creative problem solving or knowledge development, there is no concept of "what you want men to do." It is by definition of purpose left undefined. Furthermore, the classic approach omits human needs and concerns. Human error, the variability of performance, motivations, and individual goals are rarely considered with this approach. These limitations provide the basis for a neo-classical approach.

The most fundamental principle in the *neo-classical perspective* is the *Hawthorne effect* (Mayo, 1933, 1945; Whitehead, 1938; Roethlisberger and Dickson, 1939; Roethlisberger, 1941), which emerged out of a series of human productivity experiments at Western Electrics Hawthorne Works facility outside of Chicago, IL. The Hawthorne effect is a phenomenon where individuals alter their behavior as a result of their awareness that they are being observed and measured. This introduced a social-psychological dimension to workplace productivity. Mayo's work in particular asserted that humans were unique and motivated differently and are not interchangeable parts. His work has been criticized as counter to Taylor's scientific management approach. However controversial, Mayo paved the way for organizational theorists to study motivation, employee satisfaction, peer social relations, and leadership in organizations.

Within the context of social link theory (Chapter 4), motivation, employee satisfaction, and peer relations can perhaps be better understood. An individual worker will initially affiliate with an organization in order to access financial or other resources of benefit. The nature of this organization becomes a social circle with an inherent group culture. Individuals are driven to form social relations to gain acceptance. Individuals conform to the organizational culture to achieve validation and prestige within the culturally defined norms. These standards might be performance driven, qualification based, or derived from interpersonal relations. Financial incentives might provide a sense of value in certain organizations, such as financial investment firms or large corporations. We argue that it is not so much the money as it is the value it conveys that is important. In the US military, status is derived from combat tours and specialized military qualifications such as ranger, or paratrooper. Financial incentives are not as effective as status in motivating this workforce. Other extensions of the Hawthorne studies have revealed non-financial incentives for productivity. Thus, an understanding of social link theory provides an alternate frame of reference to understand group and individual motivation. This is at the heart of informal power brokerage. One may ask why a low ranking individual, who is highly central within an organization's social network, may have high informal power, which is more influential in motivating group behavior than the formal power a designated leader is assigned. We contend that this is precisely the issue that Mayo was concerned with understanding, and social network analysis provides an insightful frame of reference to conceptualize this phenomenon.

The most recent perspective in organizational theory is the *environmental perspective*. This approach argues that there is no universally best practice for organizing, leading, or decision making for a particular type of organization. Organizational design must therefore not only consider the function of the organization, but it must also contend with internal social dynamics of the individual members, current and desired organizational culture, informal power brokers and opinion leaders, changing external environment that may affect the organizations purpose, and more. This approach requires a manager to understand the social terrain of their employees and recognize that the organization may not actually be interacting as programmed.

Regardless of perspective, organizations can be classified into one of six structure types (Daft and Armstrong, 2009): prebureaucracy, bureaucracy, postbureaucracy, functional, divisional, and matrix. The *prebureaucracy* structure lacks any standardization of role or task and a central leader makes all decisions. This type of structure is usually confined to small organizations engaged in completing simple tasks.

The *bureaucracy* is defined as an organization that has standardized roles and is focused on efficiency. Weber proposes three criteria for a bureaucracy: (i) clearly defined roles and responsibilities, (ii) hierarchical structure, and (iii) value for merit (as defined by the organizational culture). This structure is well suited for larger organizations engaged in well-defined activities. Again, the primary advantage of the bureaucracy is efficiency; however, it is usually achieved at the cost of reduced flexibility and innovation.

The *postbureaucracy* (Heckscher and Donnellon, 1994), also referred to as *organic*, is the opposite of a bureaucracy. There are no standardized roles, responsibilities, or formal leadership. There is no hierarchy. Decisions are made through discourse and consensus. This type of organization is common among cooperatives, community organizations, and nonprofit organizations. It encourages participation and empowers individuals who may not otherwise have a voice in the decision-making process. The primary advantage of this organization is its flexibility, innovation, and commitment of membership; however, these goals are achieved at the cost of efficiency.

The *functional structure* is defined as an organization where groups are formed based on specialized expertise. For example, there may exist a sales team, an engineering division, and an accounting office. Or in a manufacturing process, casting, milling, shaping, joining, and assembly might all fall in separate divisions, where production flow moves back and forth between divisions and competes with other product lines for processing. Functional structures increase skill capacity through divisional mentorship, training, and oversight. However, they lose some efficiency in cross-divisional collaboration that might inhibit the development of new products and cross-disciplinary problem solving. This structure is best suited for organizations that conduct a limited set of tasks or produce a limited amount of products at high volume and low cost.

The *divisional structure* is constructed such that it has each functional element within it and can operate almost independently from other divisions. Sometimes divisions are organized geographically such as the United States, Australia, or EU divisions. Alternately, divisions may be organized around product lines such as consumer goods, commercial goods, and industrial goods. Each division may even have its own marketing and sales departments. The advantage of this structure is its flexibility to adjust to changing environments and operate independently; however, it comes at the cost of reduced functional expertise.

The *matrix structure* combines the functional and divisional structures. The organization frequently uses teams to accomplish tasks. Matrix structures may have *weak/functional* management, where a project manager oversees the product, whereas functional managers retain most decision making and control over the functional employees. Other organizations may have a *strong/project* management, where project managers control the project and functional leaders provide technical expertise and assign resources as needed. *Balanced* management shares responsibility equally between functional managers and project managers. This organizational structure strikes a hybrid between other structures with the associated strengths and weaknesses.

Depending on the organization's purpose and function, subgroup analysis provides a valuable frame of reference for understanding organizations. Social links may be defined between individuals as affinity, advice seeking, communication, collaboration, perceived usefulness, frequency of interaction, or other forms of relationship. Identifying subgroups by defined organizational structure allows the network analyst to evaluate whether the organizational structure is performing as designed or if a new structure might improve performance. Additionally, methods such as Newman grouping can identify self-organizing subgroups that may provide insight into organizational design.

The purpose of a manager or leader in any organization is to define goals and motivate organization members to maximize their productivity and efficiency. Social network analysis provides powerful tools and theory to understand important social dynamics that affect every aspect of an organization. The remainder of this chapter will present several techniques and theories for understanding subgroups within social networks. Identifying subgroups also allows an investigator to study networks of organizations in addition to individual actors.

5.3 RANDOM GROUPS

Consider for a moment that the probability of a node being connected to another node occurs completely at random with the flip of a coin (probability is 0.5). Then, the probability of getting a unique sequence of 1s and 0s is $(\frac{1}{2})^{(n-1)}$. Therefore the probability of two nodes being exactly structurally equivalent (have exactly the same connections) is also $(\frac{1}{2})^{(n-1)}$. In an eight-node network for example, the probability of two nodes being structurally equivalent is $(\frac{1}{2})^{(8-1)} = 1/128 = 0.0078$. If we looked at a network of only 30 nodes this probability would be 0.0000000186. Therefore, even in small networks the probability of two nodes even knowing similar people is quite small. However, in practice it is not unusual for people to have similar acquaintances. This is because there is some reason for the social interactions, such as shared membership in an organization or some of the social forces presented in Chapter 4. The reason for structural grouping in networks is precisely what makes subgroup analysis so important.

5.4 HEURISTICS FOR SUBGROUP IDENTIFICATION

Networks can be partitioned into subgroups based on the similarity between nodes. There are several ways a node can be similar, however, which has led to multiple approaches for assigning group membership. Three approaches are presented in this chapter: Attribute Defined, CONCOR (Breiger et al., 1975), and Newman–Girvan (2004).

5.4.1 Attribute Defined

The simplest approach to assign nodes to subgroups is by a predefined attribute. This is often done in organizational studies where the attribute is the work section within an organization. For example, a company might have sections such as shipping and receiving, production, sales, customer service, research and development, and management. There is usually a desire to understand how collaborative these sections are with one another. One might hope that there is a lot of interaction between customer service and research and development or between production and sales for example. Assigning individuals by a defined attribute allows exploration of intergroup collaboration.

Assigning nodes to groups based on a predetermined attribute is essentially a straightforward method to defining the groups without the use of any algorithm or heuristic approach. The attributes may represent node characteristics to study potential homophilous social circles, also known as *Blau-space*. Alternatively, the attribute might be the pre-designated grouping such as formal structures within an organization. When groups are pre-designated, subgroup analysis methods such as block models can be used to explore organizational collaboration and function.

5.4.2 Consecutive Correlation (CONCOR)

Consider an adjacency matrix for a social network. The CONCOR heuristic correlates the rows or columns of the adjacency matrix, creating a new matrix that contains the correlation between node connections. CONCOR then takes iterative correlations of this new matrix. This process will converge such that nodes are correlated with a $+1$ or a -1, thus partitioning the network into two groups. This procedure can be executed multiple times, each time bifurcating the network in two. Therefore, executing the algorithm three times should partition the network into eight groups. It is possible, however, that no nodes would be assigned to one of the groups in the case of structural equivalence between nodes.

CONCOR is typically used to identify role redundancy within a network. There is a certain amount of redundancy that is necessary within an organization. Identifying nodes that are similar in their connections can assist the analyst in strengthening and coordinating activities of redundant nodes. In other cases, there may be too much redundancy, such as after a merger, acquisition, or joint venture. Management may wish to reassign individuals to other activities within the organization or make other labor decisions. CONCOR can assist by identifying similar individuals.

5.4.3 Newman–Girvan Grouping

Another approach is the Newman–Girvan grouping, commonly referred to as simply Newman grouping. The Newman grouping iteratively identifies edges that are high in betweenness. This is similar to betweenness centrality; however, the algorithm looks for the edges that fall on geodesics most frequently. The highest edge in betweenness is removed and the edge betweenness is recalculated. The next edge that is highest in betweenness is removed and so forth. This process is repeated until a specified number of subgroups are found or until a specified ratio of in-group links to out-group links is achieved. Thus, the number of subgroups can be specified or the algorithm can determine the number of subgroups.

The Newman group is excellent for identifying network clusters. The clusters are naturally formed groups that essentially have no probability of occurring at random, as shown earlier. An analyst can explore the network to attempt to identify the reason for the social grouping. This is a common approach when there are no attribute defined groups.

Check on Learning

Which subgroup analysis method is used to find clusters of related nodes and/or cliques?

Answer

Newman–Girvan grouping. The attribute defined groupings do not necessarily consist of related nodes. They could be anything that the analyst defines. The CONCOR grouping identifies nodes with similar connections, which usually identifies role redundancy, but not necessarily clusters. The Newman–Girvan grouping iteratively deletes links with high betweenness, revealing clusters of related nodes.

5.5 ANALYSIS METHODS

We examine three analysis methods for investigating subgroups in a network. They are the group membership method, the hierarchical clustering method, and the block model method (White et al., 1976). These methods will be demonstrated using example network graph and adjacency matrix (Fig. 5.1 and Table 5.1).

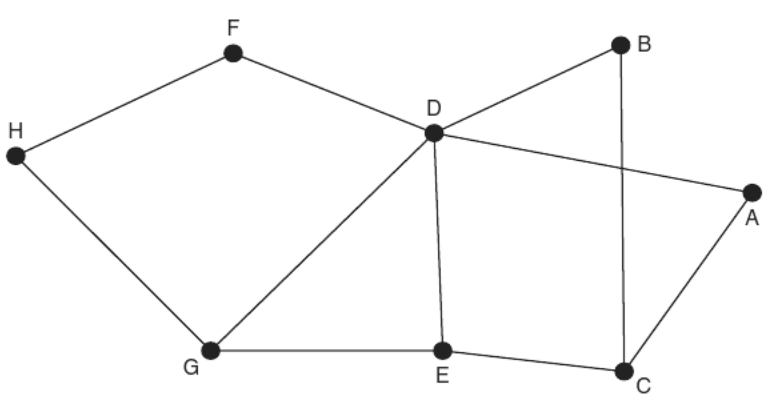

FIGURE 5.1 Example network.

TABLE 5.1 Adjacency Matrix for Example Graph

	A	B	C	D	E	F	G	H
A	0	0	1	1	0	0	0	0
B	0	0	1	1	0	0	0	0
C	1	1	0	0	1	0	0	0
D	1	1	0	0	1	1	1	0
E	0	0	1	1	0	0	1	0
F	0	0	0	1	0	0	0	1
G	0	0	0	1	1	0	0	1
H	0	0	0	0	0	1	1	0

5.5.1 Group Membership

The *group membership* analysis method investigates the assignment of nodes to different groups. It can be seen in Table 5.2 and Table 5.3 that the Newman grouping and CONCOR grouping will group the nodes in the network differently. In the Newman grouping, the groups are close together in the network. Looking once more at the Example Network graph (Fig. 5.1), Group 1 comprising nodes A, B, and C appears at the right of the network. Group 2 consisting of nodes D, E, and F appears in the center of the network. Nodes G and H, forming Group 3, appear at the left of the network. In contrast, the groups in the CONCOR grouping are based on the correlation or similarity of connections. Nodes A, B, and H are all similar in their connection to node D and their lack of connections to nodes E and F. As node C is connected to node E, it is excluded from this group. Nodes C, D, and G are similar in their connections to Group 1; however, none of them are connected to each other. This group may fulfill similar or redundant roles in the

TABLE 5.2 Newman Grouping

Group	Size	Members
1	3	A,B,C
2	3	D,E,F
3	4	G,H

TABLE 5.3 CONCOR Grouping

Group	Size	Members
1	3	A,B,H
2	3	C,D,G
3	4	E,F

organization. Nodes E and F are similar in their connection to Group 2 and their lack of connection to Group 1.

Check on Learning

Which grouping algorithm do we need in order to find out the following information from our social network of employees and clients? There are only two node types in our network: agent (employee) and agent (client). Employee node information includes their name, store location, department, gender, and pay level. Client node information includes their name, address, age, and gender.

1. Does gender matter when linking employees and clients? We want to know which female employees connect with female clients, female employees with male clients, and finally male employees with male clients.
2. We want to remove duplications in the relationships between employees and clients. Which employees are connected to the same clients?
3. We want to know who works in the same department. All employees have links with the other employees in their department but few links with employees in other departments.
4. We want to minimize our pay costs. Which employees present the highest cost in pay and which are the lowest pay cost?
5. Who are the clients who connect with employees in more than one department?

Answer

1. Newman grouping.
2. CONCOR grouping.
3. Attribute grouping on department, and Newman grouping.
4. Attribute grouping on pay level.
5. CONCOR grouping.

5.5.2 Hierarchical Clustering

The *hierarchical clustering diagram* depicted in Figure 5.2 provides a visual chart of the similarity between nodes in a network. The dots along the top row indicate that at some level all of the nodes are isolated. At the next level, nodes G and H have the closest relationship. At the next level, nodes A and C have the closest relationship. At the next level, nodes E and F have the closest relationship. Then, at the next level, node D is added to the E–F group. Then B is added to the A–C group. At this point, the network is clustered into three groups, which is what the Newman grouping returned as the optimal number of subgroups. However, you

```
Hierarchical Clustering Diagram

          B A C D E F G H


   Level   2 1 3 4 5 6 7 8
  -----   - - - - - - - -
      0    . . . . . . . .
 -0.066    . . . . . . XXX
 -0.000    . XXX . . . XXX
 +0.053    . XXX . XXX XXX
 +0.111    . XXX XXXXX XXX
 +0.161    XXXXX XXXXX XXX
 +0.161    XXXXX XXXXXXXXX
 +0.000    XXXXXXXXXXXXXXX
```

FIGURE 5.2 Hierarchical clustering chart produced in ORA for the example network of Figure 5.1.

could continue grouping and connect the D–E–F group with the G–H group and finally connect the whole network as one complete group. This diagram provides some insight into the group formation and the sensitivity of the network to a grouping threshold.

5.5.3 Block Model

Once the groups have been determined, a *block model diagram* depicted in Figure 5.3 provides some very important insight into group dynamics. The nodes are arranged in the rows and columns of the adjacency matrix such that they are listed next to other nodes within their group. Lines are drawn between the groups. Compare the block model (Fig. 5.3) to the original adjacency matrix, Table 5.1. In this case the original adjacency matrix is listed in the same order as the subgroupings, but this is not always the case.

Block densities can be determined by calculating the number of links in each block divided by the number of possible links. For this example, in the upper left block between nodes A,B,C, and themselves (1,2,3) there are $n(n-1)$ or six possible links. Recall that the diagonal of the adjacency matrix is undefined, disregarding the possibility of reflexive links. There are actually four links present


```
Block Model
           1 2 3   4 5 6   7 8                 1 2 3   4 5 6   7 8
         ------------------------            ------------------------
    1 A |     1 | 1     |     |         1 A |       |       |     |
    2 B |     1 | 1     |     |         2 B | 0.67  | 0.33  | 0.0 |
    3 C | 1 1   |   1   |     |         3 C |       |       |     |
         ------------------------            ------------------------
    4 D | 1 1   |   1 1 |   1 |         4 D |       |       |     |
    5 E |     1 | 1   1 |     |         5 E | 0.33  |  1.0  | 0.33|
    6 F |       | 1 1   | 1   |         6 F |       |       |     |
         ------------------------            ------------------------
    7 G |       |     1 |   1 |         7 G |  0.0  | 0.33  | 1.0 |
    8 H |       | 1     | 1   |         8 H |       |       |     |
         ------------------------            ------------------------
```

FIGURE 5.3 Block model chart produced in ORA for the example network of Figure 5.1.

```
Block Reduced Network

     1 2 3
   1 1 0 0
   2 0 1 0
   3 0 0 1

   Network density: 0.333333
```

FIGURE 5.4 Block reduced chart produced in ORA for the example network of Figure 5.1.

out of the six possible, so the block density is 0.67. The block connecting A, B, C to D, E, and F (4,5,6) has n^2 or nine possible links, as the diagonal is included in this block and all potential entries are possible. The block density is therefore three out of nine possible or 0.33. The block densities provide insight into the coordination and connections between subgroups in the network. Figure 5.4 depicts the construction of a *reduced network* where the nodes of this new network are the subgroups and the links are connections between subgroups. This is extremely powerful for exploring collaboration within an organization. Subgroups with high informal power and brokerage or diffusion reach can also be determined.

There are three ways to determine the connections between subgroups. The *fully connected* approach would argue that for a subgroup to be connected to itself or another group, all of the nodes of one group must be connected to all of the nodes of the other group. Under this criterion, subgroup 2 is connected to itself because D is connected to E and F and E is connected to F. Subgroup 3 is connected to itself because G and H are connected to each other. Subgroup 1 is not connected to itself because node A is not connected to node B.

The *minimally connected* approach would argue that if one member of one group is connected to one member of another group, there is a connection between the groups. Under this criterion, the only groups that are not connected are Groups 1 and 3, because there are no connections between node G to A, B, or C and there are no connections between node H to A, B, or C.

The most widely accepted method for determining subgroup links in a reduced network is to compare the block density with the original density of the network (input network density). In this example, the original network density is 0.393. The only block densities that exceed the input network density are those along the diagonal. Thus the input network density approach would identify cohesive subgroups.

The *external/internal link analysis* can provide some further insight into how well groups collaborate with each other within the network. This analysis compares the number of internal links within the subgroup to the number of links that nodes in the subgroup have with nodes outside the subgroup. The silo index provides a score that ranges from −1 to 1 to show how isolated a particular subgroup is. If 50% of the links from a subgroup are internal or external, then the silo index is 0. If a subgroup has more internal links than external links the silo index will be positive (Table 5.4).

TABLE 5.4 External/Internal Link Analysis

Group	Internal link count	External link count	Internal links%	Silo index
1	4	6	40	−0.2000
2	6	10	37.50	−0.2500
3	2	4	33.33	−0.3333

Example 5.1

For the network depicted in Figure 5.5 determine the Newman grouping.

The first step of performing a Newman grouping is to label the edges so that they can be treated as nodes for betweenness calculations. Figure 5.6 illustrates the process.

Next, we calculate edge betweenness for the entire network using the methods given in Chapter 2. In this case, edge "f" has the highest betweenness and is therefore removed. See Figure 5.7. At this point it is intuitively clear to see that edge "a" and edge "b" have the same betweenness. We randomly choose "b,", but

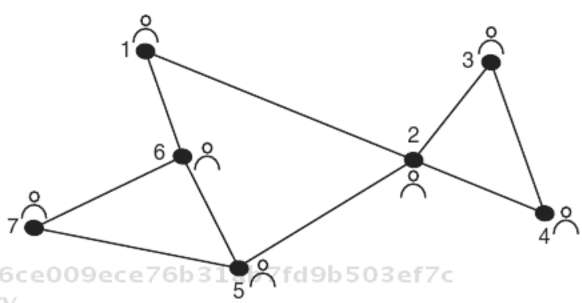

FIGURE 5.5 Example network for Newman grouping.

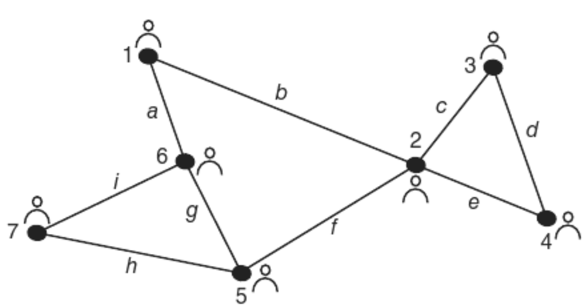

FIGURE 5.6 Network with edge labels.

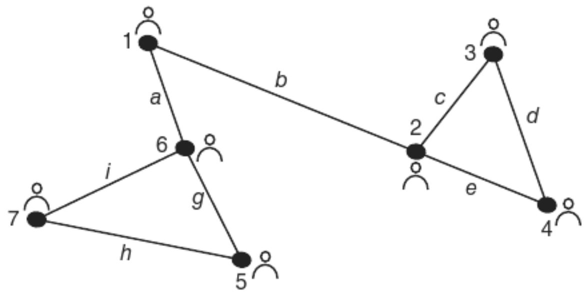

FIGURE 5.7 Highest betweenness edge removed.

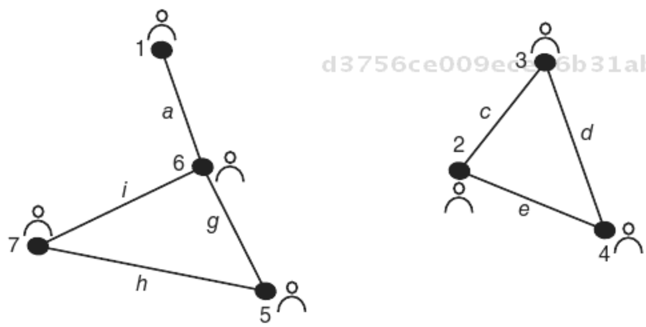

FIGURE 5.8 Newman subgroup emerges.

one might choose "a" as well based on some prior iteration of the algorithm. After removal of edge "b", we see the first fracture of the network into two separate pieces illustrated by Figure 5.8.

Calculating edge betweenness again reveals that edges "i" and "g" have the same betweenness value. Choosing "i" arbitrarily produces the network depicted in Figure 5.9.

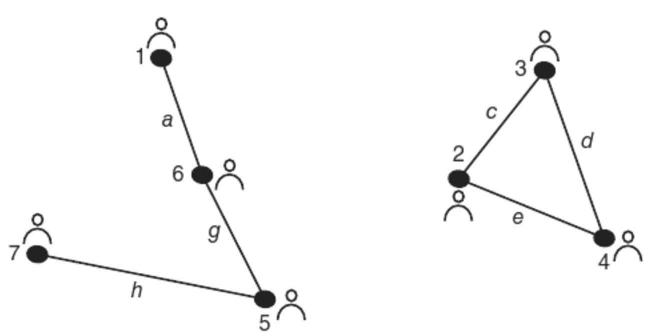

FIGURE 5.9 Iterate of Newman algorithm.

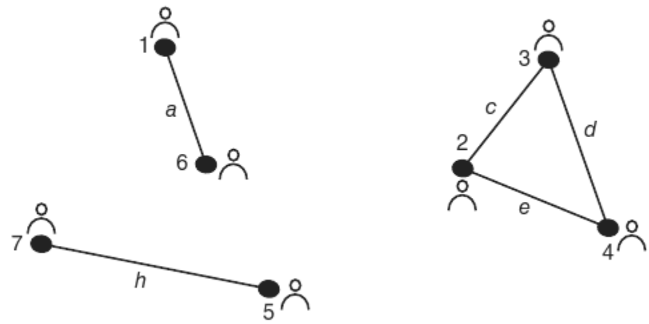

FIGURE 5.10 Iterate of Newman algorithm.

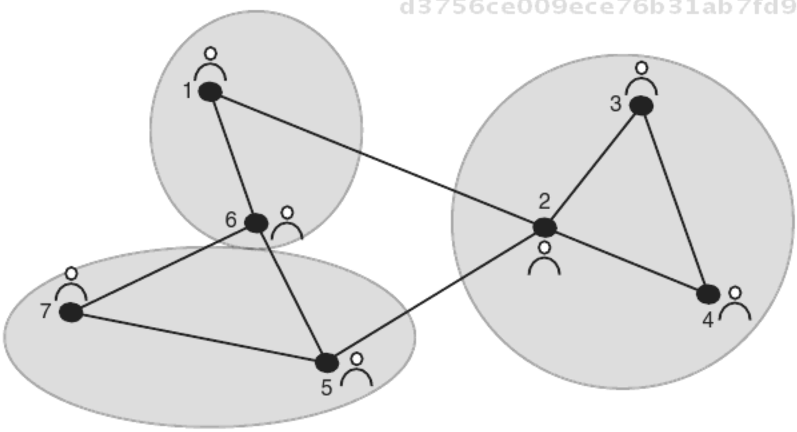

FIGURE 5.11 Newman groups.

The final edge betweenness calculation identifies “g” as the edge to be removed giving us the final version of the network with three clusters identified. Figure 5.10 depicts the network after the final iteration of the Newman grouping algorithm.

Figure 5.11 is the Newman grouping for the network. Now we can see that the closest ties are nodes 1 and 6, 5, and 7, and 2, 3, and 4.

Example 5.2

Using the Newman grouping of Example 5.1, create a hierarchical clustering diagram for the social network.

To create the hierarchical clustering diagram by hand, label a table’s columns with the closest ties next to each other, and the rows by the number of steps (or levels) it takes to proceed from a completely connected network to a completely

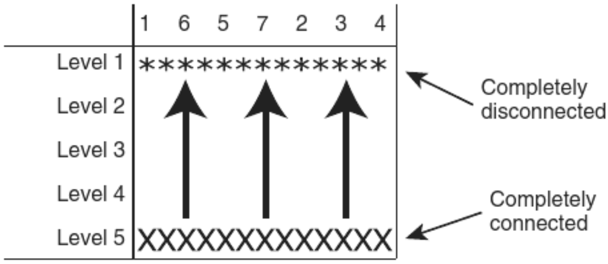

FIGURE 5.12 Empty hierarchical clustering diagram.

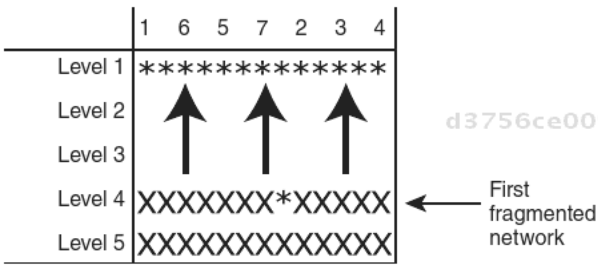

FIGURE 5.13 Hierarchical clustering diagram begins to fill.

	1 6 5 7 2 3 4
Level 1	*************
Level 2	XXX*XXX***XXX
Level 3	XXX*XXX*XXXXX
Level 4	XXXXXXX*XXXXX
Level 5	XXXXXXXXXXXXX

FIGURE 5.14 Full hierarchical clustering diagram.

disconnected network (Fig. 5.12). Next, taking note of successive iterates of the Newman algorithm, determine the structure of the first nonconnected network. In this case, it appears in Figure 5.8. We note the structure, and our diagram slowly begins to fill with very useful information as to the emergence of subgroups based on structure (Fig. 5.13).

The final form of the hierarchical table shows the progress of the Newman algorithm from a connected network to a fully fragmented network with no ties (Fig. 5.14).

5.6 SUMMARY

Where the network is an organization, the type and structure of the organization will provide the foundation for groupings within the network and the density of interactions between groups. Factors such as the hierarchy (bureaucratic, post-bureaucratic, or organic) and structure (functional, divisional, matrix) will be reflected in the groupings evident in the organizational network graph.

Networks can be grouped using different algorithms; the choice of algorithm will depend upon the nature and structure of the network and the objectives of the research. Clusters can be identified based on the position of nodes in the network, the structure of their relationships with surrounding nodes or by specific attributes that nodes may display. Subgroup analysis provides the analyst with information regarding the clustering of nodes, the existence of cliques, whether the clusters are internally or externally focused, how the clusters link together, and the effect all of these may have on the network and its operations.

The following are the main points relating to subgroup analysis covered in this chapter:

- The Newman–Girvan algorithm (Newman groups) groups nodes based on progressively removing edges high in betweenness, reflecting naturally formed clusters.
- The CONCOR algorithm divides the network into two groups multiple times to identify nodes that are similar in their connections.
- Attributes associated with nodes can be used to define subgroups, removing the need to perform analyses using the Newman and CONCOR algorithms.
- Hierarchical clustering and block model diagrams illustrate in a tabular format the grouping of nodes into clusters and illustrate how these clusters are connected.
- An analysis of external versus internal links for subgroups will indicate the extent of isolation (siloing) from the rest of the network.

CHAPTER 5 LAB EXERCISE

Subgroup Analysis

Learning Objectives

1. Analyze subgroups in a network using the Newman clustering algorithm.
2. Analyze subgroups in a network using the CONCOR clustering algorithm.
3. Analyze subgroups in a network using Attributes.
4. Visualize a network and color nodes by Newman, CONCOR, or Attributes.

Enter the data for the network Subgroups1 into ORA. The tables of data for this network can be found in Table B.8 in Appendix B. Save the meta-network as *Subgroups1.xml*.

Visualize the network and save the network diagram as a .jpg. The *Subgroups1* network should look like as follows: