Introduction to
Social Network Analysis

NOAA OFFICE FOR COASTAL MANAGEMENT

Social Science Tools for Coastal Programs

Introduction to Social Network Analysis

About This Publication

Some of the most challenging decisions in coastal management stem from the relationship between people and the environment. NOAA provides technical assistance to coastal management professionals addressing complex human-based problems. This publication, "Introduction to Social Network Analysis," is the next addition in a series of guides developed to bring information to this audience about the use of social science tools in their field of work. For more information, or to obtain additional copies, contact our office at *coastal.info@noaa.gov.*

NOAA's Office for Coastal Management

"Coastal management" is the term used by communities and organizations striving to keep the nation's coasts safe from storms, rich in natural resources, and economically strong. The national lead for these efforts is NOAA's Office for Coastal Management, an organization devoted to partnerships, science, and good policy. This agency, housed within the National Ocean Service, oversees major initiatives that include the National Coastal Zone Management Program, Coral Reef Conservation Program, Digital Coast, and National Estuarine Research Reserve System.

Table of Contents

Introduction

If someone asked how decisions are made in your office, you might point to the organizational chart. Org charts represent the official structure, but it's often the informal communication network that plays an equal if not more important role.

A formal and informal structure exists for most decision-making processes, from the team level to the highest ranks in management. Understanding how the informal structure works will make an organization stronger. To aid in this task, people turn to a social network analysis.

A social network analysis is a research approach that draws from an array of disciplines, such as sociology, anthropology, marketing, and business management, to explore human interactions. Commonly referred to as SNA, a social network analysis uncovers human and organizational connections and subsequently "maps" visual depictions of a social network. Quantitative measures are used to further describe and characterize relationships. A social network analysis can be used to accomplish the following:

- Identify teams and individuals that play central roles
- Discover isolated teams and individuals
- Detect information-sharing bottlenecks
- Identify opportunities for knowledge-flow improvement
- Accelerate the flow of knowledge and information across functional and organizational boundaries
- Improve the effectiveness of formal communication channels
- Target opportunities where increased information flow will have the most impact
- Raise awareness of existing informal networks

This guidebook describes social network analysis basics including key concepts and theory; data collection; how to compile and assemble data; an explanation of common centrality measures; and method limitations. This guidebook contains introductory level material. To become more knowledgeable, a brief list of related resources and additional readings is available at the end.

Network Structure Basics

A network is often defined as “an informal, interconnected group or association of persons.” Metaphorically, a social network could be likened to a food web. Food webs are used by ecologists to understand the relationships among an array of organisms that consume one another for nutrition and energy exchange. In a strikingly similar way, social network models depict social interaction and information exchanges among people.

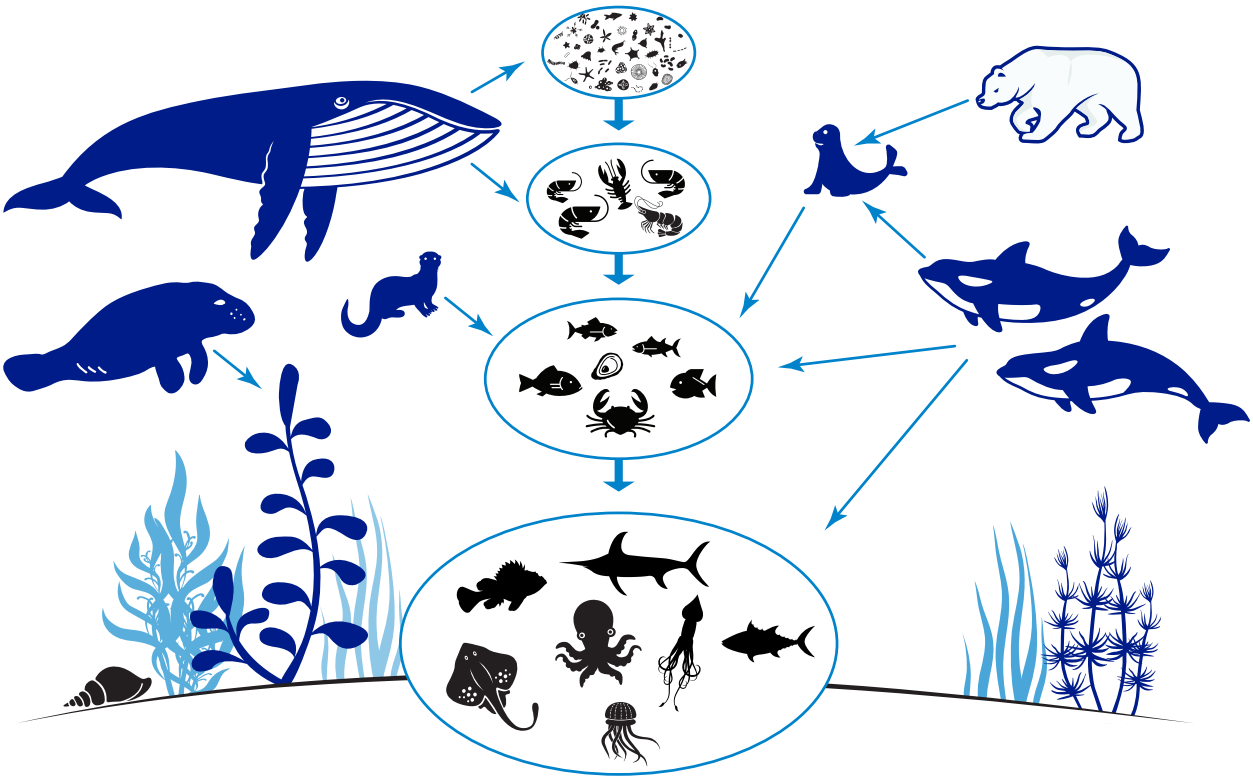

Figure 1: Food web to describe feeding relationship among species in a community

At its most basic, a social network is made up of a series of nodes and ties.

- Nodes (also referred to as actors or vectors) represent the people, organizations, or other entities within a network.
- Ties (or edges) are the lines that represent the connections among and between the various nodes.

For a social network analysis, the assemblage of nodes and ties is commonly referred to as a sociogram, or more informally, a network map. Sociograms can be hand-drawn or developed using sophisticated, computer-generated renderings.

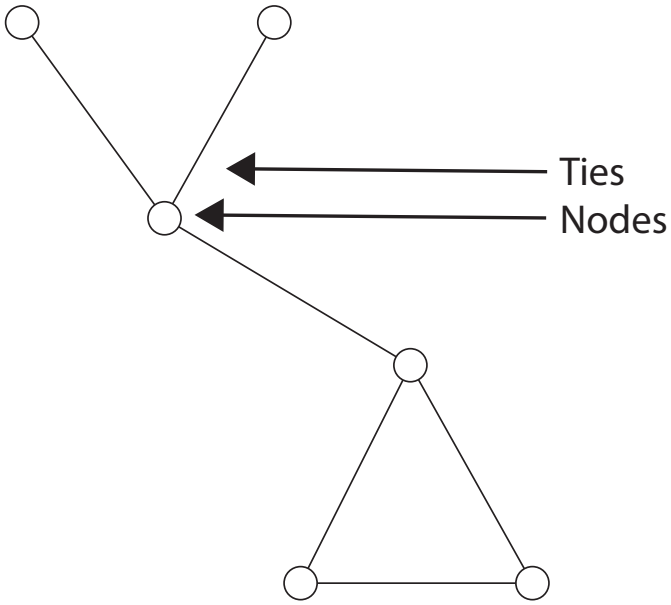

Figure 2: Depiction of undirected ties and nodes

Ties can be either directed or undirected (see Figure 2). A directed tie is shown by an arrow on one or both ends. An undirected tie is displayed with no arrows and is considered to be reciprocal by default.

Figure 3 shows an undirected network that displays reciprocal friendships among six coworkers. For example, Johanna is a friend of Katrina, and Katrina is also a friend of Johanna. The same can be said of all other actors where a tie is present. It is important to note that while such attributes are valid in both directions, they aren't necessarily equal. Katrina is friends with Gabe, Charles, and Johanna, but perhaps she likes Johanna most of all. This level or degree of friendship is not discernible in this sociogram. The presence and absence of ties among the actors give way to a number of quantitative measures, which are discussed later.

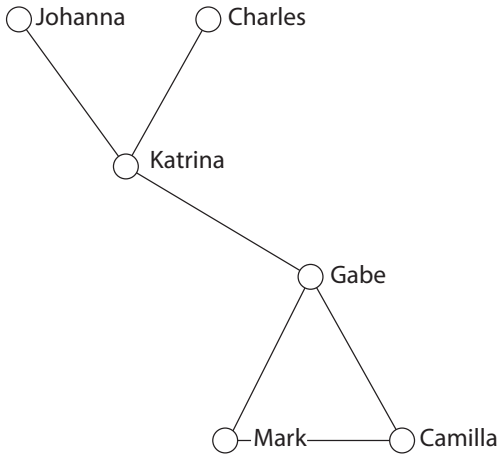

Figure 3: Depiction of actors in an undirected network

Directed ties offer the ability to show relationships that are not reciprocal. Let's look at the same group of actors (Figure 4). Instead of friendship, this visual represents information sharing on the job. According to the direction of the arrows, Katrina *receives* information from Gabe, Johanna, and Charles, but not Mark or Camilla. Katrina *provides* information to Johanna, but to no one else in the network.

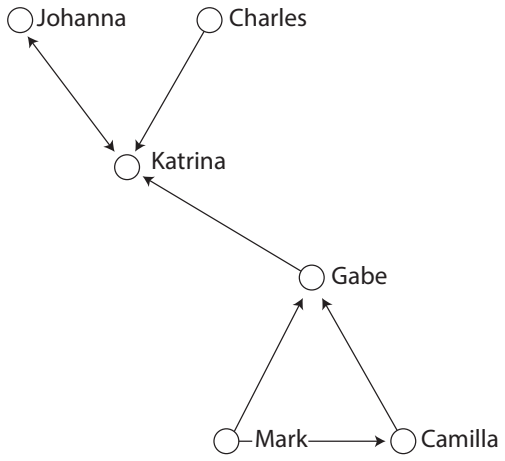

Figure 4: Depiction of actors in a directed network

Most commonly, networks display a single collection of actors, such as the one shown in Figure 4. These networks are referred to as single-mode (or one-mode) networks.

Two-mode networks display interrelationships from two different sets of actors or data. The two-mode network (Figure 5) is an example of a team of workers and the specific projects on which they work. This sociogram may not highlight direct, person-to-person ties, but it does offer insight into areas of weakness and opportunities for collaboration. Also notice Charles in the upper-left corner. Charles, referred to as an isolate in this network, is not tied to any of the numbered projects. While Charles may seem to have zero relevance to this network, the fact that he has no ties is actually valuable information and should be considered during data analysis.

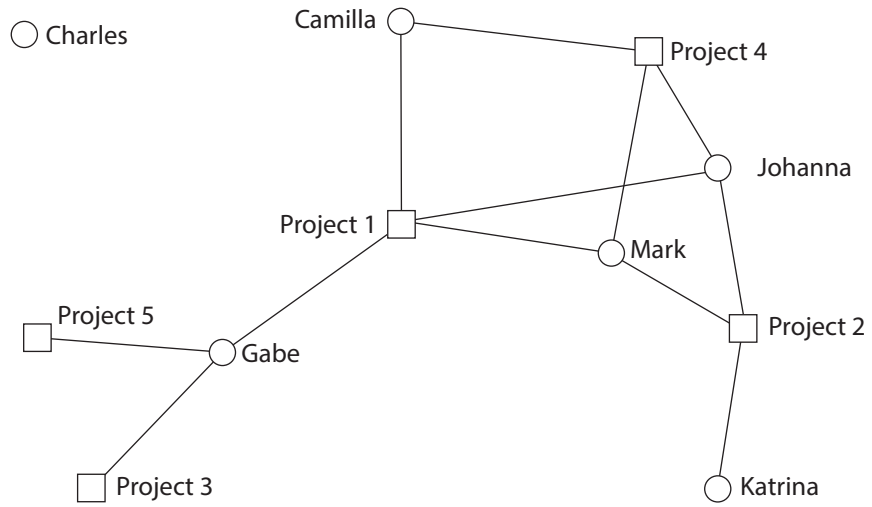

Figure 5: Depiction of actors and project activities in a two-mode network

Getting Started

What Do I Want to Learn?

Similar to any research-based endeavor, a clearly defined research question is key. Ask yourself—

- What am I interested in learning?
- What information will give me the ability to implement some sort of meaningful action?
- Who can offer answers to my question?
- Am I aware of, and do I have access to, these individuals?
- What time scale am I interested in learning about?
- What are the time, staff, technical, and financial resources available for this work?

Keeping these items in mind will help determine if social network analysis is a useful approach for your needs.

What Data Are Needed?

A social network analysis can be applied to any human transaction: professional information exchange, friendship, gossip, trust, loaning money, computer-based communication, and more. The information needed may already exist through secondary sources such as personnel or financial records, but more often than not, this isn't the case.

Sometimes the social network analysis need can be answered with a single question: Who do you trust? Who is the first person you go to for support? Who is your main contact in the office for flooding-related questions?

When more than a few simple questions are needed, perhaps a survey, or even a qualitative, interview-based approach, is more practical. Contextual information about actors may also be necessary. Contextual attributes can include gender, age range, job position, and number of years on the job.

How Are These Data Collected?

Data collection for a social network analysis can take many forms: human observation, surveys, and interviews, to name a few. But what if you don't know who these people are, or how to find them? This is a common challenge.

Sometimes the network is clearly defined and accounted for, such as a roster of students, an employee list, or known family members. When this is not the case, a common approach for unearthing a network of actors is to use snowball sampling, a sampling strategy based on leads provided by others.

Figuratively, the snowball rolls downhill and gets larger as it goes. If you collect data through a survey, for instance, you will likely find a few key individuals who can offer insight regarding the research question. This is your starting point. In the survey, ask participants to identify others they interact with who may also have pertinent insight. By doing this repeatedly, your survey sample begins to increase, subsequently building out the social structure of the network. After several iterations (collecting new contacts and repeating the survey), the contacts suggested typically become redundant; the same suggestions are made time and time again. This redundancy means you are reaching the extent of this network of individuals. This approach is tedious and time-consuming, but the process can provide high-quality results. Professional assistance using people trained in this method is recommended.

What about Analysis Software?

You have an array of social network analysis tools to choose from, including many that are free or low cost. (A few diverse examples include UCInet, Inflow, Pajek, R, and Gephi.) Primary considerations include ease of analysis (quantitative measures) and visualization, both of which are discussed in the following sections. Answering the following questions will also help you narrow your search.

- Will my social network analysis be used for research purposes?
- Do I intend to share visuals from my analysis in presentations and newsletters with my peers?
- Generally speaking, how many people are on my network? Tens? Hundreds? Thousands?
- Am I skilled, or do I have access to someone who is highly trained in social network analysis?
- Do I plan to conduct data analysis on my own, or with minimal assistance?
- Do I have a budget to purchase software (and obtain professional help)?

Data Management

The relational data of social network analysis follows a fairly simple format. A binary matrix is used to create the collection of nodes and ties and can be used to calculate a number of quantitative measures. Table 1 is an example of what the relational data look like. The value of "1" indicates the presence of a tie, and the value of "0" indicates the absence of a tie.

Table 1: An example of relational data displayed in a binary matrix

	Greg	Gail	Ronnie	Christina	Neil	Percy	Brenda	Lee	Tasi	Calvin	Eydie	Isaiah	Harriet	Maya	Claire	Amir	Evelyn
Greg	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0
Gail	1	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0
Ronnie	0	1	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0
Christina	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0
Neil	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0
Percy	1	1	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0
Brenda	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0
Lee	0	0	0	0	0	1	1	0	1	1	0	0	0	0	0	0	0
Tasi	0	0	0	0	0	0	0	1	0	1	1	0	1	1	1	0	0
Calvin	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	1	0
Eydie	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0
Isaiah	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0
Harriet	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	1
Maya	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	1	1
Claire	0	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0
Amir	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0
Evelyn	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0

In this example, all ties are reciprocal, meaning that relationships go in both directions. For example, in row one we see that Greg is connected to Gail (value represented by the number "1"), and in row two Gail is connected to Greg (also represented by the number "1"). However, if one of these values changed to a zero, then we would see a *directed tie* as described earlier, where you would see an arrow point in a single direction in the sociogram.

To complement the relational data, and offer further analytical options, an attribute data set is added. Attribute data are also simple and use numeric values to either quantify or categorize qualities possessed by network actors. In this attribute data (Table 2), office managers are identified by the number "1," and non-managers by the number "2." Similarly, those working in scientific positions are identified by the number "1" and non-scientific positions by the number "2." Attribute data allow the researcher to create visual displays within the sociograms, as shown in Figure 10.

Table 2: Example of an attribute data table

	Manager	**Scientist**
Greg	2	1
Gail	2	2
Ronnie	1	2
Christina	2	2
Neil	2	2
Percy	2	1
Brenda	2	1
Lee	1	2
Tasi	1	1
Calvin	2	2
Eydie	2	1
Isaiah	2	2
Harriet	2	1
Maya	2	1
Claire	1	2
Amir	2	2
Evelyn	2	2

Network Density and Centrality Measures

There are a number of quantitative measures used to analyze social networks. Following are several basic and frequently used measures.

Network Density

A simple, overall measure of a social network is network density. By definition, network density is a measure of how many connections there are between actors compared to the maximum possible number of connections that could exist between actors. Simply put, the higher the number of connections, the more dense the network. The network density value is typically represented as a coefficient, ranging from 0 to 1.0. A network with a coefficient value of 1.0 possesses all possible ties between all actors, such as in Figure 6, below.

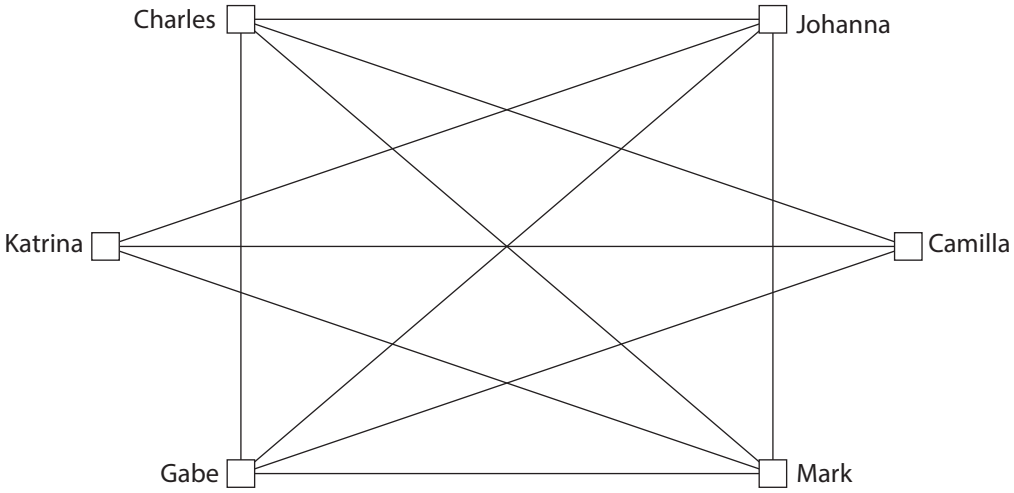

Figure 6: A network displaying the maximum number of ties, which has a density coefficient of 1.0

With fewer ties comes a sparser network and a reduced density coefficient. This is shown in the sociogram displayed previously in Figure 4, which has a network density value of 0.4.

In practical terms, information has the ability to flow more freely in a densely connected network. But keep in mind that increased density should not always be an intended goal or viewed as favorable. For example, in a professional office network, a high-density network may be counterproductive if employees receive too much information or unnecessary information. However, if a social network is being used to indicate friendship among coworkers, then a high-density network may be viewed as highly favorable. Interpreting the qualitative value of network density can be highly dependent on the situation, and what, precisely, is being measured.

Centrality Measures

There are numerous data tests that can be applied in social network analysis. The three primary measures used to assess the value of an actor's position in a network are degree, betweenness, and closeness centrality.

Degree Centrality

Degree centrality is a simple count of the number of ties or "connections" each actor in a network possesses. This measure assesses how well-connected actors are in a network, and their ability to direct influence. In the Figure 7 sociogram, Brenda possesses the highest degree centrality score – 7. From a practical standpoint, this measure helps us find people who are well connected or popular. These individuals are also likely to possess lots of information.

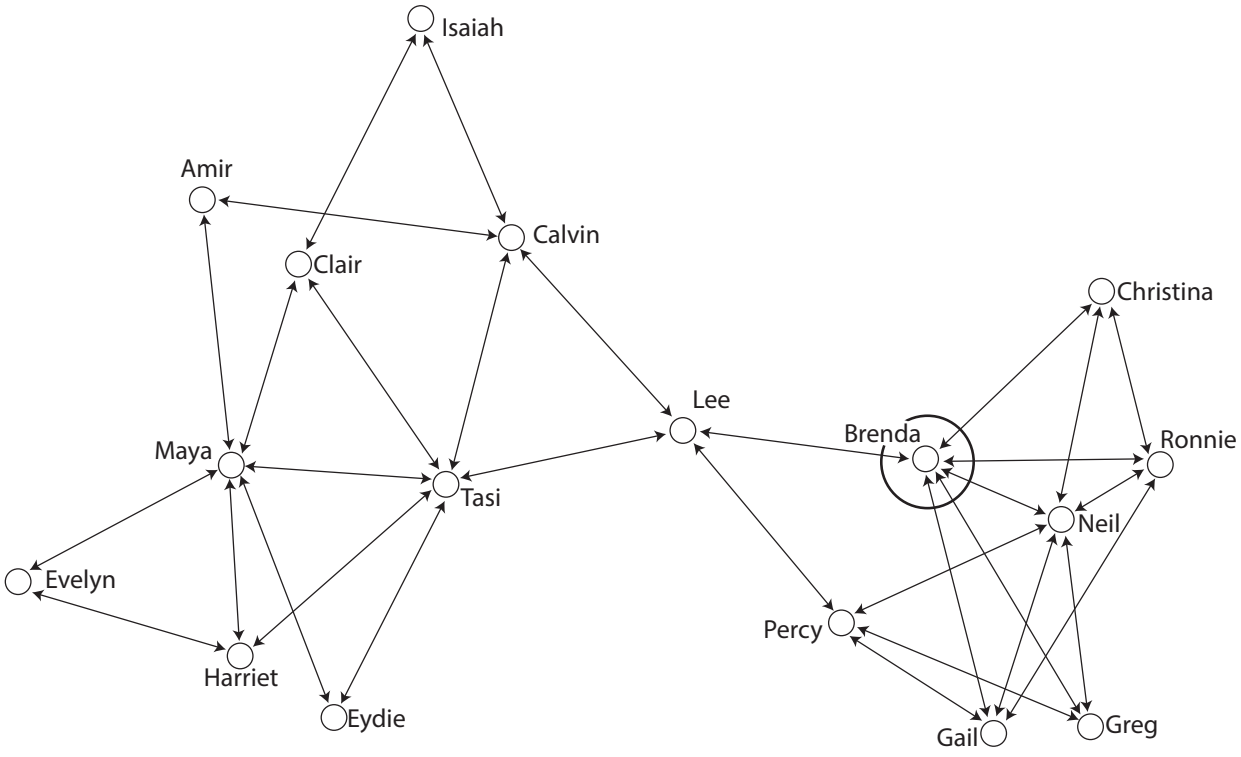

Figure 7: Sociogram highlighting the actor with the greatest degree centrality

Betweenness Centrality

A second commonly used measure in a social network analysis is betweenness centrality. This measure identifies actors' positions within the network by their likelihood of making connections to others in the network. A high betweenness value is indicative of actors possessing a favored position in the network and an increased likelihood of having the ability to exert greater influence on what happens in the network. Alternatively, those individuals are also potential points of failure should they leave the network.

In our example, Figure 8, Lee possesses the highest betweenness because he is between Calvin and Tasi on one side, and Brenda and Percy on the other. Though Brenda has greater degree centrality, Lee is likely a more important actor in this network.

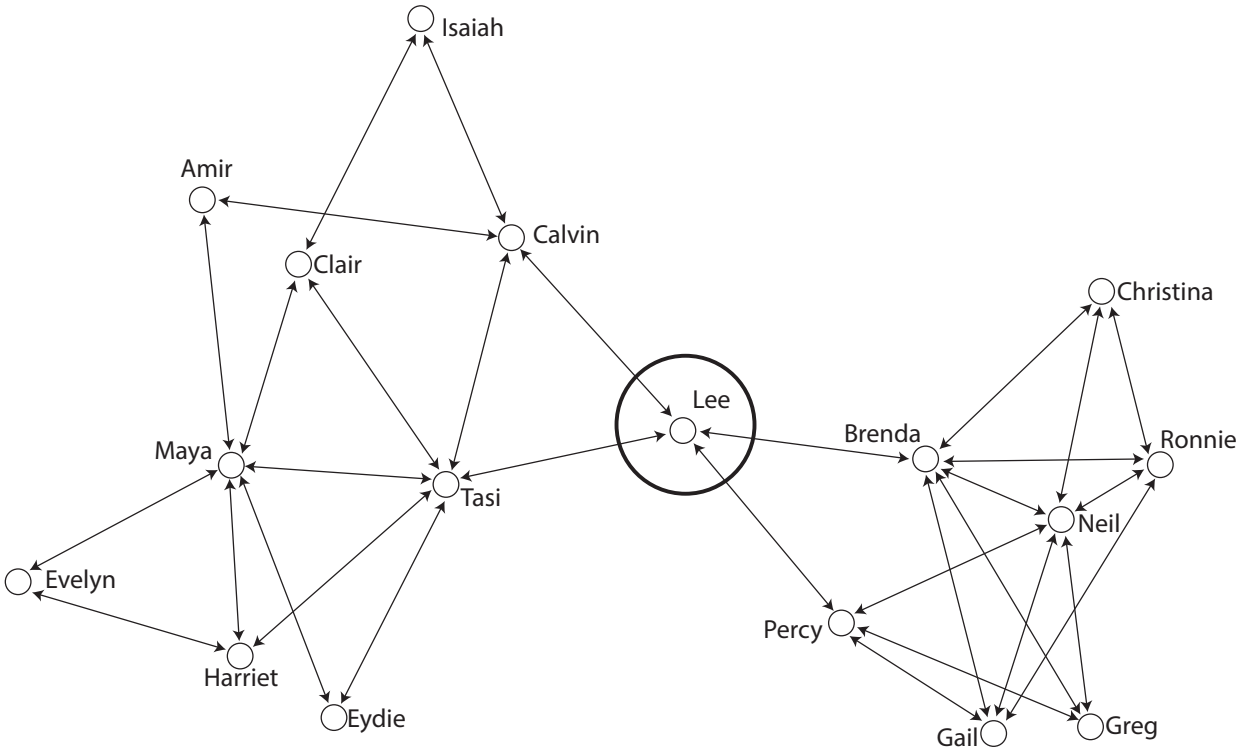

Figure 8: Sociogram highlighting the actor with the greatest betweenness centrality

Closeness Centrality

Closeness centrality measures how a person (actor) can reach all others with the fewest number of "steps" from node to node. Those with high closeness centrality typically have easy access to others in the network and may have more of a collective perspective as to what is happening in the network. As with betweenness centrality, Lee has the highest closeness centrality score (Figure 9) because he is best positioned in the network with the fewest number of steps to all others. A close second in the sample sociogram is Tasi.

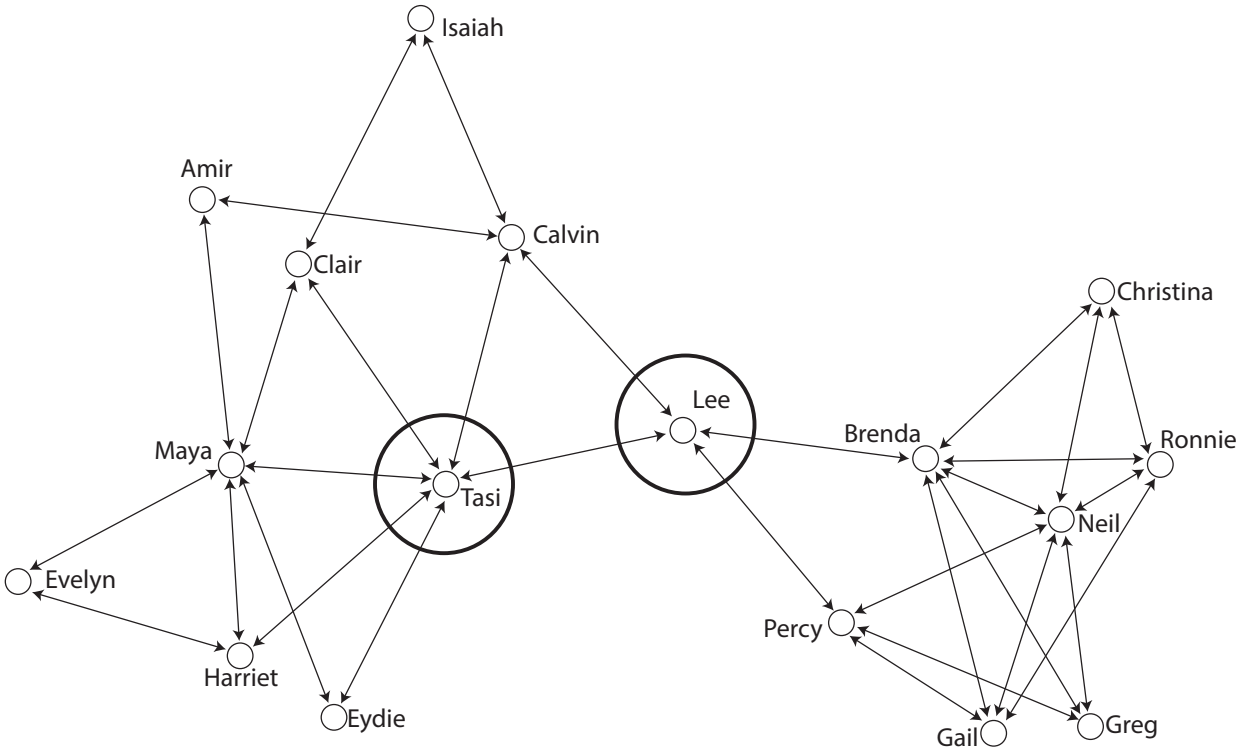

Figure 9: Sociogram highlighting actors with the highest measure of closeness centrality

Network Attributes

Unique attributes of each actor can be added to a sociogram to help further describe the characteristics and qualities of the actors and their relationships.

Attributes can be numeric (years of service to an organization), categorical (brand of car one drives), or qualitative (names). Actor attributes can be visually displayed in a sociogram by node shading or colors, by node size, by node shape, and through the use of labels.

These visual attributes help describe some social network dynamics at a glance. In the office worker example, Figure 10, nodes represent those individuals with managerial positions (shaded node = managerial), and the shape highlights the number of years on the job (triangle = 0 to 10 years; circle = 10+ years).

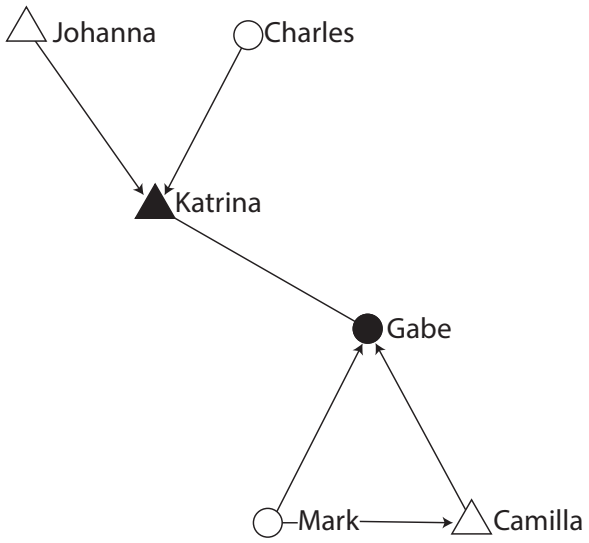

Figure 10: Actor attributes displayed by node shape and shading

Network Limitations

As with any research or analysis technique, social network analysis has its own collection of limitations to consider.

Trust

First and foremost, there is a substantial element of trust that must be addressed. Typically, within the social sciences, data and findings are reported anonymously. Quite contrarily, social network analysis data are "named" data that highlight who people are and their relationships with others. Paradoxically, a sociogram without some identifying characteristic (e.g., names) is not actionable or useful. The researcher must therefore take great care in how these data are collected and shared. Sharing or reporting social network analysis data in a group setting can lead to discomfort among those named in the network, which could eventually lead to conflict.

When collecting these data, the researcher must clearly communicate how information will be used, and shared, to minimize potential discomfort and conflict. This is where the critical element of trust comes into play. The researcher must have an established record of trust among study participants and must strictly adhere to the terms set forth with study participants. Even then, conflict may arise. Much thought and consideration is required in this area.

Self-Reported Data

Another challenge has to do with the self-reporting aspect of the data, such as data acquired through a survey. Social desirability bias helps explain why this can be problematic.

Social desirability bias is the tendency of a survey respondent to answer questions in a manner they think will be viewed as favorable by others. This could mean exaggerated reporting of connections to people they feel are good, intelligent, or influential, and under-reporting ties to those they deem as less desirable.

Here's a practical example. You are interested in understanding how environmental regulatory decisions are made in your county, and a social network analysis survey is used to uncover the communication network among all those involved. An actor in this network may overstate or understate their true degree of interactions or relationships based on personal values and who they feel most reflects the ideals they care about, making social desirability bias a possible downfall of the self-reported data.

Missing Data

Beyond the challenges of inaccurate data, there is a critical challenge when accounting for missing data. A social network analysis often has missing data, whether through incomplete responses from individuals or a lack of participation. This challenge is twofold. Missing data can mean critical individuals are missing from the social network, which leads to structural holes in the network. The problem is compounded with the use of centrality measures, which then give inaccurate weight or value to certain network actors because other actors are missing.

Summary

A social network analysis is a novel and highly insightful social research method. When performed properly, a wealth of knowledge can be uncovered about social and group dynamics. There are several considerations to take into account when designing a social network analysis, including sound research design and sampling strategy, appropriate data tests, and a way to meaningfully address method limitations.

A social network analysis represents data that you collected, and may not truly and realistically represent the full range of actual dynamics within a social network. Use a scientific approach and validate the results with the network actors to help ensure the quality and reliability of the findings.

References

Borgatti, Steve, Martin G. Everett, and Jeffrey C. Johnson. 2017. *Analyzing Social Networks*. 2nd ed. London: SAGE Publications.

Bryan, Lowell L., Eric Matson, and Leigh M. Weiss. 2007. "Harnessing the Power of Informal Employee Networks." *The McKinsey Quarterly.* Number 4. Pages 1-10.

Cross, Rob, Andrew Parker, and Stephen P. Borgatti. 2002. "A Bird's Eye View: Using Social Network Analysis to Improve Knowledge Creation and Sharing." Somers, NY: IBM Institute for Knowledge-Based Organizations. Pages 1-17.

DeJordy, Rich, and Dan Halgin. 2009. *NOAA SNA Workshop. Seattle, WA, December 16-18, 2009.*

Durland, Maryann M., and Kimberly A. Fredericks. 2005. "An Introduction to Social Network Analysis." *New Directions for Evaluation*. Fall 2005, Number 107. Pages 5-13.

Freeman, Linton C., Douglas R. White, and A. Kimball Romney. 1992. *Research Methods in Social Network Analysis*. New Brunswick: Transaction Publishers.

Haythornthwaite, Caroline, Barry Wellman, and Marilyn Mantei. 1995. "Work Relationships and Media Use: A Social Network Analysis." *Group Decision and Negotiation*. Number 4. Pages 193-211.

Holden, Ronald R., and Jennifer Passey. 2009. "Social Desirability." In *Handbook of Individual Differences in Social Behavior,* edited by Mark R. Leary and Rick H. Hoyle, pages 441-454. New York: Guilford Press.

Knoke, David, and Song Yang. 2008. *Social Network Analysis.* 2nd ed. Los Angeles: SAGE Publications.

Krackhardt, David. 1987. "Cognitive Social Structures." *Social Networks*. Volume 9, Number 2. Pages 109-134.

Krackhardt, David, and Jeffrey R. Hanson. 1993. "Informal Networks: The Company." *Harvard Business Review*. July-August 1993. Pages 104-111.

Magsino, Samantha L. 2009. *Applications of Social Network Analysis for Building Community Disaster Resilience: Workshop Summary.* Washington, DC: National Academies Press.

Mandarano, Lynn A. 2009. "Social Network Analysis of Social Capital in Collaborative Planning." *Society & Natural Resources*. Volume 22, Number 3. Pages 245-260. DOI: 10.1080/08941920801922182,

Monge, Peter R., and Noshir S. Contractor. 2001. "Emergence of Communication Networks." In *The New Handbook of Organizational Communication*, edited by Fredric M. Jablin and Linda L. Putnam, Chapter 12, Pages 440-502. Thousand Oaks: SAGE Publications,. DOI: 10.4135/9781412986243.

Prell, Christina, Klaus Hubacek, and Mark Reed. 2009. "Stakeholder Analysis and Social Network Analysis in Natural Resource Management." *Society & Natural Resources*. Volume 22, Number 66. Pages 501-518. DOI: 10.1080/08941920802199202.

Wasserman, Stanley, and Katherine Faust. 1994. *Social Network Analysis Methods and Applications.* Cambridge: Cambridge University Press.

Glossary

Actor – A person, group, organization, or other entity represented in a sociogram by a node

Attribute – A descriptive element used to distinguish an actor within a sociogram

Bottleneck – An area within a sociogram where exchange among actors is limited or impaired

Centrality – A collection of quantitative measures in a social network analysis that speaks to an actor's relative importance

- **Betweenness centrality** – The location of an actor within a network, calculated by looking at how many times an actor is positioned on the shortest path connecting two other actors
- **Closeness centrality** – Measure that underscores the independence of an actor. Being close to many other actors means that a particular actor will not have to rely on intermediaries, and is in a prominent position for communication or exchange.
- **Degree centrality** – The number of people (or other entity) a single actor is connected to within a network

Density – A measure of how many connections there are between actors compared to the maximum possible number of connections that could exist between actors.

Directed network – When ties within a network are directed (represented by arrows drawn between nodes) to indicate the path of communication flow

Ego network – All of the direct connections possessed by a single actor

Isolate – An actor that is disconnected from a network, possessing no ties

Node – A point in which ties intersect or branch. In a social network analysis, nodes represent an actor in the form of a circle, square, or other common shape.

One-mode network (or single-mode network) – Network with one set of nodes that are similar to each other

Pendant – An actor within a social network that is connected by only one tie

Relational data – Data that describes the ties among actors in a network

Snowball sampling – A nonprobability sampling method where existing study subjects recruit future subjects based on their acquaintances

Sociogram (or network map) – Graphic used in social network analysis to illustrate ties or connections among actors in a network

Tie – The pathway between two nodes, represented by a line

Two-mode network – Network that possesses two different sets of nodes, and ties displayed only between nodes belonging to different sets

Undirected network – A network that disregards the direction of communication flow

NOAA OFFICE FOR COASTAL MANAGEMENT

SEPTEMBER 2020