1

BECOMING FAMILIAR WITH SOCIAL NETWORKS

Each one of us has our own social networks, and it is easiest to start understanding social networks through thinking about our own. So what social networks do you have? These might include friendship networks, your network of colleagues at work, and the network of individuals you know from participating in various clubs and organizations. It is quite normal to be a member of many different social networks, and in fact, social network analysis encourages you to think along these lines by separating your various social networks according to different relations. Thus, a friendship network would be one relation, an advice network a different one, and a dislike network still another relation.

Breaking down social networks according to relations is aided by how a researcher phrases questions to a respondent. For example, if I start to ask you a series of specific questions regarding the different kinds of social networks you have, it becomes easier for you, the respondent, to conceptualize all the different social networks to which you belong.

Look at the questions below and make an attempt to answer these questions for yourself. In most cases, these questions will generate new lists of names, and in some instances, you will find the same individuals popping up as answers time and time again:

- who is in your immediate family?
- who do you tend to socialize with on weekends?
- whom do you turn to for advice in making important decisions about your professional career?
- whom do you turn to for emotional support when you experience personal problems?
- whom would you ask to take care of your home if you were out of town, for example watering the plants and picking up the mail?

Take a moment and answer each of the above questions, writing down the names of people who come to mind for each question. You will probably notice two things: that each question generates a slightly different list, and that certain names appear repeatedly in different lists. Each list represents a different social network for you and there is most likely overlaps in these networks. We can label these lists to give a name to the relation that the social network represents. Thus, you can have a 'family' network, a 'socializing' network, a 'career advice' network, an 'emotional support' network and 'home-care' network. In addition, you can add some more information about yourself and about each person in the lists, for example, their age and gender. Below is a fictitious example for the list entitled 'family':

In Table 1.1, I have simply listed the people in Susie's immediate family. The list is Susie's social network for the relation of 'family'. All the people listed in this social network are the *actors* in the network. As you will see, actors are also referred to in social network analysis as *nodes* and as *vertices*. In social network analysis, there are often multiple terms for the same concept, as this field has developed in many different disciplines. For example, an 'actor' is a more sociological term, whereas nodes and vertices are terms derived from graph theory. The network represented by the list in Table 1.1 is a specific kind of network in SNA, which is called an '*ego network*'. Ego networks are comprised of a focal actor (called ego) and the people to whom ego is directly connected. These people to whom ego is connected are referred to as '*alters*'. In this case, Susie is the ego and the alters are Janice, Emily and John. Susie holds a tie with each family member, and each family member's gender and age has also been listed. Gender and age are considered additional information on each particular actor, and we refer to these additional pieces of actor information as *actor attributes*. Actor attributes are the same sort of attributes you come across in more traditional social science research. They include categories such as age, gender, socio-economic class, and so forth.

Table 1.1 Susie's immediate family

Name	Gender	Age
Susie	Female	25
Janice	Female	21
Emily	Female	47
John	Male	48

Through this simple example, you have already learned a fair amount about social network analysis. In particular, you have learnt some of the fundamental terminology on which social network analysis is based, namely actors, nodes, vertices, ego and ego network, alters, ties, relations and actor attributes. A social network consists of all these pieces of information, and more formally, a *social* *network* can be defined as a set of relations that apply to a set of actors, as well as any additional information on those actors and relations.

Our example above, as simple as it is, can still provide us with a means for introducing some more SNA terms. Susie's ego network of her family represents a *state relation*. State relations have a degree of permanency or durability that make it relatively easy for a researcher to detect. Examples of state relations include kinship, affective relations such as trust or friendship, and affiliations such as belonging to the same club or church. State relations stand in contrast to *event relations*, which are more temporary sorts of relations that may or may not imply a more durable relation. Examples of event relations include attending the same meeting or conference; having a cup of coffee together; sending an email; giving advice; talking with someone; and fighting with someone. Event relations may or may not indicate a more permanent state. Usually, however, we think of event relations as individual occurrences.

Actors, vertices and nodes = the social entities linked together according to some relation

ego = the focal actor of interest

alters = the actors to whom an ego is tied

tie = what connects A to B, e.g. A is *friends* with B = A is *tied* to B.

relation = a specified set of ties among a set of actors. For example, friendship, family, etc.

actor attributes = additional information on each particular actor, for example, age, gender, etc.

ego network = social network of a particular focal actor, ego, ego's alters and the ties linking ego to alters and alters to alters

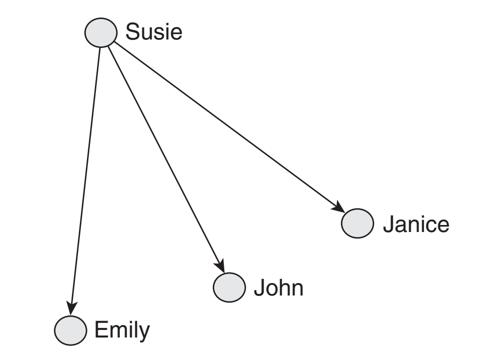
social network = a set of relations that apply to a set of social entities, and any additional information on those actors and relations

Figure 1.1 SNA terminology

DESCRIBING SOCIAL NETWORKS THROUGH GRAPHS AND GRAPH THEORY

Let's move on from this starting point. We can make a visual representation of Susie's network by drawing a graph. A *graph or digraph* is a visual representation of a social network, where actors are represented as nodes or vertices and the ties are represented as *lines*, also called *edges* or *arcs*.

When we represent social networks as graphs and describe social networks in terms of graphs, we use terms and concepts derived from graph theory, a branch of mathematics that focuses on the quantification of networks. Although social network analysis is not the same as graph theory, many of the fundamental concepts and terms are borrowed from this field, and so it is worthwhile to spend a bit of time familiarizing oneself with some of graph theory's basics.



Graphs consist of a set of undirected lines among a set of nodes. They are visual depictions of social networks.

Digraphs consist of a set of directed lines among a set of nodes.

Graphs are composed of nodes, also called vertices, which are connected through lines.

Directed lines are referred to as arcs. Undirected lines are edges.

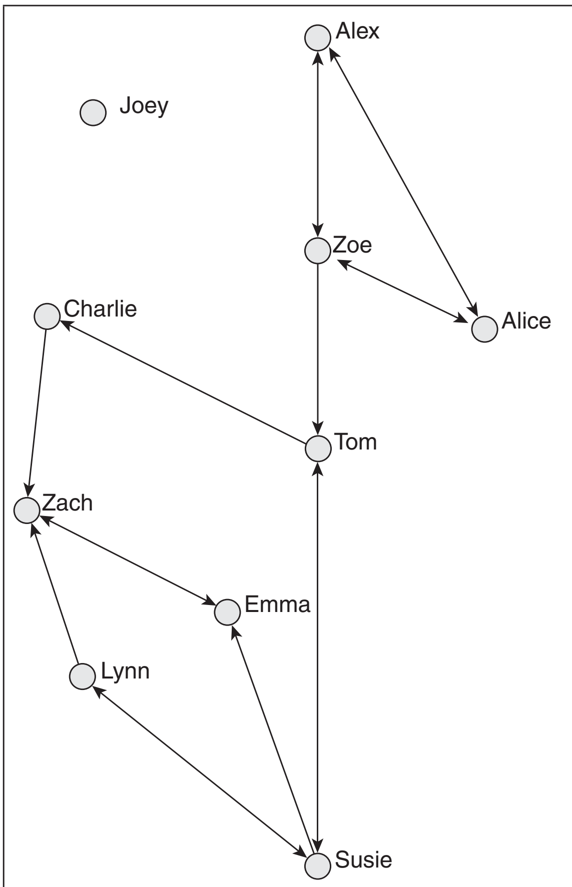
Figure 1.2 Digraph

In graph theory, ones says that there are n number of nodes and L number of lines. Thus, in the above graph, $n = 4$ and $L = 3$. In addition, we discuss how nodes are adjacent with each other. A node is adjacent to another node if the two share a tie between them. Thus, Susie is adjacent to each member of her family. She shares a tie with each member of her family.

To help you become better acquainted with some of the graph theory basics, I shall expand on our first example of a social network. A family network, as stated above, is one of many social networks of which an actor can be a member. Another network could be a friendship network. Friendship networks tend to span across different contexts: we have friends at our places of work, from our school, from our neighbourhood, and so forth. Sometimes these friends overlap, for example, a friend from our workplace might also live in our neighbourhood. For purposes of this present example, I would like to focus your attention on a very clearly bounded sort of friendship network, that is, the friends you have at work. For example, suppose I were to ask you the following question: 'Whom in your workplace would you consider to be your friend?' Most likely, your answer would not include every person with whom you work, but rather those people you feel closer to or more intimate with than the others. This would be considered your friendship network at work. Notice that this is quite different than if I were to ask you, 'Whom do you consider to be your friend?' without specifying whether or not I was interested in your workplace or not. In specifying the workplace, I have thus created a *boundary* around this particular network, i.e. I have defined what sorts of actors can be considered to be inside the network and which ones are outside my realm of interest. By specifying the boundary, I have, in essence, specified my population of interest for this particular network study. Network boundaries are an important issue which will be taken up later in this book. For now, it is good enough for you to understand that the boundary of this particular network is a particular, specified workplace.

Now suppose I were to ask the exact same question to each of your colleagues. That is, each person in your workplace were asked to nominate colleagues with whom they felt they shared a friendship tie. In asking every person at your work place the same question, I have moved from studying one particular person's ego network to studying a *complete network*. A complete network is one where an entire set of actors and the ties linking these actors together are studied. Once again, I can display a complete network as a graph, as shown in Figure 1.3.

Now I can introduce additional terms from graph theory to further describe this particular social network. You will notice that the above graph contains lines with arrowheads. The lines represent ties, and the arrowheads indicate the direction of the ties. In SNA terms, we would say this graph shows a *directed relation*, and in graph theory, a graph with directed lines is referred to as a *digraph*. The directed lines making up the digraph are referred to as arcs. Arcs have *senders* and *receivers*, where senders are the ones who nominate, and receivers are the nominees. In



In this digraph, the number of nodes is 10 ($n = 10$).

The number of lines is 17. There are 6 reciprocated arcs (so $6 \times 2 = 12$) and 5 unreciprocated ones.

Joey is an example of an isolate: he does not have a tie to any other actor in this network.

A walk refers to a sequence of nodes. An example of a walk from Zoe to Tom would be: Zoe, Alice, Alex, Zoe, and Tom.

A path is a walk where no node or line occurs more than once.

Both semiwalks and semipaths ignore the direction of arcs.

Figure 1.3 A digraph showing 'Who is friends with whom at work' the present example, the senders are the respondents who answered the question 'Whom in this workplace would you consider to be your friend?', and the receivers are those actors who were nominated by the respondents as friends.

If we did not want to pay attention to the direction of the lines, we could simply remove the arrowheads. In graph theory, a network that contains undirected lines is referred simply to as a graph, and the undirected lines are referred to as *edges*. Graphs, i.e. those which contains only edges, are considered the simplest to study, and thus you might choose to ignore the direction of the lines for this very reason. However, a graph might result from the nature of the relation being studied; for example, if the relation being studied is marriage, the marriage tie can be assumed to flow in both directions, so there is no need to represent that relation as a digraph.

You will also notice the Joey does not have any ties to other nodes in this graph. In this instance, Joey is an isolate, i.e. a node that does not have ties to any other actors in a network.

Displaying a social network as a graph or digraph invites a researcher to start describing certain aspects of the network. For example, we can discuss how close or far apart two nodes are from one another through the number of arcs or edges linking the two together. Considerations of distance between nodes involve the concepts of semiwalks, walks, semipaths and paths. I shall describe each of these briefly below:

In Figure 1.3, one can take a *walk* from Zoe to Tom by passing through Alex, Alice and Zoe. Notice that we are paying attention to the direction of the arcs in taking this walk, and that we passed by Zoe twice. Thus, a walk is a sequence of nodes, where all nodes are adjacent to one another, where each node follows the previous node, and where nodes and lines can occur more than once. The beginning node and the ending node in a walk can be different. Thus, the length of the walk is the number of lines that occur in a walk, and the lines that occur more than once in the walk are counted each time they occur. The walk from Zoe to Tom, as described above, is length 4. A much shorter walk from Zoe to Tom would be length 1.

A semiwalk simply ignores the direction of arcs. Thus, a semiwalk from Tom to Zoe is possible, even though the direction of the arc would suggest otherwise.

The notion of paths and semipaths builds on these ideas. A *path* is a walk in which no node and no line occurs more than once on the walk. Thus Zoe, Tom, Charlie and Zack would be an example of a path. Again, a semipath is a path that ignores the direction of arcs.

Although this seems like a lot of vocabulary, these terms and concepts are the building blocks for describing, analysing and theorizing social networks. In the remaining space of this chapter, we shall explore two other fundamentals regarding social network analysis. These are modes of networks and network matrices.

NETWORK MATRICES

The previous section focused on the display of networks as graphs and digraphs, and the terminology from graph theory used to describe these. Although visualizing networks as graphs and digraphs is useful, the reliance solely on these visual representations can become cumbersome and even chaotic as a network grows in size. For this reason (among others) network data are also organized as network matrices.

You are probably already familiar with a case-by-variable matrix: this is a matrix where rows in the matrix represent the individual cases, e.g. the respondents in your study, and columns represent certain variables related to these cases, such as age, gender, nationality and so on. Such a case-by-variable matrix is shown in Figure 1.4, with the cases represented by numbers in the rows, and the variables shown as columns. Each cell in this matrix contains a value that corresponds to the different levels of measurement (dichotomous, nominal, ordinal, interval and ratio). The values for age correspond to the level of interval, gender to dichotomous (where 1 = male and 2 = female), and nationality to nominal (where 1 = UK, 2 = USA, 3 = Canada and 4 = Australia).

A network matrix is slightly different to this case-by-variable matrix. In a network matrix, data are organized as case-by-case matrices (called adjacency matrices) or by case-by-events (called incidence matrices). The cells represent the presence or absence of ties. A cell is the intersection of a row with a column. Figure 1.5 is a very simple adjacency matrix. Here, each actor is represented twice: once in the row and once in the column. This matrix represents the graph based on friendship as we discussed above.

You will see that all the names in this social network are listed in the rows, and these names are abbreviated at the top of each column (thus, Susie becomes S and Emma becomes E, etc.). In addition, the rows and columns have been numbered. For example, Susie is referred to as row 1 and column 1.

	Age	Gender	Nationality
1	19	1	4
2	18	1	3
3	21	2	2
4…	22	2	1

Figure 1.4 Example of a case-by-variable matrix

		1 S	2 E	3 Z	4 T	5 L	6 Z	7 A	8 J	9 A	10 C
1	Susie	0	1	0	1	1	0	0	0	0	0
2	Emma	0	0	1	0	0	0	0	0	0	0
3	Zach	0	1	0	0	0	0	0	0	0	0
4	Tom	1	0	0	0	0	0	0	0	0	1
5	Lynn	1	0	1	0	0	0	0	0	0	0
6	Zoe	0	0	0	1	0	0	1	0	1	0
7	Alex	0	0	0	0	0	1	0	0	1	0
8	Joey	0	0	0	0	0	0	0	0	0	0
9	Alice	0	0	0	0	0	1	1	0	0	0
10	Charlie	0	0	1	0	0	0	0	0	0	0

- An adjacency matrix is a case-by-case matrix.
- A binary matrix consists of 1s and 0s.
- The rows in a matrix are represented by i and the columns by j.
- The values of cells are referred to as a (i,j). For example, the cell representing Susie's tie with Emma is a(1,2) = 1.
- Senders are found in rows; Receivers in columns.
- The diagonal of the matrix represents a sender's tie with herself. The diagonal tends to be ignored in SNA.

Figure 1.5 Matrix of a friendship network

The adjacency matrix in the above example records who sends a tie to whom through the use of 1s and 0s. Thus, Susie, in row 1, sends a tie to Emma, in column 2, and this is recorded by inserting a 1 in the cell intersecting row 1 and column 2 (highlighted above). In social network analysis, we use notation to designate which cell we are speaking about: thus, the individual rows in a matrix are referred to as *i*; the columns as *j*; and the value in a particular cell as the letter *a*. Taken together, this notation allows one to specify a particular cell in matrix as a(*i,j*). In the above matrix, we can refer to the cell representing Susie's nomination of Emma as 'a(1,2) = 1'.

Because all the values in the cells are 1's and 0's in Figure 1.5, we call this matrix a *binary adjacency matrix*. A binary matrix contains only 1s and 0s in its cells. As a binary matrix records senders to receivers as rows to columns, the diagonal of this matrix represents the relationship of the sender to itself. In most situations, social network analysts find the diagonal in a matrix uninteresting, as in most social situations, the analyst is not interested in actors' relationships with themselves, but rather the relationships actors have with one another. Thus, the cells along the diagonal are usually recorded as 0s or they are ignored altogether.

The above matrix is also an example of what we call an *asymmetric* matrix. An asymmetric matrix is one that records the direction of ties in a social network. A *symmetric* matrix, by contrast, contains data for an undirected network. This distinction between asymmetric and symmetric matrices is made clearer in Figure 1.6. In the asymmetric matrix, the top right half of the diagonal does not match the bottom left half. In the symmetric matrix, both the upper and lower half of the matrix are the same, as these data are undirected. Another way of thinking about these undirected ties is to see these ties from senders to receivers as being *reciprocal*. For example, Susie nominates Emma as a friend and Emma nominates Susie as a friend.

Asymmetric matrix

		1 S	2 E	3 Z	4 T	5 L
1	Susie	-	1	0	1	1
2	Emma	0	-	1	0	0
3	Zach	0	1	-	0	0
4	Tom	1	0	0	-	0
5	Lynn	1	0	1	0	-

Symmetric matrix

		1 S	2 E	3 Z	4 T	5 L
1	Susie	-	1	0	1	1
2	Emma	1	-	1	0	0
3	Zach	0	1	-	0	1
4	Tom	1	0	0	-	0
5	Lynn	1	0	1	0	-

Figure 1.6 Asymmetric vs symmetric matrices

By contrast, in the asymmetric matrix, the ties may or may not be reciprocal. For example, Susie sends Lynn a tie and Lynn reciprocates that tie, but Susie also sends Emma a tie and Emma does not reciprocate that tie.

In addition to issues of symmetry, and how they reflect the direction of ties, a matrix can also convey the intensity of a tie by the values found within the cell. Thus, for example, a cell containing the value of 4 represents a stronger or more intense tie than a cell containing the value of 3, 2 or 1 (the strength of ties is a topic for the next chapter, and so more discussion will be given at that time). Such a matrix is referred to as a valued matrix or valued network, to convey the fact that values greater than 1s or 0s are also contained in the matrix. Figure 1.7 is an example of a valued matrix.

		1 S	2 E	3 Z	4 T	5 L	6 Z	7 A	8 J	9 A	10 C
1	Susie	0	1	0	3	4	0	0	0	0	0
2	Emma	0	0	3	0	0	0	0	0	0	0
3	Zach	0	4	0	0	0	0	0	0	0	0
4	Tom	3	0	0	0	0	0	0	0	0	1
5	Lynn	3	0	1	0	0	0	0	0	0	0
6	Zoe	0	0	0	1	0	0	1	0	5	0
7	Alex	0	0	0	0	0	1	0	0	5	0
8	Joey	0	0	0	0	0	0	0	0	0	0
9	Alice	0	0	0	0	0	4	4	0	0	0
10	Charlie	0	0	1	0	0	0	0	0	0	0

Figure 1.7 Valued matrix

This valued matrix is based on the same friendship relation discussed earlier, but here, we have added information to the data: here some of these friendship ties are stronger than others. Thus, 1 = a friend or more likely an acquaintance, but 5 = a very close friend. Valued graphs do not necessarily reflect ideas of tie strength or intensity; the values can also represent, for example probabilities (e.g. the probability of a tie being present or absent).

Social network analysts organize social network data into matrices for a number of reasons. Top of the list is that the display of network data in graph form can easily become confusing and even chaotic the larger a network becomes. It is difficult to see how the network is structured and any interesting pattern gets lost. By structuring the data into a matrix, we can then start running a variety of quantitative analyses to start picking out the structural features and overriding patterns in the data.

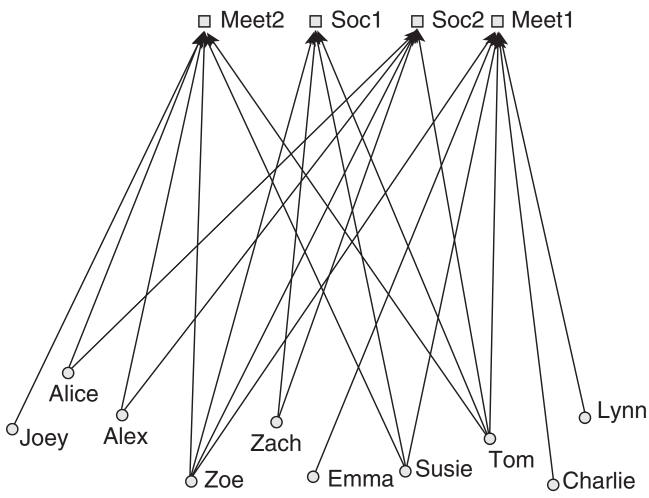
ONE-MODE AND TWO-MODE NETWORKS

In the previous section we looked at matrices of complete networks. The notion of a complete network can be broken down further into the concepts of one-mode networks and two-mode networks. Our previous examples regarding friendship have been examples of one-mode networks. One-mode networks are networks where we study how all actors are tied to one another according to one relation, like friendship. One-mode networks are structured as adjacency matrices, and can be either binary or valued. With two-mode networks, we look at how actors are tied to (or affiliated with) particular events. Examples of two-mode networks include attendance at meetings and membership in an organization. Whereas one-mode networks deal with one set of data (one relation pertaining to one set of actors), two-mode networks deal with two different sets of data. A common example of a two-mode network is one where rows pertain to actors and the columns to events these actors attend. Thus, actors can be tied to one another via events, and similarily, you can conceptualize events as being linked together via actors (more on this dual-portrayal of these data can be found below). Actors and events are only one example of two-mode networks. Another example is that of affiliation networks, i.e. ones where actors are shown in the rows and their affiliations to particular third bodies, e.g. organizations, are depicted in the columns. You could imagine a range of different kinds of two-mode networks; the important distinction here from one-mode networks would be the idea that two separate kinds of entities are being reflected in the matrix, where one set of entities is found in the rows of the matrix, and another in the columns.

The matrices in which two-mode network data are organized are referred to as *incidence matrices*, and when displayed visually, they are referred to as *bipartite graphs*. An example of an affiliation network is given in Figure 1.8, displayed as both a bipartite graph and an incidence matrix.

The digraph and incidence matrix shown in Figure 1.8 show a set of actors and their attendance (or non-attendance) to four separate events. These events are two meetings and two social events. An actor's attendance to one of these events is shown in the incidence matrix through inserting a 1 in the cell below the particular event. Thus, for example, Susie has attended meetings 1 and 2 (M1 and M2) and social event 1 (S1). In the digraph, the actors' attendance is conveyed through arcs sent from the actors to the events.

Bipartite diagraph of actors and events



Incidence matrix of actors and events

		1 M1	2 M2	3 S1	4 S2
		-	-	-	-
1	Susie	1	1	1	0
2	Emma	1	0	0	0
3	Zach	0	0	1	1
4	Tom	1	1	1	1
5	Lynn	1	0	0	0
6	Zoe	1	1	1	1
7	Alex	0	1	0	1
8	Joey	0	1	0	0
9	Alice	0	1	0	1
10	Charlie	1	0	0	0

Figure 1.8 Bipartite diagraph and incidence matrix

From the incidence matrix, we can derive two new adjacency matrices: a case-by-case matrix and an event-by-event matrix. In Figure 1.9 you can see these two new matrices and their respective digraphs.

Figure 1.9 shows how a single incidence matrix can be transformed into two adjacency matrices. Although there exists some special analyses for dealing with two-mode data, in general, most social network analyses are performed on one-mode data, i.e. on adjacency matrices. Thus, it is important to understand how two-mode data can be transformed into one-mode data.

One-mode network (case-by-case)

		1 S	2 E	3 Z	4 T	5 L	6 Z	7 A	8 J	9 A	10 C
		-	-	-	-	-	-	-	-	-	-
1	Susie	3	1	1	3	1	3	1	1	1	1
2	Emma	1	1	0	1	1	1	0	0	0	1
3	Zach	1	0	2	2	0	2	1	0	1	0
4	Tom	3	1	2	4	1	4	2	1	2	1
5	Lynn	1	1	0	1	1	1	0	0	0	1
6	Zoe	3	1	2	4	1	4	2	1	2	1
7	Alex	1	0	1	2	0	2	2	1	2	0
8	Joey	1	0	0	1	0	1	1	1	1	0
9	Alice	1	0	1	2	0	2	2	1	2	0
10	Charlie	1	1	0	1	1	1	0	0	0	1

Two-mode network (affiliation-by-affiliation)

		1 M1	2 M2	3 S1	4 S2
		-	-	-	-
1	Meet1	6	3	3	2
2	Meet2	3	6	3	4
3	Soc1	3	3	4	3
4	Soc2	2	4	3	5

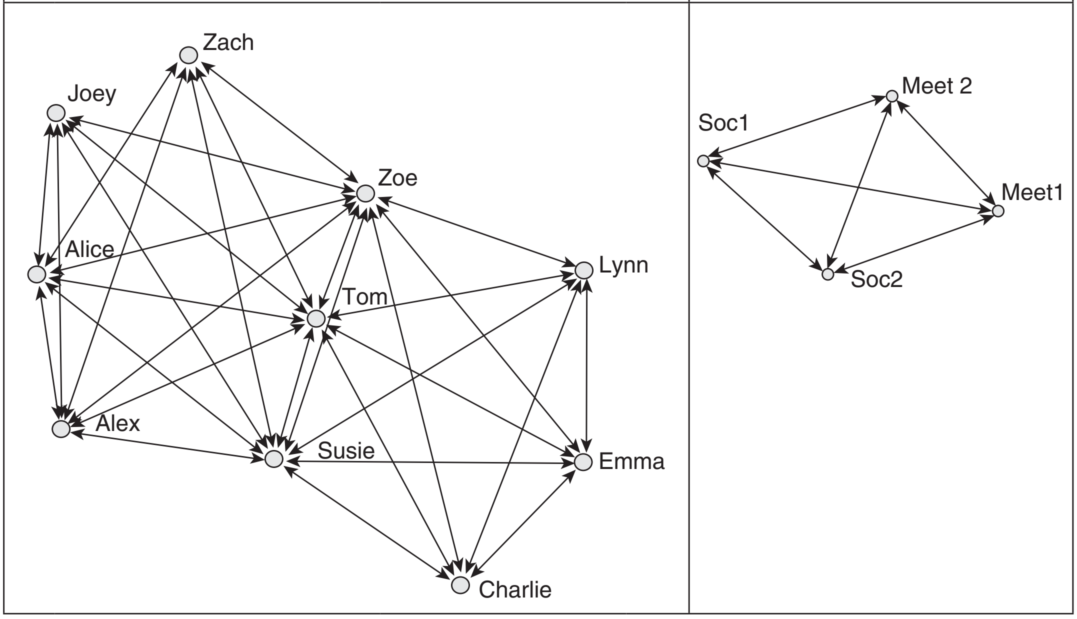


Figure 1.9 One-and two-mode networks

SUMMARY AND CONCLUSION

In this chapter we covered some fundamental concepts regarding social networks and social network analysis. In particular, you learned what comprises a network, namely actors and social ties, the graph theoretic terms for describing those networks, e.g. graphs, nodes and lines, and some basics in structuring the data as matrices and as one-mode or two-mode networks.

In Section 2, you will learn about networks from a 'levels' approach, starting with the 'actor' level and ending with the 'complete network' level. Before doing this, however, I will continue my introduction to social networks by offering you a brief history to the field. This can be found in the next chapter.