# Chapter 1

## INTRODUCTION TO SOCIAL NETWORK ANALYSIS

Social networks are as old as the human species. As small bands of hunter-gatherers spread around the globe, their survival depended on cooperative strategies for pursuing game and finding good foraging grounds. Ties of family and extended kin were crucial to raising the next generations. With increased size and density of agrarian settlements, succeeded by expanding urban civilizations, networks grew increasingly complex and indispensable for merchants involved in long-distance commerce and armies engaged in conquest. Palace and court intrigues ran on gossip, rumor, and favor-trading among political factions. Scientific and technological advances necessitated information flows through invisible colleges of experts. Social networks have a truly ancient lineage yet are seldom noted nor well understood by their participants.

People today commonly envision social networking as clusters of coworkers going for lunch or coffee, teams of dormmates playing basketball or softball, and bunches of friends chewing the fat. Yes, those small groups are all social networks. To give a formal definition, a *social network* is a set of actors, or other entities, and a set or sets of relations defined on them. In the three preceding examples, the first actors are coworkers and the relations are lunchmate and coffeemate; the second actors are residents of the same dorm and playing sports is the relation; the third network is friends gossiping leisurely. Applying the definition to diverse social settings, we can easily uncover numerous social networks, some more formal than the three previously described. For example, a college academic unit has a social network composed of faculty members, staff, students, and administrators. Multiple sets of relations suffuse such networks: collegial relations among faculty members, faculty advising graduate students, faculty instructing undergraduates, and administrators supervising faculty and staff. A police department is also structured as a formal social network, in which officers at the same rank are colleagues, whereas a quasimilitary chain of command establishes hierarchical authority relations. Typical order from top down would consist of chief of police, deputy chief, captain, lieutenant, sergeant, corporal, patrol officer.

Although people typically conceive the actors in social networks as human beings, they can just as well be collective entities or aggregated units, such as teams, groups, organizations, neighborhoods, political parties, and even nation-states. For example, corporations can engage in cooperative

and competitive relations to pursue many outcomes, such as jointly developing new technologies and products or acquiring greater market shares (Knoke, 2001). Interorganizational relations take many governance forms, from contractual agreements to equity stakes (Child, 2005; Yang, Franziska, & Lu, 2016). Inside organizations, work groups and teams often engage in knowledge transfers or information sharing to facilitate innovation and improve task performance (Tsai, 2001). International relational networks also emerge and evolve, including military alliances and conflicts, trade partnerships and disputes, human migrations, intelligence exchanges, and technology sharing and embargoes (Yang et al., 2016, Chapter 8).

Nonsocial networks are prevalent in many domains: technology networks, computer networks and the Internet, telephone networks and electrical power grids, transportation and logistics networks, food delivery, and patent-citation networks. They share some similarities with social networks, except that instead of actors their units are physical entities, such as computers and transformers, and their relations are transmission and delivery lines such as Ethernet cables, wireless connections, airline routes, and interstate highways. We mention nonsocial networks primarily to note that networks are the subjects of studies by many disciplines besides the social sciences. Those investigations illuminate and inspire one another, engendering strong momentum to improve network knowledge, including social network analysis (Knoke & Yang, 2008). For example, after mathematicians developed graph theory, computer scientists applied it to construct optimal computer networks. Social network scholars can borrow algorithms from computer and mathematical sciences to decipher communication networks among friends, coworkers, and organizations.

Sociology built a long tradition of examining the social contexts of social networks. Founding fathers such as Georg Simmel, Émile Durkheim, and Max Weber promoted a structural perspective in the study of human behaviors. Social psychologist Jacob Moreno (1934) was directly responsible for laying the foundation of modern social network analysis. With Helen Jennings, Moreno invented *sociometry* to draw maps visualizing individuals and their interpersonal relations, revealing complex structural relations with simple diagrams. Moreover, Moreno and other pioneering social network scholars endeavored to explain how network structures affect human behaviors and psychological states (Freeman, 2004). On the one hand, we can better understand people's actions and decisions by examining their social networks because networks provide participants with both opportunities and constraints. On the other hand, the formation and change of social networks themselves have been the object of many research projects. An important sociological principle is social *homophily*, which asserts that people tend to form positive relations with others similar to themselves.

Actors could be attracted to others based on similarity of attributes—such as gender, age, race, ethnicity, or socioeconomic status—or similarity of behaviors—such as life experiences, political preferences, religious beliefs, or hobby interests. In this perspective, social relations are outcomes, or dependent variables, occurring because actors share some of the independent variables listed previously.

Social network analysis was vitally important to the inception of economic sociology, a major specialty in sociology. In his classical article applying sociology to economic actions, Mark Granovetter (1985) criticized the undersocialized view of economists in which human decision making is driven solely by subjective expected utility maximization. Surprisingly, Granovetter likewise disapproved of the oversocialized view of sociologists in which human actions are determined solely by norms and social roles. So how does one avoid both under- and oversocialized explanations of human behaviors? The answer, quite obviously, is by using social network analysis: by looking at actors' social networks, we can better understand their decisions and actions. Social networks generate localized norms, rules, and expectations among their members, which reinforce mutual trust and sanction malfeasance. Thus, by examining how social networks actually operate as both causes and consequences of human perceptions and actions, theorists and researchers avoid accepting either oversocialized or undersocialized perspectives. More importantly, although Granovetter (1985) emphasized economic behaviors, his arguments are very relevant to many social pursuits, such as making friends, casting votes, looking for a job, seeking promotion, finding a therapist, searching for emotional support, and locating instrumental help.

Early sociological and anthropological research on social networks inspired other disciplines to investigate the mechanisms instigating network formation in those fields. Over the past half century, mass communication, strategic management, marketing, logistics, public administration, political science, international relations, psychology, public health, criminology, and even economics begin introducing ideas and methods of social network analysis into those disciplines. For example, Zeev Maoz (2012) analyzed international trade and military alliances as network processes. He found that international trade follows a preferential attachment or bandwagon process: all nations want a quick and short connection to a few key nations in the global trade network, resulting in a highly condensed, single-core structure. In contrast, for military alliances, nations tend to partner with countries sharing similar political ideologies and regime structures. This homophily preference produces a network configuration consisting of multiple small military alliance clusters that are only sparsely interconnected (see also Yang et al., 2016, p. 198).

We would be remiss not to mention social media as an explosively growing component of social networks. Facebook, Twitter, LinkedIn, WeChat, and other apps facilitate a massive amount of daily information exchange among billions of users. Much social networking nowadays occurs in virtual spaces as users contact one another via computers, laptops, iPad tablets, and smartphones linked together by Ethernet cables or wireless. Computer communication networks and human social networks converge, engendering innumerable research opportunities and challenges for social and computer scientists. How does one best search, capture, aggregate, store, share, process, reduce, and visualize vast volumes of complex data generated by online social networkers (Press, 2013; Lohr, 2013)? John Mashey, chief scientist at Silicon Graphics, is often credited with coining the term Big Data, which he described in a slide presentation as "storage growing bigger faster" (1998, p. 2). Exponentially bourgeoning quantities of structured and unstructured information have revolutionized businesses, nonprofits, and governments. For social network researchers, Big Data is a trove of rich relational databases and a smörgåsbord of computer tools for data mining, information fusion, computational intelligence, machine learning, and other applications (de Nooy, Wouter, Mrvar, & Batagelj, 2018). Although Big Data enhances organizational operations and outcomes, it also raises numerous ethical and privacy challenges, such as the rise of surveillance state capacities to predict and control populations (Brayne, 2017; Madden, Gilman, Levy, & Marwick, 2017). Russian manipulation of the 2016 U.S. presidential election was only the most notorious of innumerable criminal abuses of Big Data on social media platforms. Calls for governmental regulation of social media companies encounter conundrums of how to protect platforms and safeguard free speech while prohibiting dangerous content (Berman, 2019). The fate of our democracy hangs in the balance.

In sum, social network analysis is a vibrant multidisciplinary field. Peter Carrington and John Scott called it "a 'paradigm', rather than a theory or a method: that is, a way of conceptualizing and analyzing social life" (2011, p. 5). We believe the network paradigm has roots in and thrives on the integration of three elements: theories, methodologies, and applications. For theories, network analysis demands serious commitment that prioritizes actor interdependence and connectivity, emphasizing structured relations among social entities. For methodologies, network analysis borrows eclectically from diverse disciplines, collaborating across the aisles to create innovative procedures. For applications, people increasingly use their networking skills to navigate along complex interorganizational pathways to acquire desired goods and services, such as better healthcare, shopping bargains, and recreational experiences.

This volume updates the second edition of *Social Network Analysis* by Knoke and Yang (2008). In addition to providing a general overview of fundamental methodological topics, we cover new developments of the past decade. Our approach is didactic, aimed primarily at graduate students and professionals in many social science disciplines, including sociology, political science, business management, anthropology, economics, psychology, public administration, public health, and human resources. College faculty could assign it as a text in graduate-level courses, use it for workshops at professional association meetings or summer instructional institutes, or study it to learn more about networks on their own. Graduate and advanced undergraduate students interested in social network analyses can read it to get a jump-start on their social network skills and intellectual aspirations. Professionals face many challenges in developing social network research, such as how to design a social network project, details and problems that may arise during network data collection, and alternative techniques for analyzing their social network data. Social network scholars may find this volume a useful brief refresher or reference book. For more advanced texts, we suggest Easley and Kleinberg (2010); Dorogovtsev and Mendes (2014); Lazega and Snijders (2015); de Nooy, Mrvar, and Batagelj (2018); and Newman (2010).

We frequently illustrate concepts and methods by referring to substantive social network research problems, citing examples from children's playgroups to organizations, communities, and international systems. We tried to write with a precision and freshness of presentation using concise language that minimizes technical complexities. The book consists of five substantive chapters. Chapter 2 introduces fundamental network assumptions and concepts, as applied to a variety of units of observation, levels of analysis, and types of measures. It contrasts relational contents and forms of relations and distinguishes between egocentric and whole networks. The structural approach emphasizes the value of network analysis for uncovering deeper patterns beneath the surface of empirical interactions. Chapter 3 concerns issues in collecting network data: boundary specification, data collection procedures, cognitive social structures, missing data, measurement error, and collecting online social media and Big Data. In Chapter 4, we discuss basic methods of network analysis, including graphs and matrices; centrality, prestige, and power; social distance, paths, walks, and reachability; transitivity and cliques; and size, centralization, density, and different measure of equivalence for pairs of actors or entities. Chapter 5 gives an overview of more-advanced methods of network analysis, including ego-nets; clustering, multidimensional analysis, and blockmodels; 2-mode and 3-mode networks; community detection; and exponential random graph models. The final

section concludes with some speculations about future directions in social network analysis.

After years of painstaking efforts, network analysts developed several computer packages to facilitate social network data collection and analyses. Softwares vary on many dimensions, such as operating systems, affordability, learning curves, and strengths and weaknesses. We attached an Appendix that summarizes some useful packages and contrasts them on those dimensions. We remain most impressed, however, with the breadth and user-friendly qualities of UCINET (Borgatti, Everett, & Freeman, 2002) as both a teaching and a research tool for smaller-scale social network analyses. Consequently, we used it to make this edition whenever we demonstrated social network analysis methods.

# Chapter 2

## NETWORK FUNDAMENTALS

In this chapter, we discuss fundamental concepts for understanding social network analysis methods. We use terms and definitions most widespread and accepted by academic researchers but in instances of disagreement defer to sociological perspectives. We cite many examples from diverse disciplines that illustrate these basic concepts. Interested readers should read numerous publications to deepen their understanding of how network analysis methods can be applied to investigate substantive problems in their fields.

To clarify the distinctive social network perspective on social action, a contrast to individualistic, variable-based approaches may be insightful. Many social science theories, possibly a large majority, assume that actors make decisions and act without regard to the behavior of other actors. Whether analyzed as utility-maximizing rational calculations or as drive-reduction motivation based on causal antecedents, such explanations primarily consider only the characteristics of persons while ignoring the broader interaction contexts within which social actors are embedded. In contrast, network analysis explicitly assumes that actors participate in social systems connecting them to other actors and that their relations comprise important influences on one another's behaviors. Central to the theoretical and methodological agenda of network analysis is identifying, measuring, and testing hypotheses about the structural forms and substantive contents of relations among actors. This distinctive structural-relational emphasis sets social network analysis apart from the individualistic, variable-centric traditions still prevalent in much social science theory and research. We see encouraging signs that many social science disciplines are increasingly embracing structural-relational explanations of social action.

### 2.1 Underlying Assumptions

The network perspective emphasizes *structural relations* as its key orienting principle. Siegfried Nadel, the great British anthropologist, proposed a relational definition of social structure: "We arrive at the structure of a society through abstracting from the concrete population and its behaviour the pattern or network (or 'system') of relationships obtaining 'between actors in their capacity of playing roles relative to one another'" (Nadel, 1957, p. 12). By network, he meant "the interlocking of relationships

whereby the interactions implicit in one determine those occurring in others" (p. 16). By separating structural forms from their empirical contents, structural analysts can uncover the underlying systems of roles that arise from interdependent activities of the persons performing those roles. Nadel further contributed to nascent network science by suggesting that matrix methods could graphically depict network relations. Nadel's conceptualization of networks as relational social structures was widely adopted by social network theorists and researchers over the ensuing decades of development. For example, Harrison White and his colleagues defined social structure as "regularities in the patterns of relations among concrete entities; it is not a harmony among abstract norms and values or a classification of concrete entities by their attributes'' (White, Boorman, & Breiger, 1976, pp. 733–734). More recently, the core mechanisms in Crossley and Krinsky's (2016) relational approach to sociology are interactions, relations, and networks. In network analyses, the entities may be individual natural persons, small groups, organizations, or even nation-states. Some types of network entities lack agency, such as documents posted on websites and participatory events such as sports matches and social movement protests. The patterns of relations connecting members of one or more sets of entities comprise the macrosocial contexts, or overall relational structures, that influence actor perceptions, attitudes, beliefs, decisions, and actions. The primary objectives of network analysis are to measure and represent these structural relations accurately and to explain both why they occur and what their consequences are.

Social network analysis rests on three underlying assumptions about structural relations and their consequences. First, structural relations are often more important for understanding observed behaviors than are such characteristics as race, gender, age, socioeconomic status, and political ideology. For example, research on voting behavior and social movement participation found that egocentric network structures more strongly influence people's choices than respondent attributes (Diani, 2004; Huckfeldt & Sprague, 1987; Knoke, 1990). Many actor attributes remain unaltered across the numerous social settings in which they participate (a woman's age, race, and education remain unchanged whether at home, at work, and at church). In contrast, many structural relations occur only at specific time-and-place locales and either vanish or are suspended when participants are elsewhere (e.g., student-teacher and doctor-patient relations do not exist outside school and clinic settings, respectively). A man holding a menial factory job requiring little initiative may be the dynamic leader of his church and an enthusiastic softball team player. Such behavioral differences are difficult to reconcile with unaltering gender, age, and status attributes but comprehensible on recognizing that

people's structural relations can vary markedly across social contexts within which they are embedded. The structural-relational explanations favored by network analysts depart markedly from substantialist approaches premised on static ''thing-concepts'' as their primary units of analysis: essences, self-action, norm-based conformity, rational choice, and variable-centric and social identity approaches (Emirbayer, 1997). In assuming that patterned relations influence social entities apart from their attributes, network analysis offers distinctive theoretical and empirical explanations of the origins of social action.

Second, social networks affect actor perceptions, beliefs, and actions through diverse structural mechanisms that are socially constructed by relations among entities. Direct contacts and more-intensive interactions dispose people and organizations to be better informed, more aware, and more susceptible to influencing or being influenced by others. Indirect relations through intermediaries (in popular imagery, agents who broker connections for their clients) also bring exposure to new ideas and potential access to useful resources that may be obtained through exchanges with others. For example, in a classic network study by Mark Granovetter (1973), job seekers typically obtained less useful information from their intimate circles, whose members already shared and circulated the same intelligence, than from their weaker and more distant social contacts. Relational structures provide complex pathways for assisting or hindering flows of knowledge, gossip, and rumor through a population (Fang, McAllister, & Duffy, 2017). A variety of structural-relational factors explains racial differences in the spread of HIV/AIDS infections among young men who have sex with men (Mustanski, Birkett, Kuhns, Latkin, & Muth, 2015) and the propagation of financial distress through the international banking network during the global financial crisis of the aughts (Kojaku, Cimini, Caldarelli, & Masuda, 2018). Physical illness, mental health, and recovery from substance abuse are strongly affected by people's social support networks (Cullen, Mojtabai, Bordbar, Everett, Nugent, & Eaton, 2017; Stevens, Jason, Ram, & Light, 2015), with social media exerting some unusual impacts (Lu & Hampton, 2017; Pallotti, Tubaro, Casilli, & Valente, 2018). Structural relations are vital to building cohesion and solidarity within a group but may also reinforce prejudices and intensify conflict with outgroups (Bliuc, Faulkner, Jakubowicz, & McGarty, 2018; Roversi, 2017). Competitive and cooperative relations enable innovation in corporate supply chains (Delgado-Márquez, Hurtado-Torres, Pedauga, & Cordón-Pozo, 2018), mobilization for collective action by social movements (Diani, 2016), and the operation of ''dark networks'' for drug trafficking, immigrant smuggling, and terrorist campaigns (Wu & Knoke, 2017). By channeling information, money, and other types of resources to particular

structural locations, networks help to create interests and shared identities and to promote shared norms and values. Network analysts seek to uncover the mechanisms through which social relations affect social entities and to identify the contingent conditions under which particular mechanisms operate in specific empirical contexts.

The third underlying assumption of network analysis is that structural relations should be viewed as dynamic processes. This principle recognizes that networks are not static structures but are continually changing through interactions among people, groups, or organizations. In applying their knowledge about networks to leverage advantages, network entities also transform those structural relations, both intentionally and unintentionally. For instance, in an intervention experiment to reduce conflict and bullying among students in 56 schools, experimenters comprehensively measured every school's networks, then randomly selected "seed groups" of 20 to 32 students to be encouraged to take public stands against conflict (Paluck, Shepherd, & Aronow, 2016). Disciplinary reports of conflict fell by 30% in the treatment schools compared to control-group schools, but the effect was stronger for seed groups containing more students who attracted greater student attention. Apparently, those popular students changed their network peers' beliefs and behaviors by publicly stigmatizing conflict and bullying as less socially normative. Such dynamics exemplify the more general "micro-to-macro problem" in the theory of social action (Coleman, 1986). The core issue is how large-scale systemic transformations emerge out of the combined preferences and purposive actions of individuals. Because network analysis simultaneously encompasses both structures and entities, it provides conceptual and methodological tools for linking changes in actors' microlevel choices to macrolevel structural alterations. The increased availability of longitudinal datasets, especially large online networks, coupled with methodological developments for analyzing multilevel relations, are accelerating research on cross-level dynamic processes (Lazega & Snijders, 2015; Snijders, Steglich, & Schweinberger, 2017). Likewise, developments in temporal exponential random graph models (TERGMs) and stochastic actor-oriented models (SAOMs), such as SIENA, hold great promise to advance our understanding of network dynamics (Leifeld & Cranmer, 2019; Leifeld, Cranmer, & Desmarais, 2018).

## 2.2 Entities and Relations

The two indispensable elements of any social network are entities and relations. Their combination jointly constitutes a social network, as described in the next subsection. *Entities* may be individual natural persons or

collective actors such as informal groups and formal organizations. Common examples of individual actors include children on a playground, high school students attending a prom, employees in a corporate work team, staff and residents of a nursing home, and terrorists operating in a covert cell. Collective actors might be firms competing in an industry, voluntary associations raising funds for charities, political parties holding seats in a parliament, and nations signing a military alliance. Other types of entities lack human agency, such as bills debated in a legislature, dances attended by students, and books read by library patrons. Sometimes networks are comprised of diverse types of entities, such as a healthcare system consisting of doctors and nurses, patients, clinics, hospitals, laboratories, insurance companies, and governmental regulations.

A *relation* is generally defined as a specific kind of contact, connection, or tie between a pair of entities, or *dyad.* Relations may be either *directed,* where one actor initiates and the second actor receives (e.g., advising, selling), or *undirected,* where mutuality occurs (e.g., conversing, collaborating). A relation is not an attribute of one entity but is a joint dyadic property that exists only so long as both participants maintain their association. An enormous variety of relations among individual and collective entities may be relevant to representing network structures and explaining their effects. At the interpersonal level, children befriend, play with, fight with, and confide in one another. Employees work together, discuss, advise, trust, undermine, and betray. Among collectivities, corporations exchange goods and services, communicate, compete, sue, lobby, and collaborate. In healthcare systems, physicians refer patients to specialty clinics, pharmacies, laboratories, hospitals, imaging centers, nursing homes, and hospices. Which specific type of relation a network researcher should measure depends on the particular objectives of the research project. For example, an investigation of community networks will likely examine various neighboring activities, whereas a study of banking networks would investigate financial transactions. Of course, some analyses scrutinize multiple types of relations, such as the political, social, and economic ties among corporate boards of directors. We present a general classification of relational contents in the next subsection.

Social science researchers rely heavily on measuring and analyzing the attributes of individual or collective units of analysis, whether through survey, archival, or experimental data collection. Although attributes and relations are conceptually distinct approaches to investigating social behavior, they should not be viewed as mutually exclusive options. Instead, many entity attributes can be reconceptualized as relations connecting dyads. For example, a nation's annual volumes of exports and imports are characteristics of its economy. But, the amount of goods and services exported and

imported between all pairs of nations represents the structure of trading networks in the global economy. Patents awarded to scientists employed at high-tech firms indicate companies' research innovations, but patent-citation networks reveal how knowledge flows through industries (Zhang, Kong, Zheng, Wan, Wang, Hu, & Shao, 2016). The number of friends indicates a child's popularity, but only network analyses of all dyadic friendship choices can uncover important cliques and clusters. Relations reflect emergent dimensions of complex social systems that cannot be captured by simply displaying a variable's distribution or averaging its members' attributes. Structural relations potentially influence both individual behaviors and systemic outcomes in ways not reducible to entity characteristics. For example, efforts to control sexually transmitted infections among injection drug users and sex workers require knowledge of both social and geographic distances among street people. Researchers identified 101 "hotspots" of high-risk activities in Winnipeg, Canada, where "the combination of spatial and social entities in network analysis defines the overlap of vulnerable populations in risk space, over and above the person to person links" (Logan, Jolly, & Blanford, 2016). An experiment in a large environmental nongovernmental organization found that "boundary spanners"—individuals who cross internal boundaries, such as departmental or geographic location, via their informal social networks—were more likely to diffuse innovations, although positions in a formal organizational hierarchy mediated this activity (Masuda, Liu, Reddy, Frank, Buford, Fisher, & Montambault, 2018). The strong inference is that exclusively focusing on actor attributes loses many important explanatory insights provided by network perspectives on social behavior.

## 2.3 Networks

A *social network* is a structure composed of a set of entities, some of whose members are connected by a set of one or more relations. These two fundamental components are common to most network definitions; for example: "a network contains a set of objects (in mathematical terms, *nodes*) and a mapping or description of relations between the objects or nodes" (Kadushin, 2012, p. 14). Different types of relations identify different networks, even where observations are restricted to the same set of entities. Thus, the friendship network among a set of office employees very likely differs from their advice-seeking network. Stating that connections exist among members of a network does not require that all members have direct relations with all others; indeed, sometimes very few dyads have direct links. Rather, network analysis considers both present and absent ties and

possibly also variation in the intensities or strengths of the relations. A configuration of empirical relations among entities identifies a specific *network structure,* the pattern or form of that network. Structures can vary dramatically in form, ranging from isolated structures where no actors are connected to saturated structures in which everyone is directly connected. More typically, real networks exhibit intermediate structures in which some entities have more numerous connections than others. A core problem in network analysis is to explain the occurrence of different structures and, at the entity or nodal level, to account for variation in linkages among entities. The parallel empirical task in network research is to detect and represent structures accurately using relational data.

The first researcher credited with using the term *social network* was John A. Barnes (1954), an anthropologist who studied the connections among people living in a Norwegian island parish. Barnes viewed social interactions as a ''set of points some of which are joined by lines'' to form a ''total network'' of relations (Barnes, 1954, p. 43). The informal set of interpersonal relations composed a ''partial network'' within this totality. Barnes drew on the work of Jacob Moreno (1934), whose hand-drawn *sociograms* of lines and labeled points displayed children's likes and dislikes of their classmates. We discuss methods for representing networks visually as graphs and mathematically as matrices in Chapter 4. From anthropology and sociology, network ideas and methods diffused over the past half century to many disciplines, which adapted them to prevailing theories and problems. For historical overviews of the origins and diffusion of network principles, see Freeman (2004, 2011); Knox, Savage, and Harvey (2006); Kadushin (2012); and Scott (2017).

If network analysis were merely a conceptual framework for describing how a set of actors is linked together, it would not have excited so much interest and effort among social researchers. But, as an integrated set of theoretical concepts and analytic methods, social network analysis offers more than accurate representations. It proposes that, because network structures affect actions at both the individual and systemic levels of analysis, network analysis can explain variation in structural relations and their consequences. J. Clyde Mitchell's (1969, p. 2) definition of social networks emphasized their impacts on outcomes: ''a specific set of linkages among a defined set of persons, with the additional property that the characteristics of these linkages as a whole may be used to interpret the social behavior of the persons involved.'' The first edition of this book underscored this perspective: ''The structure of relations among actors and the location of individual actors in the network have important behavioral, perceptual, and attitudinal consequences for the individual units and for the system as a whole'' (Knoke & Kuklinksi, 1982, p. 13). Similarly, Barry Wellman

(1999, p. 94) wrote, ‘‘Social network analysts work at describing underlying patterns of social structure, explaining the impact of such patterns on behavior and attitudes.’’

## 2.4 Research Design Elements

Three elements of network research design shape the measurement and analysis strategies available to researchers: social settings, relational form and content, and level of data analysis. Every network data collection project must involve making explicit choices about these elements before beginning fieldwork. Varying combinations of them generate the wide range of social network investigations published in the research literatures of numerous disciplines.

**Social Settings.** The first steps in designing a network study are to choose the most relevant social setting and to decide which entities in that setting comprise the network entities. Ordered on a roughly increasing scale of size and complexity, a half-dozen basic units from which samples may be drawn include individual persons, groups (both formal and informal), complex formal organizations, classes and strata, communities, and nation-states. Some two-stage research designs involve a higher-level system within which lower-level entities comprise the actors. Common examples are hierarchical social settings such as corporations with employees, schools with pupils, hospitals with physicians, municipal agencies with civil servants, and universities with colleges with departments with professors.

The earliest and still most common network projects select small-scale social settings—classrooms, offices, factories, gangs, social clubs, schools, villages, artificially created laboratory groups—and treat their individual members as the actors whose relations comprise the networks for investigations. Recent examples include bullying and homophobic teasing among middle school students (Merrin, De La Haye, Espelage, Ewing, Tucker, Hoover, & Green, 2018), helping and gossip networks among employees of a Turkish retail clothing company (Erdogan, Bauer, & Walter, 2015), and the effects of ethnic diversity on the spread of word-of-mouth information in two matched rural Ugandan villages (Larson & Lewis, 2017). Small settings have considerable advantages in sharply delineated membership boundaries, completely identified populations, and usually researcher access by permission from a top authority. However, network analysis concepts and methods are readily applied to larger-scale formations, many of which have porous and fuzzy boundaries, including clandestine networks. Examples include peer network origins of adolescent dating behavior (Kreager, Molloy, Moody, & Feinberg, 2016), criminal organizations in

communities of Calabria, Italy (Calderoni, Brunetto, & Piccardi, 2017), and strategic alliances among multinational corporations in the Global Information Sector (Knoke, 2009).

**Relational Form and Content.** Network researchers must decide on which particular relations to collect data. Relations among pairs of social actors have both form and content, a dichotomy that Georg Simmel (1908) proposed in his classic analyses of association. The two elements are empirically inseparable and only analytically distinguishable. *Contents* are the interests, purposes, drives, or motives of individuals in an interaction, whereas *forms* are modes of interaction through which specific contents attain social reality. Simmel argued that the task of sociology is to identify a limited number of forms—sociability, superiority, subordination, competition, conflict, cooperation, solidarity—that occur across a wide range of concrete settings, social institutions, and historical contexts. A particular form can vary greatly in content. For example, the basic forms of superordination and subordination are ever present in government, military, business, religious, athletic, and cultural institutions. Conversely, diverse contents like economic interests and drives for power are manifested through forms of competition and cooperation.

The form-content dichotomy also applies to social network analysis. *Relational form* is a property of relations that exists independently of any specific contents. Two fundamental relational forms are (a) the intensity, frequency, or strength of interaction between pairs of entities and (b) the direction of relations between both dyad members—null, asymmetric, or mutual choices. *Relational content* refers to its ''substance as reason for occurring'' (Burt, 1983, p. 36). Substantive content is an analytic construct designed by a researcher to capture the meanings of a relation from the informants' subjective viewpoints. When people are asked, ''please identify your close friends, friends, and acquaintances" in some social setting, the intended relational content is "friendship." The results of this query depend on how each actor first conceptualizes the meanings of the three proffered response categories and then classifies the other actors according to recollections of diverse interpersonal interactions. Obviously, people may vary markedly in their interpretations of both the friendship labels and those activities that they consider to be indicators of greater or lesser intimacy. Friendship dyads are never precisely reciprocated and the level of intimacy may be very unequal; for example, one dyad member considers the second person a "best friend," but the second member views the first person as a "friend." The National Study of Adolescent Health (Add Health) found that girls and Asian Americans were most likely to have reciprocated friendships, whereas interracial friendships were much less common than friendships between students of the same race (Vaquera & Kao, 2008).

The choice of relational content, also called *type of tie,* is largely determined by a project's theoretical concerns and research objectives. A study of healthcare networks could inquire into people's interpersonal sources of trusted information and advice about health-related matters, whereas a project on political networks might ask them to identify others with whom they discussed or participated in political affairs. Some substantive problems imply that more than one analytically distinct relational content should be investigated, in which case measuring and simultaneously analyzing two or more types of ties (i.e., *multiplex networks*) is an appropriate strategy. For example, psychologists asked 132 undergraduates at Midwestern University to list their Facebook friends who fulfilled each of five social functions (i.e., types of ties): sharing social activities, discussing personal matters, providing instrumental support, providing emotional support, and sharing success and happy events (Gillath, Karantzas, & Selcuk, 2017). Students with higher attachment avoidance were likely to ascribe fewer multiplex social roles to their networks' members, implying a lower degree of social trust.

Inexplicably, network analysts have conducted little research on the connections among diverse domains of relational contents. Ronald Burt (1983) examined survey respondents' perceptions of relational contents and uncovered substantial confusion, redundancy, and substitutability among the 33 questions posed to a sample of Northern Californians. He concluded that just five key questions would suffice to recover the principal structure of relational contents in the friendship, acquaintance, work, kinship, and intimacy domains. However, we still need much more research on the similarities and differences of meanings that people attach to commonly used relational terms and labels in a wide variety of network settings. A cognitive map of the structural connections among relational content domains would enable researchers efficiently and accurately to select specific contents most relevant to their theoretical and substantive concerns.

Until that desideratum arrives, in the spirit of Simmel we propose a small typology of generic contents:

- Transaction relations: Entities exchange control over physical or symbolic media, for example, in gift giving or economic sales and purchases.
- Communication relations: Linkages between entities are channels through which messages may be transmitted.
- Boundary penetration relations: Ties consist of membership in two or more social formations, for example, voluntary associations or social movement organizations.

- Instrumental relations: Actors contact one another in efforts to obtain valued goods, services, or information, such as a job, an abortion, political favors, or religious salvation.
- Sentiment relations: Perhaps the most frequently investigated networks involve actors expressing their feelings of affection, admiration, deference, loathing, or hostility toward one another.
- Authority/power relations: These types of ties, usually occurring in formal hierarchical organizations, indicate the rights and obligations of position holders to issue and obey commands.
- Kinship and descent relations: These bonds of blood and marriage reflect relations among family roles.

**Levels of Analysis.** After deciding the social setting and the relational forms and contents, researchers have several alternative levels at which to analyze the structures in data that they collect for social network projects. Details of appropriate measures and methods appear in Chapters 3 through 5, but here we summarize four conceptually distinct levels of analysis that analysts could investigate.

The simplest level is the *egocentric* network, consisting of one actor (*ego*) and all other actors (*alters*) with which ego has direct relations as well as the direct relations among those alters. This set is also called ego's "first zone," in contrast to second and higher zones consisting of all the alters of ego's alters, and so on. If a network's size is *N* actors, an egocentric analysis would have *N* units of analysis. Each ego actor can, in turn, be described by the number, intensity, and other characteristics of its linkages with its set of alters, for example, the proportion of reciprocated relations or the density of ties among its alters. An egocentric analysis of incarcerated California youths indicated that respondents reporting no close friendships within the facility had lower postinterview misconduct than those who nominated peers, suggesting an influence or amplifying effect of friends on misbehavior (Reid, 2017). In some respects, egocentric analysis resembles typical attribute-based survey research, with a respondent's individual characteristics such as gender, age, and education supplemented by measures derived from that person's direct network relations. Egocentric network research designs are well suited to surveys of respondents who are unlikely to have any contact with one another. The 1985 General Social Survey of the adult U.S. population (Marsden, 1987) pioneered procedures for identifying and eliciting information about a respondent's alters, which we describe in some detail in Chapter 3.

A second level of analysis is the *dyadic network,* consisting of pairs of actors. If the order of a pair is irrelevant—as in marital status where persons

are either unmarried, cohabiting, married, separated, or divorced—a sample of $N$ actors has $(N^2 - N)/2$ dyadic units of analysis. But, if the direction of a relation matters, as in giving orders and taking advice, then the sample contains $(N^2 - N)$ ordered dyads. The most basic questions about a dyad are whether a specific type of tie exists between two actors, and, if so, what is the intensity, duration, or strength of that relation? A closely related issue is whether a dyad without a direct tie is nevertheless indirectly connected via ties to intermediaries (e.g., brokers, go-betweens). Typical analyses seek to explain variation in dyadic relations as a function of pair characteristics, for example, the homophily hypothesis that "birds of a feather flock together" or the complementarity hypothesis that "opposites attract." Dyadic empathy—"a combination of perspective taking and empathic concerns for one's romantic partner"—is associated with higher sexual satisfaction, relationship adjustment, and sexual desire of first-time parents (Rosen, Mooney, & Muise, 2017, p. 543).

A third level of network analysis is, unsurprisingly, *triadic relations*. A set of $N$ actors has $\left(\frac{N}{3}\right)$ triples, the number of ways to take $N$ actors, three at a time. All possible combinations of present and absent directed binary relations among the actors in a triple generates a set of 16 distinct triad types. A basic descriptive question for empirical network analysis regards the distribution of observed triads among the 16 types, a summary tabulation called the *triad census.* Substantive research on triadic structures concentrated on sentiment ties (liking, friendship, antagonism), with particular interest in balanced and transitive triadic relations (e.g., if A chooses B and B chooses C, does A tend to choose C?). Because we lack space to review triad analysis methods, interested readers should consult the research program of James Davis, Paul Holland, and Samuel Leinhardt (Davis, 1979) and a comprehensive treatment by Wasserman and Faust (1994, pp. 556–602) for details.

Beyond the three microlevels, the *whole network* (also called *complete network*) is the most important macrolevel of analysis. Researchers use the information about every relation among all $N$ actors to represent and explain an entire network's structural relations. Typical concerns are the presence of distinct positions or social roles within the system that are jointly occupied by the network actors and the pattern of ties within and among those positions. Although a whole network has $N$ actors and $(N^2 - N)$ dyads (assuming directed relations and self-relations are generally ignored), these elements add up only to a single system. Examining the causes or consequences of structural variation at the whole network level of analysis typically involves measures of the global structural properties.

An example is a Dutch online social network of more than 10 million users living in 438 municipalities (Norbutas & Corten, 2018). Communities with higher network diversity were more economically prosperous than less-diverse communities, whereas greater network density at the community level was negatively associated with prosperity.

The four levels of network analysis imply that emergent phenomena at one level cannot be simply deduced from knowledge of the relations at other levels. For example, transitivity of choice relations is a substantively important variable for theories of friendship formation (''a friend of my friend is my friend''), which can be observed at the triadic level but not at the egocentric or dyadic level. For another illustration, Mark Newman (2001) found that coauthorship networks in biomedical research, physics, and computer science were each structured as "small worlds," where only five or six steps were necessary to connect random pairs of scientists. However, biomedical research was dominated by many people with few coauthors, in contrast to other disciplines characterized by a few people with many collaborators (see also, e.g., Ebadi & Schiffauerova, 2016; Maggioni, Breschi, & Panzarasa, 2013). The adaptability of network principles and procedures to investigate structural relations across multiple levels of analysis underlies its bourgeoning popularity for theorizing about social action and guiding empirical research.